Evaluation of the PBF-Funded Project "Countering Hate Speech Through Education and Advocacy for Improving Social Cohesion in Sri Lanka"
(Dec 2021-March 2024)

Final Report

17 May2024

Tammy Smith, Team Leader
Shashik Dhanushka, National Evaluator
Ammaarah Nilafdeen, Research Assistant

Project/outcome Information		
Project/outcome title	Countering Hate Speech Through Education and Advocacy for Improving Social Cohesion in Sri Lanka	
Project Number	00129535 (From MPTFO-Gateway)	
Corporate outcome and output		
Country	Sri Lanka	
Region	Asia and the Pacific	
Date project document signed	1 December 2021	
Project dates	**Start**	**Planned end**
	3 December 2021	1 March 2024
Total committed budget	$3,000,000	
Project expenditure at the time of evaluation	71.4 %	
Funding source	The Secretary-General's Peacebuilding Fund	
Implementing party[1]	UNDP and UNICEF	

Evaluation information		
Evaluation type (project/ outcome/thematic/country programme, etc.)	Project evaluation	
Final/midterm review/ other	Final	
Period under evaluation	**Start**	**End**
	3 December 2021	1 March 2024
Evaluators	Tammy Smith and Shashik Dhanushka	
Evaluator email address	tammyannsmith@gmail.com	shashikdhanushka@gmail.com
Evaluation dates	**Start**	**Completion**
	1 February 2024	1 June 2024

[1] This is the entity that has overall responsibility for implementation of the project (award), effective use of resources and delivery of outputs in the signed project document and workplan.

Table of Contents

Abbreviations and Acronyms

BBS	Bodu Bala Sena
CCA	Common Country Assessment
CEJ	Center for Equality and Justice
CGBV	Cyber Gender-Based Violence
CoP	Community of Practice
COVID-19	Corona Virus 2019
CPA	Center for Policy Alternatives
CRD	Crisis Risk Dashboard
CRPOs	Child Rights Protection Officers
CSO	Civil Society Organization
EWS	Early Warning System
FGDs	Focus Group Discussions
GEWE	Gender Equality and Women's Empowerment
GiZ	German Deutsche Gesellschaft für Internationale Zusammenarbeit
GM	Gender Marker
HACT	Harmonized Approach to Cash Transfers
KII	Key Informant Interviews
LGBTQI+	Lesbian, Gay, Bisexual, Transgender, Queer, Intersex+
LTTE	Liberation Tigers of the Tamil Eelam
M&E	Monitoring & Evaluation
MoV	Means of Verification
MPTF-O	Multi-Partner Trust Fund Office
NCEASL	National Christian Evangelical Alliance Sri Lanka
OHCHR	Office of the High Commissioner for Human
PBF	Peacebuilding Fund
PBSO	Peacebuilding Support Office
PVE	Preventing Violent Extremism
RCO	Resident Coordinator's Office
SDCF	Sustainable Development Cooperation Framework
SDGs	Sustainable Development Goals
SFCG	Search For Common Ground
SGBV	Sexual and Gender Based Violence
ToR	Terms of Reference
UN	United Nations
UNDP	United Nations Development Program
UNFPA	UN Population Fund
UNICEF	UN Children's Fund
YSOs	Youth Service Officers

I. EXECUTIVE SUMMARY

The project "Countering Hate Speech Through Education and Advocacy for Improving Social Cohesion in Sri Lanka" was initiated to address the rising issue of online and offline hate speech, which undermines social cohesion in the country. Funded by the Secretary-General's Peacebuilding Fund (PBF) and implemented by UNDP and UNICEF, the project spanned from December 2021 to March 2024 with a committed budget of $3,000,000. The primary goal was to reduce the prevalence and impact of hate speech through education and advocacy among a wide array of stakeholders, thereby fostering a more cohesive society. This final project evaluation, coming in the last month of project implementation, was commissioned to assess a range of criteria, including relevance, effectiveness, efficiency and impact among others, to determine the efficacy of the project's approach and implementation to address these pressing challenges in Sri Lanka.

The evaluation is positioned as both an accountability measure to the project's donor and stakeholders as well as a means to consolidate learning about what worked well and what could be improved upon. The learning function of the evaluation is particularly important, given that various aspects of the project are novel and have implications beyond Sri Lanka, such as the development of a Code of Practice for social media companies, while other aspects piloted initiatives for which national stakeholders have dedicate plans to scale up, such as digital literacy training within secondary schools.

The project's aim of countering hate speech reflects the need to address longstanding social divisions that have prevailed throughout the country. Sri Lanka has a complex socio-political history marked by ethnic and religious tensions that have periodically resulted in violence and driven social fragmentation. Historical contextual factors, such as the language policy disputes since 1956 and the protracted civil war, have institutionalized divisions among different ethnic groups. More recently, the 2019 Easter Sunday attacks and later the economic and political crises, exacerbated by the COVID-19 pandemic, have intensified social divides and hate speech and highlight the urgent need for interventions to promote social cohesion and counter hate speech

During the project's implementation, the socio-political environment was volatile. The government's mishandling of economic policies as well as public frustration with perceived rampant corruption led to mass protests and the eventual ousting of the Rajapaksa government in July 2022. Notwithstanding a number of impressive achievements of the project, these upheavals, coupled with ongoing issues like fuel scarcity, posed significant challenges to the project's execution, particularly in terms of engaging government and non-government partners.

Evaluation scope, approach and methodology

The evaluation team was tasked with evaluating all aspects of the project, from the project's inception through implementation and achievement of results from the outcome to the impact levels. The ToR presented evaluation questions for eight criteria, including Relevance, Effectiveness, Coherence, Efficiency, Impact, Sustainability, Gender and Human Rights Inclusion, Project Management and Monitoring. Given that this is a peacebuilding project funded by the PBF, the evaluation team added two more criteria, including Conflict Sensitivity and Risk Management, and whether the project evidenced any Catalytic effects.

The project is structured around two outcomes with distinct pathways of work, expected results, and implementing partners and stakeholders. Outcome 1, implemented solely by UNDP with non-financial

support from the Resident Coordinator's Office (RCO) sought to generate much-needed data on patterns of hate speech and encourage national actors to use that data as the basis for evidence-based advocacy for policy change and to counter incidents of hate speech. Outcome 2, on the other hand, was implemented by both UNICEF and UNDP to capacitate an array of actors, including provincial journalists, religious leaders, education professionals, women, secondary students young people engaged in clubs, and both influencers and users of social media on how to detect, avoid and respond to hate speech. The specific outcomes and outputs are included below:

Outcome 1: Systematic Research and evidence provide the foundation for the countering of online hate speech for effective advocacy and preventive action

Output 1.1: Improved access to dynamic data and analytical tools for national stakeholders to systematically track online hate speech, social cohesion indicators, and alongside related SDG 16 targets as a basis for evidence-based advocacy and targeted responses.

Output 1.2: Strengthened civil society capacities for early identification of spread of hate speech, building community resilience and social cohesion and advocacy with social media platforms.

Outcome 2: National stakeholders actively engage in and lead early action to prevent and address hate speech and to support social cohesion through producing online and offline alternative narratives.

Output 2.1: Increased capacity of national stakeholders, including local governance actors, community leaders, civil society and youth to prevent and address hate speech.

Output 2.2: Journalists and other media personnel are trained and capacitated in promoting ethical and more cohesive media and communications for countering hate speech and supporting social cohesion.

Output 2.3: Adolescents and youth have the skills to engage through informal spaces in responding to the drivers of harmful or dangerous speech and conflict.

Output 2.4: Secondary students in targeted schools have the required media and information literacy skills to help prevent and counter hate speech.

In terms of geographic focus, Outcome 1 was mainly implemented through national actors located in Colombo, while Outcome 2 included more diverse targeting, including work with journalists in all nine provinces, as well as advocacy and training of other stakeholders in Colombo, and Northern, Eastern, Uva and Central Provinces. The project also targeted community leaders, schools and clubs in Kandy, Ampara, Batticaloa, Jaffna and Puttalam Districts, as well as two urban centers in Badulla District. Beyond these intentionally targeted lines of work, the project also sought a universal blanketed approach by supporting influencers to generate positive social media content and launching games on social media meant to increase awareness and skill of general users to detect hate speech. As detailed in the evaluation report, the long list of key project stakeholders arrayed across very different interaction spaces presented a challenge to monitoring meaningful project outcomes and negatively affected evaluability of the project.

Given the complexity of the project approach, the methodological approach for the evaluation was multi-method and sought to be gender- and conflict-sensitive. The evaluation team, comprised of a team leader, a national evaluator, and a research assistant, employed a participatory mixed methods approach to ensure a robust assessment of the project's design, implementation, outcomes and impacts. This

approach included content analysis of official documents, semi-structured interviews, focus group discussions (FGDs), and a review of project-specific data collected by implementing partners. The team ensured inclusivity by conducting interviews and discussions in Sinhala, Tamil, and English, depending on the respondents' preferences, and by purposively sampling participants from various stakeholder categories such as UN staff, government counterparts, project beneficiaries, and development partners.

Data collection for the evaluation took place from March 5 to March 31, 2024, across multiple locations including Colombo, Kandy, Hatton, Batticaloa, and Jaffna. The evaluation team reached 272 respondents (146 men and 126 women) through key informant interviews (KIIs) and FGDs. The team used purposive sampling to select participants from six main categories: UN programming and leadership colleagues, project implementing partners, contractors of specific project tasks, government counterparts, representatives from development partners, and direct project stakeholders such as religious leaders, journalists, social media influencers, teachers, students, and youth club members. This approach allowed the evaluation to capture a broad range of perspectives and ensure a balanced representation of opinions across gender, religious, and ethnic lines. Interview and FGD questions and structure were designed to maximize gender sensitivity and inclusivity, including the framing of questions, specific focus on gender-inclusion within our documentary review, and sex-segregated FGDs to provide an open and safe space for young girls to share their experiences and views.

The data collection process was designed to maximize validity and reliability through triangulation of evidence from different sources. The team used data collected from project stakeholders to validate or problematize evidence presented in periodic reports to the donor and other fund recipient-generated claims. Interviews and FGDs were recorded, summarized, and analyzed using Atlas.ti software to identify patterns and emergent themes. The evaluation also included control groups among secondary students and young people in Clubs to compare learning outcomes and better attribute changes to the project's interventions.

Summary of key findings

Relevance. The project’s design was fundamentally aligned with addressing the key factors and actors contributing to social divides and hate speech that undermine social cohesion in Sri Lanka. It remained relevant through a series of crises, including economic turmoil and political upheaval, which highlighted the persistent issue of hate speech as a driver of weak social cohesion. The targeting of beneficiaries, including religious leaders, media professionals, and young people, was highly pertinent, ensuring the engagement of key societal influencers. The alignment of the project with national and UN priorities, such as the Sustainable Development Goals (SDGs) and the UN Sustainable Development Cooperation Framework (SDCF), further reinforced its relevance. For example, the collaboration with the Ministry of Education on curriculum reform and with the Ministry of Mass Media on professionalization initiatives underscored the project's integration into existing national frameworks, enhancing its sustainability and impact.

Effectiveness. The project’s approach demonstrated considerable success in leveraging a systems approach to address hate speech, particularly through initiatives embedded within existing structures like the Ministry of Education's curriculum reforms and the Ministry of Mass Media's professionalization efforts. These measures ensured that project impacts extended beyond individual learning, fostering broader institutional and societal transformations. One of the standout achievements was UNDP's work in generating critical data on hate speech, coupled with support to a Community of Practice that brought key partners together to share information and align advocacy efforts, the project significantly shaped

national strategies and policies. The development and promotion of a voluntary Code of Practice for social media companies, the first of its kind in the Asia-Pacific region, represents a notable success. Additionally, professionalization programs for traditional media, especially among provincial journalists, conducted in partnership with the Ministry of Mass Media, aimed to elevate media standards and reduce incendiary reporting.

The evaluation revealed varying degrees of effectiveness across different stakeholder groups. For religious leaders, the project made important strides in sensitizing them to the negative impacts of hate speech, both online and offline. Workshops conducted under the project enhanced their awareness, encouraging them to fact-check information and reflect on their own communication practices. These workshops also facilitated inter-religious dialogues, breaking down stereotypes and fostering cooperation among leaders of different faiths. Despite these successes, the late mobilization of female religious leaders and the hierarchical barriers within religious institutions posed challenges to broader and more inclusive engagement.

For provincial journalists, the project offered critical training that addressed gaps in knowledge and ethical practices, through training that was based on an independent needs assessment to tailor the approach for journalists in all nine provinces. Many provincial journalists, who had not previously received formal training, reported significant improvements in their understanding of professionalism and ethical obligations. However, the short duration of the training limited its long-term impact, and journalists highlighted the need for ongoing support to embed these practices into their daily work. Additionally, systemic issues such as low wages, lack of resources, and editorial pressures to produce sensationalist content were not fully addressed, hindering the potential for sustained change in media practices.

Youth and Children's Club members, as well as students, were engaged through training programs aimed at improving their understanding of hate speech and promoting digital citizenship. In Northern and Eastern provinces, partnerships with organizations like Sarvodaya facilitated in-depth training over several months, while in Central and Uva provinces, shorter, more concentrated sessions were conducted by local education authorities and Child Rights Protection Officers (CRPOs) and Youth Service Officers (YSOs) trained by T-Field. These efforts increased awareness among young participants, who reported greater sensitivity to the consequences of hate speech and a propensity to fact-check information before sharing it. However, the rushed implementation in some areas and varying degrees of training exposure resulted in inconsistent learning outcomes. While many young people could identify and counter hate speech, there were notable gaps in understanding gender-based hate speech and the broader implications of harmful online behavior. Additionally, small-scale social impact projects undertaken by Club members were often disconnected from the project's primary goals, focusing more on community development rather than directly addressing hate speech and social cohesion.

Aside from these stakeholder specific results, the evaluation found that the independent implementation of various project components without sufficient coordination and connection among project stakeholders and implementing partners led to missed opportunities for greater effectiveness. For instance, creative content produced through print, broadcast, and social media campaigns, while individually compelling, lacked a cohesive strategy to amplify their reach and effectiveness. Furthermore, the short-term nature of many contracts and the rushed implementation timeline limited the potential for sustained impact and thorough monitoring of changes at a meaningful level. This was compounded by the absence of a unifying outcome and clearly articulated indicators, which hindered the project's ability to measure its overall impact on social cohesion effectively.

Impact and Sustainability. The evaluation found mixed results in terms of the project's impact and sustainability. On the positive side, embedding activities within existing systems, such as the Ministry of Education's curriculum reforms and the Ministry of Mass Media's professionalization initiatives, has shown indications that these workstreams may be highly impactful. These efforts have facilitated a nationwide approach to tackling hate speech, providing the best chance that the project's impacts can be sustained over time. The systemic approach taken by the project, particularly through partnerships with government ministries and the establishment of a Community of Practice (CoP), has shown early signs of generating meaningful and lasting change. Additionally, the work through UNDP and the Resident Coordinator's Office (RCO) in generating critical data on hate speech has contributed to shaping national strategies and policies, supporting the creation of a voluntary Code of Practice for social media companies, which will be the first of its kind in the Asia-Pacific region once fully adopted.

Despite these positive developments, the project has also faced substantial challenges that have hindered its overall impact and sustainability. One major barrier has been the fragmented implementation of various project components, which lacked sufficient coordination and coherence. This disjointed approach has limited the project's ability to achieve collective outcomes and has raised questions about the relevance of certain activities, such as the numerous small-scale social impact projects implemented by Children's and Youth Clubs. While these projects allowed participants to practice leadership and planning skills, they were largely disconnected from the broader project aims of reducing hate speech and promoting social cohesion. Furthermore, the absence of robust monitoring tools to capture specific learning or changes in attitudes and behaviors has limited the evaluability of the project's impacts.

In terms of sustainability, the project has laid a strong foundation through its work with existing institutions, but several areas require further attention to ensure long-term success. For instance, the professionalization programs for traditional media, while promising, need ongoing support not only to solidify learning outcomes but to address market incentives that drive media outlets to publish inflammatory content. The success of the voluntary Code of Practice for social media companies as well as A code under development for mass media companies will depend on close monitoring and the willingness of these companies to adhere to ethical standards. Future initiatives should allow for longer-term engagements and more comprehensive training to ensure that changes in attitudes and behaviors are deeply rooted and sustained over time.

Finally, the project's impact on social cohesion has been mixed. While there have been notable successes in raising awareness and promoting positive narratives among young people, religious leaders, and media professionals, the short duration of the project and the limited exposure to training for many participants have impeded deeper and more widespread changes over the implementing timeframe. The evaluation also highlighted the need for a more focused approach to addressing gender-based hate speech, including cyber gender-based violence, as well as a stronger emphasis on engaging religious leaders who can play a crucial role in fostering social cohesion. Strengthening these areas and ensuring better coordination among different project components will be essential for achieving more profound and lasting impacts in future initiatives.

Catalytic Effects. Financially, the project catalyzed $207,500 in additional funding, including $16,500 from the Office of the High Commissioner for Human Rights (OHCHR) for monitoring discrimination against minority groups and $191,000 from the United States Department of State for augmented data collection and monitoring from 2022 to 2024. Programmatically, the project influenced strategic frameworks and policy changes within the UN system, particularly shaping the SDCF and UNDP's SDG 16 Flagship Program. These frameworks now prioritize addressing hate speech as a key assumption driving efforts to enhance

social cohesion in Sri Lanka. Additionally, the Department of Education in Central Province plans to utilize its own resources to partner with the University of Peradeniya to refine curricula and piloting an intensive course on good citizenship and combating hate speech, an indication of strong local ownership that was unanticipated by the project design team.

Efficiency. The project's innovative approach in leveraging existing systems to deliver outputs efficiently. For instance, the project's Steering Committee utilizes existing structures to ensure cohesion with broader efforts and reduce duplication of oversight structures. The collaboration with the Ministry of Mass Media for professionalization programs and the Ministry of Education for curriculum reforms, moreover, ensured that the project's initiatives were integrated within established structures, thus optimizing resource use and fostering sustainability. The project's emphasis on using data-driven approaches, such as the systematic monitoring of social media platforms by Hashtag Generation, also exemplified strategic resource allocation that facilitated timely responses to emerging hate speech incidents.

Despite these strengths, the project encountered challenges that impeded its overall efficiency. While some of these challenges reflect the difficult operational environment in which the project was implemented, others are a result of programming choices by UNDP and UNICEF. The high volume of relatively short-term contracts issued to implementing partners, for example, was a major constraint, with an average contract length of just 7.2 months. This short-term nature led to inefficiencies, as partners had to invest time in understanding the project's context and establishing their operations only to wind down shortly after making initial progress. Additionally, the independent implementation of various project components without sufficient coordination resulted in missed opportunities for synergies and collective impact. For example, creative content produced through print, broadcast, and social media campaigns ran on separate timelines and lacked a cohesive strategy to amplify their reach and effectiveness.

Furthermore, the project's budget allocation revealed imbalances that affected its efficiency. A substantial portion of the budget was dedicated to non-programmatic activities and resources, such as monitoring, visibility, and documentation, which, while important, reduced the funds available for substantive programming. Implementing partners reportedly had to cover budget gaps using their own limited resources, which strained their capacity and hindered the overall effectiveness of the interventions. Moreover, the project's lack of integration and joint implementation across different outcomes further reduced programmatic coherence and efficiency, as partners often operated in isolation without benefiting from shared knowledge and resources.

Gender Responsiveness and Human Rights

The project's findings on gender responsiveness and human rights highlight both successes and areas for improvement. Positively, the project made significant strides in including traditionally excluded groups in its activities. For example, work by UNDP and the RCO, filled evidence gaps that enabled national stakeholders to address hate speech, including gendered hate speech. Reports by organizations such as Hashtag Generation, Factum, and Deloitte revealed increasing hate speech against women and the LGBTQI+ community, who were targeted in 53% of hate speech incidents monitored between January and April 2023, up from the 25% noted in 2021. These findings enabled stakeholders to advocate for gender provisions within a revised Code of Practice for media and tech companies, which, if fully adopted, could have far-reaching effects. Furthermore, training provided through the Center for Equality and Justice

supported 30 young social media content producers, including women and LGBTQI+ individuals, resulting in impactful online campaigns against cyber gender-based violence.

Despite intentions to ensure broad inclusivity, some project aspects, such as the support for religious leaders, faced significant barriers in achieving the project's aspiration of advancing gender equality and women's empowerment. At the time of the evaluation, financial reports from 2023 revealed that only $45,000 of the allocated $150,000 was expended in this programmatic output, and UNICEF faced significant barriers to include female religious leaders in most workshops and training. Additionally, the needs assessment of provincial journalists conducted by Verité did not consider the specific barriers faced by female journalists, resulting in recommendations that were not tailored to their needs. The project's results framework was also critiqued for lacking clarity, deficiencies in sex-disaggregated data, and stretching definitions to meet gender-based targets. Financial execution rates for gender equality and women's empowerment (GEWE), moreover, fell below the required 30%, leading to a potential downgrade of the project's gender marker rating from gender marker (GM) 2 to GM1, despite the project's substantive positive impacts in some areas.

Conflict Sensitivity and Risk Management

The project demonstrated a strong awareness of inter-communal differences and context-specific triggers of tension. Positive findings include the project's sensitive handling of religious considerations that stakeholders largely appreciated. Additionally, UNICEF successfully integrated key concepts of addressing hate speech into the Ministry of Education's curriculum reforms by taking a sensitive and cautious approach to language to build acceptance and national ownership. These efforts reflect a nuanced understanding of the complex social fabric and an adaptive approach to avoid exacerbating existing tensions.

However, the evaluation also highlighted areas where the project's conflict sensitivity and risk management could be improved. The decision not to openly address caste discrimination, due to fears of provoking tensions, was seen as a necessary but limited approach that may need revisiting in the long term, given regional and intra-communal differences regarding caste-based hate speech. Additionally, the project faced challenges in managing risks associated with the participation of state actors. For example, some journalists were suspicious of government-facilitated training, which undermined certain aspects of the training's effectiveness. Moreover, the presence of government-affiliated individuals in reporting networks facilitated by CPA created concerns over confidentiality and safety among participants. These findings suggest the need for a more strategic approach in addressing such sensitivities and ensuring that risk management measures effectively offer a sense of protection to all stakeholders.

Coherence

The establishment of the Community of Practice (CoP) under Outcome 1 was a notable success that not only avoided potential overlap with an existing CoP but helped foster coordinated approaches among CoP members. The CoP became a dynamic mechanism for sharing data, analysis, and coordinating advocacy and action, growing its membership from 14 to 22 active members. UNDP's strong support to the CoP enhanced its institutionalization by facilitating regular meetings, document sharing, and collaborative efforts, thereby enhancing coherence with external actors and driving collective efforts to counter hate speech. Additionally, the data generated through social media monitoring and incident

reporting contributed significantly to sensitizing development partners to emergent problems and informed their policy and funding decisions.

Despite these achievements, the lack of operational and substantive interaction between Outcome 1 and Outcome 2 significantly reduced programmatic coherence. The two project outcomes were implemented largely independently, falling short of the true spirit of joint programming. This disjointed implementation led to missed opportunities for project stakeholders to amplify each other's messages or benefit from creative outputs produced within the project. Implementing partners often lacked awareness of the project components managed by others, resulting in inefficiencies and duplicated efforts. The absence of a unifying outcome and clearly articulated indicators further compounded these issues, making it challenging to measure the project's overall impact on social cohesion effectively.

Project Management and Monitoring

Echoing findings from an earlier, similar project, the evaluation found that the project's results framework was not adequately designed to measure higher-order changes, which negatively impacted project cohesion and evaluability. The independent functioning of the project's two main outcomes led to missed opportunities for synergy and unified impact. The lack of a unifying outcome or strategic impact statement meant that each outcome remained as an isolated intervention, which hindered the ability to measure the project's meaningful contributions to social cohesion. For future projects, it is recommended to set a strategic outcome above the individual outcomes to facilitate tracking the overall impact of efforts aimed at reducing hate speech and enhancing social cohesion.

Another significant challenge was the high volume of relatively short-term contracts issued to implementing partners, which contributed to inefficiencies and constrained the ability to achieve long-term impact. Many partners expressed frustration that their work had just begun to yield meaningful results when the contracts ended, leading to a rush to deliver within limited timeframes. This was further exacerbated by UNDP's internal review processes to approve contracts and deliverables such as social media content, which delayed the launch of time-sensitive content for some implementing partners. Additionally, the project suffered from a lack of operational and substantive interaction between UNDP and UNICEF, which further reduced programmatic coherence. Implementing partners often worked in isolation, unaware of other project components, leading to missed opportunities for knowledge sharing and resource optimization.

Recommendations

Recommendation 1: Focus on Gender-Based Hate Speech. Future projects designed by the UN Country Team should place a more significant focus on gender-based hate speech, including cyber gender-based violence (CGBV), or more comprehensively integrate gender responsiveness into other approaches. The consistent and overwhelming incidence of hate speech against women and LGBTQI+ individuals underscores the need for a dedicated gender-responsive approach. The project has already demonstrated productive results in this area with modest budgets and short-term contracts, indicating high relevance and demand from national stakeholders. Harnessing partnerships not only with women and the LGBTQI+ community but also with allied men could strengthen these efforts.

Recommendation 2: Engage Religious Leaders Intensively. The evaluation highlights the importance of religious leaders as key influencers capable of fostering social cohesion within their communities. Future initiatives should design dedicated workstreams to engage religious leaders across different faiths. This

approach should consider the diverse organizational structures, cultural-religious tenets, and language barriers among religious groups. The project should also include efforts to involve female religious leaders, despite the inherent challenges. Strengthening intra-religious bonds and promoting inter-religious understanding through targeted initiatives can significantly advance the project's goals.

Recommendation 3: Broaden Access to Conflict Data Platforms. The Crisis Risk Dashboard (CRD) and Early Warning System (EWS) are innovative tools that could be better utilized if access is broadened to a wider array of activists and policy actors. This inclusion would enhance the platforms' sensitivity to emergent conflict drivers and ensure their relevance. If making the existing platforms more widely available is not feasible, the creation of an external-facing platform using select data should be explored. Additionally, systems should be established to better track the usage of this information to make informed decisions about its continued usefulness and attract financial support.

Recommendation 4: Develop a Sustainable Plan for the Community of Practice (CoP). UNDP should support the CoP in developing a one- to two-year plan to become self-sustaining. Current levels of support from UNDP are not sustainable. National actors should take ownership of the CoP and determine its future role. UNDP can assist in planning for a gradual reduction of support, offering capacity building if needed and establishing clear benchmarks for tracking progress towards independence.

Recommendation 5: Capitalize on Creative Content for Training and Education. UNDP and UNICEF should capitalize on and repurpose the creative content produced during the project, such as social media campaigns, commercials, and games. This content can be used to invigorate training and education efforts with a variety of stakeholders, including journalists, students, and community members. Specifically, identifying and utilizing outputs like Moonshot's game or Time and Space's commercials for training in Children and Youth Clubs can demonstrate value-for-money and responsiveness to stakeholders' learning preferences.

Recommendation 6: Integrate Provincial Initiatives into National Curriculum Reforms. As UNICEF anticipates the rollout of the Ministry of Education's curriculum reforms, it should ensure that provincial-level education initiatives addressing hate speech are incorporated into the national plan. The pilot in Central Province offers an opportunity to inform national efforts. Joint learning sessions between national and provincial actors should be planned to integrate effective pedagogical approaches and content, ensuring that local ownership and drive are maintained without conflicting with national reforms.

Recommendation 7: Adopt a Sustained Approach to Training. Training efforts need a more sustained approach, with ongoing opportunities for master trainers and participants to share experiences and reinforce learning. Short-term sessions are insufficient to change deeply entrenched behaviors. A series of training and learning opportunities over months or years should be planned, focusing on consistent, long-term engagement. Special attention should be given to training teachers, CRPOs, and YSOs, ensuring they have access to frequent training and structured opportunities for experience sharing.

Recommendation 8: Continue Professionalization of Journalists. UNDP and the Ministry of Mass Media should continue to offer training and professionalization opportunities to journalists, including provincial journalists. Future training should consider the specific needs and barriers of female journalists. Additionally, addressing industry incentives that drive the demand for inflammatory content is crucial. A holistic approach involving all media professionals, including company owners and managers, can promote ethical, fact-based journalism.

Recommendation 9: Promote Cross-Community Interactions Among Youth. UNICEF should look for opportunities to promote cross-school or cross-club sharing and joint events to reduce social separation experienced by young people. Twinning schools or clubs from different ethno-religious communities and organizing joint activities such as study visits or sporting events can help break down social barriers. These efforts should be embedded in broader initiatives that support long-term personal growth and relationship development among young people.

These recommendations, if implemented, can significantly enhance the effectiveness and sustainability of efforts to combat hate speech and promote social cohesion in Sri Lanka. By focusing on systemic approaches, engaging key stakeholders, and ensuring long-term support for training and education, future projects can build on the foundations laid by this initiative and achieve more profound and lasting impacts.

Conclusion

The project "Combating Hate Speech and Promoting Social Cohesion in Sri Lanka" has made significant strides in addressing the factors contributing to hate speech and social division. By mobilizing a diverse range of stakeholders, including media professionals, religious leaders, and youth, the project has fostered promising approaches to promoting a more cohesive society. The approach, particularly through partnerships with government ministries and the creation of a Community of Practice (CoP), shows early signs of generating sustained, impactful change. However, the evaluation highlights areas needing improvement, such as better coordination across project components, more effective monitoring of change, and ensuring long-term sustainability.

The evaluation revealed that the project was most effective when it worked through existing systems, like the Ministry of Education's curriculum reforms and the Ministry of Mass Media's professionalization initiatives. Embedding the work within these national institutions provides the best opportunity for nationwide, enduring change. These efforts, moreover, aim to ensure that the project's impacts extend beyond individual learning to broader institutional and societal transformations. Despite these achievements, the evaluation noted that the lack of a unifying outcome and clear indicators hindered the project's ability to provide convincing evidence of its overall impact on social cohesion. Future initiatives should aim for a more cohesive integration of these elements and ensure an approach to measurement that captures whether these project pathways elicit the desired or expected change.

Other notable successes include the RCO and UNDP's support for data generation on hate speech, which shaped national strategies and policies, and the development of a voluntary Code of Practice for social media companies, the first in the Asia-Pacific region. Additionally, UNICEF's collaboration with the Ministry of Education on curriculum reforms, noted above, has integrated concepts of digital citizenship, critical thinking, and inter-group tolerance into mandatory coursework, significantly impacting students in the formal education system. Training programs for young social media content producers and traditional media journalists have introduced these stakeholders with the skills needed to recognize and counter hate speech, resulting in impactful media content and campaigns. However, better coordination and continued capacity building are necessary to sustain these achievements.

The evaluation of the project's gender responsiveness highlighted several successes but also areas for improvement. Positively, the project managed to include traditionally excluded groups, with evidence from reports by organizations like Hashtag Generation and Deloitte revealing a rise in hate speech incidents targeting women and the LGBTQI+ community from 25% in 2021 to 53% in early 2023. This data

facilitated advocacy for gender provisions within a revised Code of Practice for media and tech companies and resulted in impactful online campaigns against cyber gender-based violence led by social media content producers trained and supported by the project. However, challenges included underutilization of funds for supporting female religious leaders and gender-blind approaches in certain assessments and trainings. Given these challenges, financial execution rates for gender equality and women's empowerment activities fell below the required 30%, potentially downgrading the project's gender marker rating from GM2 to GM1, despite substantive positive impacts in some areas. UNDP and UNICEF are urged to consider a more prominent and intentional focus on addressing gender-based hate speech in future efforts.

II. INTRODUCTION

II.A. The context of Sri Lanka during project design and implementation

While the phenomenon of online hate speech and its impact on social cohesion in Sri Lanka is relatively new, the basis for the various forms of divisive narratives have their roots in institutionalized societal divisions, some of which stem from the earliest days of Sri Lankan independence. As such, contemporary forms of online hate speech reflect long standing patterns of division, resentment, and "othering" in society. To understand hate speech, then, one must situate the various ways – both online and offline – that patterns of exclusion and intolerance have been produced and reproduced over time.

As early as 1956, in the contest to define Sri Lankan post-colonial identity, tensions and occasional violence was triggered by the debate over establishing Sri Lanka's official language(s), with Sinhala eventually prevailing despite heavy opposition by Tamil politicians who urged for recognition of both languages. These decisions created a spiral down of tit-for-tat actions with far-reaching effects that resulted in polarization between the communities.[2] Subsequent attempts by then-Prime minister S.W.R.D. Bandaranaike, a Sinhalese politician, and the prominent Tamil politician S.J.V. Chelvanayakam to broker a compromise triggered communal riots that resulted in loss of life and property damage. Under the weight of such extreme responses, support for a compromise was eroded and the initiative was abandoned.

Recurring communal tension and violence culminated in the Civil War from 1983 to 2009, between Sri Lanka's military and the Liberation Tigers of the Tamil Eelam (LTTE). These events, and the purported violence against civilians that marked the end of the war, have become touchstones for fomenting hate and furthering division across generations.[3]

Following the defeat of the LTTE, the nature of social division shifted from a fractured relationship between specific ethnic groups and the state to tensions between ethnic or communal groups. Buddhist extremist groups like the Bodu Bala Sena (BBS), founded in 2012, and increasingly fractious political discourse played a significant role in propagating hate speech, frequently fueled by false claims, that challenged inter-ethnic and inter-religious tolerance and social cohesion. Distrust and popular prejudice toward the Muslim community, the second-largest minority in the country grew in this period, fed by fake news spread through both online and offline channels. Other groups such as Christian religions and sects also experienced an uptick in hate speech as communal polarization deepened in the country.[4]

Tensions came to a head in 2019, with the coordinated Easter Sunday attacks on churches and luxury hotels by Islamic extremist groups, in which 271 people were killed and more than 300 were injured. These attacks occurred against a backdrop of existing anti-Muslim prejudice. The Easter Sunday attack affirmed anti-Muslim prejudices and sparked riots in the North Western Provinces, leading to widespread

[2] The shift to Sinhala, for example, prompted a change on vehicle license plates from English to Sinhala. In response, a prominent Tamil activist similarly changed his vehicle's number plate to Tamil, enraging Sinhalese groups in the south, who launched a campaign to erase Tamil names from public signs in the south.

[3] According to the Report of the Secretary-General's Panel of Expert on Accountability in Sri Lanka, as many as 40,000 civilians were killed in the final phase of war in 2009.

[4] According to the Centre for Policy Alternatives (CPA), 65 attacks on places of worship took place from 2009 to 2012, while the National Christian Evangelical Alliance reported 103 incidents of violence and intimidation against the Christian community from 2009 to 2014. According to the Sri Lanka Muslim Congress, there were 241 Anti-Muslim attacks in 2013 alone, with the majority (148) being hate or propaganda campaigns.

property damage and one death. Courageous efforts by religious and community leaders helped stop the violence, although widespread reports of discriminatory actions and fake news meant to create outrage or separation with the Muslim community persisted.

In the wake of the Easter Sunday attack, Gotabaya Rajapaksa, who had served as Secretary of Defense during the government of his brother, Mahinda Rajapaksa, was elected President in autumn 2019. His inauguration at the Ruwanwelisaya Stupa, a prominent sacred site for Sinhalese Buddhists, signaled to minority communities his pro-Sinhalese position and forewarned a worsening of inter-communal relations. Subsequent legislative proposals and Presidential actions favoring Sinhalese Buddhist nationalism triggered a spiral down in communal relations,[5] and prompted a report by the United Nations Human Rights Commission in January 2021 warning about the failures to address past rights abuses and increasing militarization, ethno-religious nationalism and intimidation of civil society actors.[6]

In 2022, deepening public discontent with power cuts, corruption, economic hardship, and democratic backsliding – among other factors – triggered a wave of popular protests. Demonstrators from diverse backgrounds took to the streets, demanding accountability and political change. Despite attempts by the government to suppress dissent, the protests persisted, fueled by a growing sense of disillusionment with the ruling regime. As pressure mounted both domestically and internationally, culminating in widespread calls for change, President Gotabaya Rajapaksa was compelled to resign, paving the way for a transition of power.

The Role of Hate Speech in Perpetuating Social Division

Since the onset of the COVID-19 pandemic in 2020, Sri Lanka has witnessed a surge in online hate speech, exacerbating inter-communal tensions and contributing to a climate of hostility among different ethnic and religious groups. While patterns of hate speech against one or another group have risen and fallen over time, the spread of COVID-19-related conspiracy theories and misinformation online, which often targeted minority communities, particularly Muslims, is emblematic of the introduction and spread of such divisive discourse. False narratives linking the pandemic to religious practices, such as practices connected to Islamic burial rights and communal gatherings, fueled anti-Muslim sentiment and incited violence against the Muslim community. Young people, who are frequent users of social media platforms, were said to be particularly susceptible to these narratives, making them vulnerable to radicalization and perpetuating inter-communal tensions.

Fueled by disinformation, ultra-nationalist rhetoric and ethnonationalist movements have gained traction online since the pandemic and social media platforms have become breeding grounds for extremist views, with individuals and groups exploiting fears and insecurities exacerbated by the pandemic to promote exclusivist ideologies and stigmatize minority communities. Women, in particular, face a multitude of targeted harassment and intimidation online, abuse that can be compounded by their ethnic or religious identities.[7] Muslim women, for example, are frequently targeted with hate speech for wearing the burqa and face restricted access to public services like schools and hospitals for veiling. Within the Sinhalese community, female activists and politicians face backlash and hate speech for challenging

[5] Neil DeVotta, "Sri Lanka: the Return to Ethnocracy," Journal of Democracy 32:1 (2021)

[6] The UNDP Crisis Risk Dashboard, for example, recorded 78 incidents related to human rights violations from March 2020 to October 2021, with most violations stemming from arbitrary arrests under the Prevention of Terrorism Act of 1979, mainly in the Batticaloa, Jaffna, and Colombo districts. Increased use of the PTA has been noted since the Easter attack, with the arrest of several noted authors and scholars based on their publications promoting extremist ideologies.

[7] Approximately 25 percent of online hate speech targets women.

traditional gender norms and virtues. The normalization of hate speech online has not only deepened inter-communal tensions but also eroded democratic norms and values, undermining efforts towards reconciliation and social cohesion.

Despite these challenges, civil society organizations, youth groups, and women's networks have been seeking to leverage social media platforms to amplify voices of moderation and tolerance, challenging hate speech and advocating for inclusive and pluralistic societies. By providing alternative narratives and spaces for constructive dialogue, these initiatives aim to empower young people and women to resist radicalization and build bridges across ethnic and religious divides.

II.B. Description of the project

"Countering Hate Speech" is a $3 million joint project between UNDP and UNICEF that was funded by the UN Secretary-General's Peacebuilding Fund (PBF) and implemented from 3 December 2021 to 1 March 2024, a period that witnessed several crises, as noted above. The project's original end date was 1 June 2023, but, citing delays due to disruptions to implementation caused by both the COVID-19 pandemic as well as the political crisis, the project received a no-cost extension that pushed the end date to 1 March 2024.[8]

As stated in the project document, the overall aim of the project is to strengthen social cohesion through building safer and more inclusive spaces, both online and offline, by supporting national partners to address online hate speech. Within this overall goal, the project also expects to advance protection of human rights and boost women's civic participation.[9]

The rationale for the project recognizes that hate speech, disinformation and divisive narratives that have circulated online in Sri Lanka have exacerbated existing societal divisions and contributed to a number of violent incidents in the past. Notable incidents include communal riots in 2018 in Kandy District and intra-communal riots in North Western Province based on anti-Muslim hate speech that circulated on social media after the 2019 Easter Sunday suicide attacks by Islamic extremists. Aside from these extreme cases, the project document highlights worrying trends in society more generally, including an uptick of 168 per cent in gender-based hate speech in the Asia Pacific region from 2019 to 2020, following socio-economic strain and breakdown in social services caused by the COVID-19 pandemic and an economic crisis. In Sri Lanka, an estimated 200 messages tagged as hate speech are said to circulate monthly on Facebook, alone. Given that this estimate does not include the influence of popular messaging platforms such as WhatsApp, Viber or Telegram, it is assumed that such instances are undercounted.[10]

Another aspect of the project's approach works with and through children, adolescents and young people to improve their critical thinking skills and digital media literacy to break the generational cycle of hate speech promotion. In addition to these youth- and child-focused activities, the project also launched a set of activities promoting digital storytelling and "gamification," and encouraging digital literacy and critical thinking among social media users to create resilience against online mis/disinformation and hate

[8] PBF/IRF-427, Hate Speech Prevention No Cost Extension, July 2023

[9] PBF/IRF-427: "Countering Hate Speech through education and advocacy for improving social cohesion in Sri Lanka," Prodoc, p. 2.

[10] PBF/IRF-427: "Countering Hate Speech through education and advocacy for improving social cohesion in Sri Lanka," Prodoc, pp. 5-6.

speech. Project partners consider these aspects of the project to be part of an innovative pilot that may be scaled up later by government partners if shown to be effective.

In addition to work with children and young people in schools and clubs, the project also sought to sensitize staff in schools, journalists, social media influencers and religious leaders on the negative impact of hate speech and build their skills in identifying and countering it. This aspect of the project was complemented by additional project activities focused on creating a more robust regulatory environment and empowering civil society actors to prevent online hate speech and more effectively respond to it when it does occur.

In addition, the project recognizes that an estimated 25 percent of hate speech in Sri Lanka is gender-based, including frequent hate speech toward young women, women candidates during election cycles, and targeted hate speech against those identifying as LGBTQI+.[11] In response, the project has dedicated 30 percent of its budget to Gender Equality and Women's Empowerment (GEWE) to address gender-based hate speech through targeted actions with women's groups and its work with youth.

To achieve these aims, the project is structured around two interlinking outcomes with underlying outputs:

Outcome 1: Systematic Research and evidence provide the foundation for the countering of online hate speech for effective advocacy and preventive action

Output 1.1: Improved access to dynamic data and analytical tools for national stakeholders to systematically track online hate speech, social cohesion indicators, and alongside related SDG 16 targets as a basis for evidence-based advocacy and targeted responses.

Output 1.2: Strengthened civil society capacities for early identification of spread of hate speech, building community resilience and social cohesion and advocacy with social media platforms.

Outcome 2: National stakeholders actively engage in and lead early action to prevent and address hate speech and to support social cohesion through producing online and offline alternative narratives.

Output 2.1: Increased capacity of national stakeholders, including local governance actors, community leaders, civil society and youth to prevent and address hate speech.

Output 2.2: Journalists and other media personnel are trained and capacitated in promoting ethical and more cohesive media and communications for countering hate speech and supporting social cohesion.

Output 2.3: Adolescents and youth have the skills to engage through informal spaces in responding to the drivers of harmful or dangerous speech and conflict.

Output 2.4: Secondary students in targeted schools have the required media and information literacy skills to help prevent and counter hate speech.

In terms of geographic focus, certain aspects of the project were aimed at the national level, in all nine provinces, other initiatives were implemented through more targeted actions in Colombo, as well as in the Northern, Eastern, Uva and Central Provinces. The project also targeted community leaders and

[11] PBF/IRF-427: "Countering Hate Speech through education and advocacy for improving social cohesion in Sri Lanka," Prodoc, p. 7.

school partners in Kandy, Ampara, Batticaloa, Jaffna and Puttalam Districts, as well as two urban centers in Badulla District. According to the project document, these project sites were selected for their multi-ethnic or multi-religious composition. In addition, the project team sought regional balance and included communities that had experienced negative incidents based on mobilization through online hate speech.[12]

Analyzing the project's Theory of Change (ToC)

The project document includes an extensive presentation of many different aspects of the project, focused mainly at the output level. The project document, however, also implicitly lays out core assumptions that guide the overall approach. This section will examine these core assumptions and probe the underpinning evidence base.

As stated in the project document, one core assumption is that past experience[13] demonstrates the value in supporting advocacy and initiatives that promote better-moderated, more inclusive, safer online spaces that feature more engaging positive content that can act as a counter to divisive narratives.[14] If we delve a bit deeper, two underlying assumptions emerge: 1) if people had access to accurate information and more positive narratives, they would reject negative ones, and 2) if the circulation of negative narratives was reduced or eliminated, people would be more open to positive messages and less apt to mobilize on the basis of divisive narratives or hate speech. The Evaluation Team will test whether these assumptions are borne out by evidence throughout implementation of the project.

A second core assumption states that evidence shows that by increasing the skills of journalists, educators, young people and women through education on digital media and critical thinking they will improve their understanding of hateful narratives and disinformation.[15] At heart, this approach assumes that people's ability to counter online hate speech is determined by their level of knowledge/understanding and skill, and that by providing access to learning and skills building, the project can help these stakeholders effectively counter online hate speech. This is a rather important assumption about how change is likely to work, and one that will be tested by the Evaluation Team.

In addition to these two sets of core assumptions, two additional underlying assumptions are implied but not explicitly stated in the project document. First, as the title and description of the project make clear, the aim of the initiative is to strengthen social cohesion. To address this level of change, the ToC should make clear that one of its fundamental assumptions is that reducing the appeal of online hate speech and promoting positive narratives can have a positive effect on social cohesion. The logic behind this change is that hate speech promotes social division that can stoke negative actions such as communal riots, discrimination and other social ills that contribute to a spiral downward; by removing the hate speech, one removes a contributor to social division that undermines social cohesion. As noted on page 8 of the project document, however, hate speech itself is a "threat multiplier" that builds off existing divisions. These existing divisions are produced and reproduced by formal and informal institutions in the social, cultural, religious, economic and political spheres that are for the most part not addressed by the project.

[12] PBF/IRF-427: "Countering Hate Speech through education and advocacy for improving social cohesion in Sri Lanka," Prodoc, p. 21.

[13] Although the project document references past experience, it does not detail what the experience is or the evidence that this experience generated. Thus, it is challenging for the Evaluation Team to determine upon what evidence base the assumptions have been developed.

[14] PBF/IRF-427: "Countering Hate Speech through education and advocacy for improving social cohesion in Sri Lanka," Prodoc, p. 26.

[15] PBF/IRF-427: "Countering Hate Speech through education and advocacy for improving social cohesion in Sri Lanka," Prodoc, p. 26.

The question remains whether by treating hate speech, which is a byproduct of the ways in which institutions foment division, the project will be able to overcome the more institutionally embedded ways that social divisions are maintained and even rewarded in Sri Lanka. This is an aspect of the project that should be closely monitored and has informed the Evaluation Team's thinking about the project's impact.

II.C. Description of the Evaluation Process and Methodology

II. C.1. Evaluation Criteria, Design and Methodology

A team of two evaluators and one research assistant conducted this evaluation between February and April 2024. The evaluation team was mandated to evaluate a broad range of ten evaluation criteria, including Relevance, Effectiveness, Coherence, Efficiency, Impact, Sustainability, Gender and Human Rights Inclusion, Project Management and Monitoring, Conflict Sensitivity and whether the project evidenced any Catalytic effects. The Terms of Reference (ToR) for this evaluation is attached to this report as Annex A.

The evaluation employed, to the greatest extent possible, a participatory approach whereby interviews and focus group discussions with key stakeholders provide and verify the substance of the findings. Given that many aspects of the project will continue after project closure, the evaluation team expects that a participatory approach will support validity of findings and greater stakeholder ownership with respect to recommendations. The types of change anticipated by the project, moreover, call for evaluation methods that directly engage stakeholders in order to determine whether new skills and knowledge have been imparted as a result of the project activities, and if so, whether new skills and knowledge have had the anticipated effect of reducing hate speech or encouraging positive narratives.

Beyond prioritizing stakeholder participation, the evaluation relies on a multi-method approach to maximize the validity and reliability of evaluation findings and to afford the team opportunities to triangulate its evidence. This is especially important given that many outcome indicators measure only the number of things – such as the number of students who were exposed to training – rather than any substantive change in their attitudes or behaviors, let alone the hoped for effect these types of changes might have on social cohesion within their environment. By launching a multi-method approach, the team aimed to go beyond validating project reporting, for example, on how many stakeholders were trained to assess to what degree this training was relevant and effective.

The evaluation's overall findings are based on the integration of content analysis of official United Nations documents and literature from external partners, semi-structured interviews conducted virtually and in person, as well as focus group discussions conducted in person.

As a first step, the team conducted an extensive document review, including: analyses of project documents, extension requests and narrative and financial reports; coordination meetings; project-related data collection instruments and outputs; strategic plans and relevant government policies and plans; and related studies, evaluations and other literature connected to contextual and thematic background, including conflict analyses. A list of consulted documents is included in Annex B. The document review enabled the team to generate a picture of project implementation inputs, outputs and challenges. Given that extensive data on online hate speech was collected by the project in outcome one, review of data analysis reports from project implementing partners contributed to the team's ability to

understand key areas of work within the context as it was unfolding and the degree of gender-responsiveness of project-generated data.

The evaluation team also closely analyzed the project's monitoring plan, including analysis of gender-sensitive indicators, baselines and targets, and sought to use relevant data from indicators in the project's results framework as well data collected directly through the project, mainly in Outcome 1. Early in project implementation the project team hired an independent M&E consultant to devise an M&E plan. This plan mainly consists of specifying Means of Verification (MoV) for indicators within the results framework and is accompanied by a document labeled "Explanatory Note," which are minutes of a final meeting with the M&E consultant. For most indicators cited in the M&E plan, the MoV are baseline and endline surveys meant to be conducted throughout the course of the project. In many instances the use of baseline and endline surveys, however, was not warranted. For example, indicator 1b seeks to track the number of "newly tracked gender-disaggregated indicators on social cohesion." Since the indicator is not actually tracking what new indicators are measuring, but instead merely counting the number of new indicators, there should have been no need for a baseline and endline survey. In other instances, since the outcomes and outputs themselves were crafted in a way that obscures the substantive change to which they are meant to contribute, the team was unable to devise relevant indicators. For example, output indicator 1.1.2 foresaw the increased capacity of CSOs engaged in the community of practice and planned a survey and interviews with CSO partners about their capacities. While their perceptions of their skills or knowledge may be one input into determining whether capacities have been raised, the project could have augmented perception data with more objective ways to measure the effectiveness of these new skills by looking at the outcome of their advocacy or public engagement efforts. The evaluation team turned to the Baseline Survey Report in hopes that some of these shortcomings or inconsistencies would be resolved. While the report offers rich data and analysis about three different stakeholder groups, namely media personnel, religious leaders and children and youth, it does not actually engage the indicators of the results framework and, thus, did not provide actual baselines for reporting.

Beyond mismatches between the indicators and their MoVs, the evaluation team also found the data provided by many of the outcome indicators to be of only marginal use for the evaluation, given that most often they do not measure outcomes but outputs and present a muddled, compound approach to measurement that makes tracki substantive achievement difficult. The section below on Project Management and Monitoring will provide additional details of the indicators within the results framework and an extensive analysis of each indicator is included in Annex D.

To fill gaps and validate existing data, the evaluation team collected primary data from March 5 to March 31, 2024, in Colombo, Kandy, Hatton, Batticaloa and Jaffna. When preparing the evaluation exercise, the evaluation team considered the gender, inter-communal, and logistical restrictions that may affect their efforts to collect data, how to mitigate these risks, and maximize inclusivity.

The research team conducted key informant interviews (KIIs) In Sinhala, Tamil and English, depending on the language with which respondents were most comfortable. Interviews or focus group discussions (FGDs) were conducted with a broad range of stakeholders using a purposive sampling strategy to select participants who fell within six categories of respondents: 1) UN programming and leadership colleagues in UNDP, UNICEF and the RCO; 2) project implementing partners; 3) others contracted by implementing partners to conduct specific pieces of work; 4) government counterparts in Colombo and at Provincial levels, and 5) representatives from development partners, 6) direct stakeholders of the project, including: religious leaders, journalists, social media influencers, teachers, CRPOs and YSOs, students, and youth and children's club members. In total, the evaluation team reached 272 respondents (146 men, 126

women) which break down across the stakeholder categories in the following way: in depth interviews with 27 UN staff and development partners (13 men, 14 women); 70 respondents through KIIs with governmental and non-governmental project partners, and religious leaders, journalists and social media content producers as project beneficiaries (45 men, 25 women); and 175 participants from secondary schools and Children's and Youth Clubs who participated in FGDs (88 young men, 87 young women). Within these categories, the team expressly sought to ensure a broadly representative range of opinions by seeking the relatively balanced inclusion of women and men, as well as religious, caste and ethnic group members throughout its data collection efforts. In addition to project stakeholders, the team conducted FGDs with control groups of students in Central Province to help the team triangulate to what degree changes in attitude among students may be attributed to the project. A comprehensive list of all KIIs and FGDs, including the location where the conversation took place, the date and the sex of those involved, is included as Annex C.

While interviews with project partners, development partners, governmental stakeholders, journalists, religious leaders and social media influencers were semi-structured and relatively open-ended to allow respondents to provide their own thoughts about the project's implementation and effects, FGDs with students and young people followed a more established structure to afford the team the chance to compare learning outcomes across groups. FGDs generally began with an introduction about the purpose of the exercise and an outline of the informed consent process before segueing into initial questions about which training or activities the students and young people had participated in. After establishing the timing and extent of their exposure to the training, FGD members were asked to read a series of statements[16] to each other and discuss whether they considered the statements to be hate speech. The approach enabled the team to identify common points of learning and skill adoption as well as whether there are differences across various communities and/or geographic locations.

Given the project's focus on sensitive topics on gendered, ethnic, caste, and faith-based hate speech, the evaluation team sought to hold separate discussions with each stakeholder type. For example, the team sought to hold a focus group discussion with Tamil girls participating in a Youth Club without the inclusion of boys or other community members in order to create an open and safe space where the girls could voice their concerns or challenges without fear of backlash, self-censorship, or stigmatization. In particular, this approach allowed the team to probe the specific and unique aspects and impacts of gendered hate speech on girls and boys, as well as other differences emerging from religious, ethnic or caste identities.

Wherever possible, the team sought to record interviews and FGDs with stakeholders. Following the interview or FGD, the team created narrative summaries of the conversations and entered them into Atlas.ti (a qualitative data analysis computer software package), where they were coded for key words and themes and analyzed. By entering metadata about KII or FGD participants' identities), the evaluation team has been able to analyze patterns across stakeholders to identify emergent themes and to investigate whether recurring themes are seen within specific sets of stakeholders.

To identify the geographic locations for its field data collection, the evaluation team identified where key actions connected to the project were implemented. While some aspects of the project sought to have nation-wide effects, such as social media campaigns to raise awareness on the gendered nature of hate speech, others were concentrated in a handful of provinces. After reviewing the project implementation

[16] Statements included stereotyping language as well as positive sentiments about a range of ethnic or religious communities, the roles of women and men, and openness to people who identify as LGBTQI+.

plan and reporting and discussing the field data collection plan with UN colleagues, the team identified Central, Eastern and Northern Provinces as areas of high concentration of project activities outside of the capital. The team received a list of secondary schools and Youth and Children's Clubs in Central Province targeted by UNICEF and selected a random sample in Kandy and Hatton districts to enable a substantial cross selection of young people from the Sinhala, Estate Tamil, and Muslim communities. In Batticaloa and Jaffna, the team conducted FGDs with Children's Clubs as well as KIIs with implementing partners, the Divisional Secretariat, local peace activists connected to the project, Child Rights Protection Officers and a Human Resource Development Officer, journalists and religious leaders.

Within its inception report, the team had included the possibility of launching a survey on Facebook to engage the general population on questions of social cohesion and hate speech. Deployment of this survey was always considered an approach of last resort, given challenges early on to access existing data. In the end, the team opted not to launch this survey for a series of reasons, including: 1) the team was given access to data on the incidence of online hate speech, religious violence and aspects of social cohesion, thus offsetting the need to conduct an additional survey; 2) given the nature of Facebook surveys, this method could not elicit generalizable data but merely provide illustration for evaluation findings deduced by other forms of evidence. This lack of generalizability reduced the relevance of this data collection approach vis-a-vis existing data; and 3) as we elaborate in greater detail below in the section on Constraints and Limitations, the relatively brief nature of the project's engagement with many of the project's ultimate stakeholders draws into question how much impact the project could reasonably have on the general population, thus further reducing the relevance of such a survey.

II.C.2. Constraints and Limitations

As with every evaluation process, this evaluation faced several challenges, some of which are common among evaluations of peacebuilding initiatives while others are unique to this project.

Insufficient data at appropriate outcome level to fully assess Impact and elements of Effectiveness. The project was implemented over a 27-month period, which could have potentially enabled the project team to capture nascent indications of socio-cultural, political or institutional changes specific to project activities. Indicators included in the project results framework, however, focus mainly on measuring project outputs instead of higher order changes. This limitation was identified in an evaluation of an earlier project with a similar though more limited focus on addressing hate speech to promote greater social cohesion,[17] but was not adequately addressed in the current project. Without such data, understanding how inputs like training of journalists and religious leaders or alternative, positive social media content may have positively impacted social cohesion or contributed to a reduction in hate speech will be difficult because the team will not have access to baselines and monitoring data about the overall effect of training on individuals and communities.

This limitation is somewhat offset by data collection that occurred through Outcome 1, which gathered data on the incidence of online hate speech and, to a lesser degree, on elements of social cohesion. Most of these broad-based data collection efforts, however, are pitched at the national level and do not easily enable the project management team or the evaluation team to connect actions taken in the geographic areas targeted by the project to the available data on higher order changes.

[17] Final Evaluation Report "Preventing Violent Extremism (PVE) through Promoting Tolerance and Respect for Diversity in Maldives and Sri Lanka," (Jan 2022)

Short period of time that ultimate stakeholders were exposed to project outputs within Outcome 2. Even with access to data on incidence of hate speech and some forms of social cohesion, training and skills building among students, youth and children's club members and religious leaders largely occurred in the last four to five months of the project's life cycle. The degree of exposure to training, moreover, varied substantially across geographic space and raises additional questions about the role exposure to training might have played in affecting attitude or behavior change. In Northern and Eastern Provinces, for example, Club members participated in training workshops that lasted two or three days, while participants in Central Province report having received only one two- to three-hour session on identifying hate speech and their responsibilities as good digital citizens.[18] As a result, even if the evaluation team can identify changes in social cohesion or the incidence of on- or offline hate speech, it is not plausible to attribute positive changes in these areas to the actions of the project. These limitations impede the team's ability to generate findings on the criteria of Impact and Sustainability, and to a lesser degree Effectiveness. These limitations were mainly observed within elements of Outcome 2. The team sought to address this limitation by conducting FGDs with religious leaders, students and youth who did not have access to the training in order to see whether there are patterned differences between those who did and did not receive training.

The structure of the project does not lend itself to evaluation of collective outcomes. As we unpack more fully in the evaluation's findings on Coherence and Project Management and Monitoring, the project document makes clear that the ultimate aim of reducing hate speech is to promote social cohesion. The project's two outcomes, however, are designed in such a way that work in one area does not influence or affect work in the other, which has led to implementation of the project that has advanced in parallel tracks. And as noted above, the project's results framework measures change that remains largely at the output level, thus adding to the difficulty of assessing how the project's two channels may have contributed to collective outcomes. While the evaluation has been able to document a number of positive results and promising areas of work as a result of this project, the team has been constrained in its ability to evaluate the contribution of these positive outcomes to broader change related to social cohesion or beyond the immediate stakeholder groups.

III. FINDINGS

III.A. RELEVANCE

Q1. Did the intervention seek to positively affect key factors and actors that contribute to or combat hate speech? Is hate speech a driver of weak social cohesion?

Q2. Did changes in the context trigger changes in the relevance of the project's goals or strategy?

Q3. How relevant was the targeting and selection of beneficiaries reached through this project?

Q7. Did the project align to country-specific aspects of SDG 16, relevant pillars/peacebuilding outcomes of the SDCF, or government efforts to promote social cohesion?

Finding 1: The project's design was relevant to underlying factors and actors that contribute to social divides and hate speech that undermine social cohesion in Sri Lanka. The project remained relevant through a series of crises that occurred during implementation and continues to be relevant at project's end, as the country heads into an election cycle.

[18] Shorter exposure to training was noted more prominently among students than Club members.

The country context at the time of project design in mid- to late-2021 was characterized by an increasingly severe economic and political crisis. Even before the COVID-19 pandemic, analysts noted worrying trends among Sri Lanka's ethnic and religious communities, with the Sinhalese majority expressing increasing resentment at perceived preferential treatment in favor of ethno-religious minorities by the government and rampant anti-Muslim hate speech in the wake of the 2019 Easter Sunday attacks. Half of respondents in the UN's 2019 Perception Survey on Peacebuilding, moreover, attested to having personally witnessed incitement to religious-based violence and nearly all communal groups registered widespread distrust of political institutions, with most asserting that political parties represented the greatest barrier to sustaining peace in the country. Periodic analyses by the Resident Coordinator's Office since 2018 reveal persistent concerns that hate speech circulating on social media reflected widespread mistrust and resentment within society while also amplifying and encouraging such division. Given this broad view of on- and offline hate speech and its contribution to social tension, the evaluation team assesses that the project's approach was largely relevant to contextual factors and actors of conflict.

Almost immediately after project implementation began in early 2022, the government's mishandling of economic policies and a general perception of widespread corruption led to mass protests that eventually brought down the Rajapaksa government in July 2022. During this time, schools and public and private services were periodically closed. In this fraught environment, UNDP and UNICEF reported difficulty in launching certain aspects of the project in a timely manner, as the attention of non-governmental project partners was directed at the political crisis, and government officials were either difficult to reach or, eventually, replaced in office. By November 2022, the project's annual report cited delays in training for the project's media and education/youth components due to the crisis, including fuel scarcity that prevented participants from traveling to project training and changes in the government within the Ministries of Education and Mass Media, which necessitated the project team to gain the support of a new set of governmental officials. During these transitions, UNDP and UNICEF were both careful in their messaging to new officials, modulating the terminology of the project's aims to gain acceptance and understanding of new partners and working hard to foster their trust.

Reports of online hate speech monitoring by project partners revealed that the public's dissatisfaction with the Rajapaksa government and backlash against protesters and women, combined with an uptick in anti-Christian hate speech, led to a reduction of anti-Muslim hate speech in 2022, although it was unclear at the time whether the reduction would be a temporary or sustained trend.[19] In addition to the broad relevance noted above, this observation illustrates another aspect of how the project remained relevant in the midst of a fast-changing context. Specifically, while the targets of hate speech at the national level have shifted over time – from Northern Tamils in the early years after the war, to Muslims following the Easter Sunday attacks, to increasing hate speech against Christians during the COVID-19 response and into the current economic crisis, for example[20] – the phenomenon of using hate speech and mis- and dis-information on social media and in more traditional mass media to drive social division has remained constant. The project's design, which addresses a wide array of hate speech expressions, including those rooted in ethnic and religious divisions, as well as backlash against gender equality and women's empowerment, enabled the project to speak across these divides and respond to the ways in which hate speech changes and adapts over time. Reflecting on conflict dynamics in 2021, it would have been

[19] See, for example, Monitoring Hate and Dangerous Speech (Online) in Sri Lanka, 2022 Annual Report, Hashtag Generation and Analysis of Incidents of Violence and Intolerance against Religious Communities, October - December 2022, NCEASL

[20] Monitoring Hate and Dangerous Speech (Online) in Sri Lanka, 2022 Annual Report, p. 2, Hashtag Generation

plausible for UNDP and UNICEF to have proposed a project focusing on anti-Muslim hate speech and physical attacks that ramped up after the Easter Sunday attacks and continued to morph into different types of anti-Muslim hate speech with the onset of the COVID-19 pandemic. Having done so, however, would have tied the project too narrowly to one group and undermined the project's relevance as the targets of hate speech have evolved and shifted over time. Instead, the project's general approach to all forms of hate speech has enabled it to flexibly adapt to emergent trends, fill an important evidence gap, and remain relevant.

While data does illustrate the ebb and flow of hate speech against specific groups, hate speech or cyber Gender-Based Violence (GBV) against women and sexual minorities, which has constituted over 50% of online hate speech at various times over the project life cycle,[21] have remained a consistent and quite virulent problem over time.[22] This hate speech is directed at women from all walks of life and social groups, from higher profile female politicians to average women on the street. Online hate speech, moreover, reflects offline abuse and barriers to women's economic, social and political participation.[23] The project did address gender-based hate speech to an extent, especially through the interventions of two implementing partners, the Center for Equality and Justice and Search for Common Ground, which created support networks among a handful of female social media content creators and helped build their skills in responding to online abuse. A Community of Practice composed of civil society partners, facilitated by UNDP, moreover, embedded the monitoring of new dimensions of gender-based hate speech within a new Code of Practice for tech companies. Given the ever-present on- and offline discrimination that Sri Lankan women of all social strata, religions, ethnic and caste groups experience,[24] however, the project's relevance could have been amplified by an even greater focus on addressing gender-based hate speech through approaches that included not only women but allied men as well.

Since the resignation of the Rajapaksa government in July 2022, the caretaker government of President Wickremesinghe has sought to settle the political and economic crisis, though often through means that have been characterized by development partners and human rights advocates as occasionally punitive and designed to silence dissent. The government has utilized the 1979 Prevention of Terrorism Act to indefinitely detain protesters and political opposition,[25] while proposed amendments have been met with sharp criticism for not resolving the worst violations of the Act.[26] Most recently, on 24 January 2024, Parliament adopted the Online Safety Act, seen as a potential lever for controlling political opposition, over the objections of opposition and the concern of development partners. While the Cabinet has since announced plans to amend the Act, and key project partners such as Factum are a part of that process, the application of such legal measures signifies that space for political opposition, dissent and free

[21] See monthly reports of Hashtag Generation, Factum and Deloitte.

[22] From January to August 2022, for example, online hate speech targeting women exploded in the context of social and political unrest. Real time monitoring by Hashtag Generation revealed that 60.25% of online GBV was directed at women, mainly politicians and activists, as well as members of the LGBTQI community.

[23] In the 2021 Global Gender Gap Index, Sri Lanka ranked 116 out of 156 countries, having fallen from 100th place in 2018

[24] Even before reported increases in SGBV and intimate partner violence during and after the COVID-19 pandemic, the Women's Wellbeing Survey of 2019 recorded that nearly 40% of Sri Lankan women experienced sexual, physical, emotional or other abuse at some point in their lives, while 18.9% of women reported physical violence against them, a phenomenon that reaches 37.9% among women in the estate sector.

[25] Amnesty International, "End the Use of and Repeal the Draconian PTA," accessed on 1 April 2024 at https://www.amnesty.org/en/wp-content/uploads/2022/02/ASA3752412022ENGLISH.pdf

[26] OHCHR, "Call to Sri Lanka to revise anti-terrorism bill," press release, 19 January 2024.

expression appears to be at risk. This raises the third aspect of broad contextual relevance for the project, that of preserving the space for non-governmental actors, development partners, the United Nations in Sri Lanka, and others to continue to engage and promote issues related to human rights and peacebuilding broadly speaking.

Finding 2: The approach is rooted in existing national and UN priorities, including the SGDs.

The project has been firmly grounded in national and international policy frameworks, which speaks to its relevance to broader priorities and helps enhance the chances of sustainability. For example, the project's overall aims are clearly represented in pillar four on social cohesion within UNDP's Flagship Portfolio on SDG 16: on Peace, Justice and Strong Institutions and within its Country Programme Document (CPD). This is a relatively new pillar, having been adopted in 2021 when the overall portfolio was revised following a change of government. Within this pillar, UNDP has prioritized the promotion of data-driven early warning mechanisms, social dialogue, information integrity and media governance, among others, as key approaches. This prioritization carries over to the UN's Sustainable Development Cooperation Framework, which cites weak social cohesion as a main barrier to sustaining peace and promoting economic growth and prosperity and articulates the UN's work through UN SDCF Strategic Priority 5 on Social Cohesion and Governance. By extension, support for early warning mechanisms represent one of three priority areas of Sri Lanka's Peace Window in the SGD Multi-Partner Trust Fund.[27] Similarly, the project is grounded in UNICEF's CPD 2023-2027 through the CPD's Education Pillar, which enables UNICEF to take a more programmatic approach that increases the relevance and sustainability of certain key outputs beyond the project's lifecycle.

Beyond UN strategic frameworks and priorities, the project has partnered with a number of governmental and non-governmental institutions to advance their priorities. UNICEF's approach to align implementation of its work with secondary schools under Output 2.4, for example, has bolstered chances that resources and systems created by this project will not duplicate or work at cross purposes with the Ministry of Education comprehensive curriculum reforms. By opting to partner with the Ministry in this way, UNICEF has ensured that aspects of the project relating to digital citizenship, critical thinking, and support for inter-group tolerance are reflected as dedicated, mandatory coursework, and that the basic principles of this curricula will be mainstreamed throughout the reforms.

Similarly, UNDP's hard-won partnership with the Ministry of Mass Media for activities related to Output 2.2, aligns well with the Ministry's support of the media sector to improve ethics and raise the level of professionalism within journalism. Regarding work with technology companies, between 2022 and 2023, the project supported CSOs to develop and improve a Code of Practice for social media platforms and produced a Code that for the first time included separate consideration of gender-based hate speech. This Code is seen as a meaningful and relevant self-regulatory framework for social media platforms, providing an alternative to a legal framework regulated by the government.

Finding 3: Most project stakeholders and resources were relevant, with some exceptions.

As will be discussed in greater detail in the section on Coherence, the project worked with many kinds of project beneficiaries through distinct, and at times unconnected, avenues of work. The remaining analysis on Relevance will consider each major project stakeholder group, set of outputs, and/or issues regarding

[27] See the United Nations Sri Lanka SDG Multi-Partner Pooled (UNSLSDG) Fund Peace Window Vision Paper

geographic focus on its own terms, leaving questions of Effectiveness, Efficiency and Coherence to subsequent sections.

In 2019, seven in 10 Sri Lankans from all ethno-religious communities identified religious leaders as among the most trusted actors in the country, together with the courts and the media.[28] **The project's approach of incorporating religious leaders as agents of positive change, then, appears to be highly relevant to the context.** This positive disposition was confirmed during data collection in FGDs with young people, who expressed trust in and deference toward local religious leaders. Religious leaders, in addition, have been found to be less engaged on social media than other potential stakeholder groups, so facilitating their greater involvement in combating on- and offline hate speech in their communities in order to positively affect social cohesion underscored the relevance of working with them. The project document, however, indicates the goal of ensuring that 20 to 30% of religious leaders reached by the project are women. While the evaluation team recognizes the intention to ensure broad inclusivity within the project's approach, barriers faced by the implementation team within the formal structures of the religious authorities to include women highlight that expecting to reach female religious leaders with the same messaging and through the same approaches as male religious leaders was not as relevant to the context. These observations will be elaborated more in the findings on Gender and Human Rights and within Recommendations.

As noted above, Sri Lankans also express trust in the media, but evidence-based analysis also identifies media professionals as producers of mis- and disinformation and hate speech, including through the unnecessary inclusion or focus on ethnic or religious details or misogynistic reporting.[29] This observation was confirmed through field data collection among project stakeholders across all sectors – from Ministry-level stakeholders to students and young people at the local level. **The project's targeting of journalists, then, is seen to be highly relevant**. Moreover, the project's specific targeting of provincial journalists, who typically have not received training in their home regions, was seen as highly relevant and very much appreciated. As relevant as ensuring greater professionalism and awareness among journalists was, however, the team recognizes that addressing these changes is a necessary but insufficient condition, given the heavy hand that editors and media owners can play in reshaping a story. While the project reached news editors with one intervention late in the project lifecycle, in 2023, it is unclear how effective this has been, especially given the brief nature of the exposure and the popular appetite for inflammatory reporting that generates financial incentives that drive editors and media owners to publish such stories. In essence, addressing knowledge or awareness gaps among journalists strikes at only part of the problem.

The project's theory of change for engaging children and youth in and out of school was to shape their views as both responsible citizens and social media consumers, given that they are both much more present online than other stakeholder cohorts and vulnerable to positive or negative influences in their lives. **Similar to work with religious leaders, targeting children and young people was determined by the evaluation team to be highly relevant.** The means of working with young people, through Children's and Youth Club, which are present in just about every part of the country, moreover, was also seen as relevant but also a challenge. Clubs historically have brought girls, boys, young women and young men together around sporting and cultural activities. Working through Clubs to reorient their work toward

[28] 2019 National Perception Survey on Peacebuilding, p. 2.

[29] Reviewing Dangerous Content Across Sri Lanka's Digital Ecosystem, SecDev Analytics, 2021.

increasing knowledge about hate speech and online safety initially was not seen by CRPOs and YSOs as resting within their mandate nor did they feel comfortable with or have knowledge about the topics.[30]

While conducting field data collection, the evaluation team was also introduced to a number of **social impact projects implemented by Children's and Youth Club members** who underwent training on hate speech. To receive funding, projects needed to produce a tangible output that would benefit the community. The aim of this line of work was that by implementing the projects, youth would be able to practice the principles of tolerance and understanding that they were introduced to through the training. While in the field, the evaluation team witnessed social impact projects across a wide array of topics, including the establishment of a Club library, a project to share seeds with local community members, the extension of water pipes to a Club meeting house, and the construction of a fence to prevent flooding and landslides on a key thoroughfare during the rainy season. During FGDs with Club members, 100% of respondents expressed satisfaction and pride in their achievement but were unable to link the process and outcome of these small-scale projects to the aims of reducing hate speech or promoting social cohesion, raising questions about the relevance of this aspect of the programming for advancing the project's overall aims.

Finally, although not prominent within the project, specific **work with female social media influencers through the Center for Equality and Justice (CEJ) demonstrates the relevance of ensuring an inclusive approach to targeting with a dedicated focus on women and the LGBTQI+ community** when it comes to addressing online hate speech. While this point is detailed in the findings on Gender and Human Rights and in Recommendations, participants in CEJ's training indicated the importance for their own safety and overall wellbeing of having developed a strong network of social media content producers with similar aims and challenges, which was a new development for them.

III.B. ENDS

III.B.1. Effectiveness

Q4a. Did individual level change related to attitudes and behaviors toward hate speech extend to broader changes, including at the community level?

Q4b. Did the project's approach of working through key government structures on one hand and an array of other stakeholders, including young people, religious and community leaders and media, yield meaningful changes in combating or reducing hate speech?

Q5. Were there any unexpected results or unintended consequences of the results (both positive and negative)?

Q6. To what extent did the project make timely adjustments to its implementation strategy to maintain its effectiveness?

Evaluation questions pertaining to effectiveness focus on the project's work on various aspects of on- and offline hate speech, while investigation into the project's contribution to social cohesion, will be taken up in the section on Impact. The team is taking this approach because improvements in social cohesion are prime examples of the kind of social, cultural, or institutional changes we would expect at the highest level of change, i.e. the impact level. Further, we cannot assess the contribution of hate speech reduction to improved social cohesion until we first establish whether there has, indeed, been a reduction of hate speech. We also recognize that when we examine reductions in hate speech, we are not anticipating a

[30] 75% of CRPOs, YSOs or the organizations who hired them expressed this view in KIIs.

gross reduction at the national level but will examine change at a level more proximate to project stakeholders, focusing mostly at the individual and immediate community levels, where community can signify a broad range of configurations, including schools, clubs, religious, professional, and local communities. We deem this appropriate given that the project timeframe was planned for 18 months, eventually extended to 27 months, rendering assessment of wider impacts unrealistic.

As detailed below in the sections on Cohesion and Project Management and Monitoring, the project's two outcomes have been implemented largely in parallel, making collective contribution to common results across the project's outcomes and outputs difficult to discern. To assess effectiveness taking this challenge into account, the evaluation team has approached questions of effectiveness through a two-fold lens: one, corresponding to Outcome 1, that examines the project's contribution to improving the social and traditional media landscape – i.e. the policies, regulations and priorities of government, media and technology companies, and other key actors – with respect to hate speech, and another, touching upon Outcome 2, that examines the project's contribution to supporting project stakeholders to improve their understanding of the negative consequences of hate speech and their roles in promoting positive narratives.

Finding 4. The project has made positive inroads to securing important policy and regulatory changes that may help improve the overall social and traditional media landscape.

As the project document notes, one clear aim of the project was to fill a data and information gap on the incidence and impact of hate speech across various media and platforms. The project sought to expand an earlier initiative of the Resident Coordinator's Office and UNDP to systematically monitor social media platforms by contracting Hashtag Generation in early 2020. Filling this data gap was seen as crucial to mount effective and timely responses to hate speech as well as to launch advocacy campaigns with government and media and technology industry leaders to regulate and curb hate speech while preserving freedom of expression. To advance this aspect of the project, UNDP contracted five organizations and/or companies, including:

Hashtag Generation, which focused its monitoring on Facebook, TikTok and YouTube. Hashtag's approach, while producing data and analysis that by all accounts is reliable, is human-centered and quite labor-intensive, involving identification as well as cross-checking/validation by a team of individuals. Hashtag produced 27 reports over the project period, including 25 monthly reports and two annual analyses.

Deloitte's contract aimed to test whether new advances in artificial intelligence and machine learning could be configured to automate some aspects of coding to help reduce the transaction-heavy human-centered approach of Hashtag. Deloitte produced four monthly reports and one final report over the project period.

Factum was contracted to test whether Hashtag's human-centered approach could be decentralized to work through collections of male and female senior journalists based in three areas in the east and south of the country. These data miners produced monthly reports that were shared with Factum and formed the basis of Factum's final report that analyzed incidence hate speech from November 2022 to February 2023.

SecDev, a Canadian consulting group, was contracted to test to what degree other platforms could be added to create a more complete picture of online hate speech. Their monitoring focused on YouTube,

which they saw as used mostly by the Sinhala community, as well as messaging applications like Telegram, which is used more frequently by Tamil and Muslim communities, WhatsApp, and Viber. SecDev produced six monthly reports and one final report.

PAFFREL, received a short-term contract specifically to monitor online hate speech connected to anticipated local elections in early 2023. Monitoring was conducted on Facebook, YouTube, and TikTok platforms PAFFREL produced four weekly reports in February and March of 2023.

In addition to monitoring social media platforms, the project also contracted two other organizations, the **National Christian Evangelical Alliance, Sri Lanka (NCEASL)** and the **Center for Policy Alternatives (CPA)** to monitor the incidence of offline religious violence and communal or localized violence, respectively. NCEASL's efforts rely on long standing expertise within the organization for monitoring religious-based violence throughout the country to produce monthly and quarterly consolidated reports, while CPA's approach tested the feasibility of creating a network of 60 locally based "peacebuilders" who could report on and verify inter-communal incidents throughout the country. Following a period of training, on the basis of peacebuilders' regular reports and focalized incident reporting through a dedicated phone-based application, CPA produced a series of bi-weekly reports in late 2023, as well as issue-specific snapshots that provide focused conflict analyses of various issues, from water-based conflicts to land use and beyond.

Together with open-source data and analysis, these various on- and offline data collection efforts form the evidence base for two inter-related mechanisms established and maintained by the RCO and UNDP, UNDP's Crisis Risk Dashboard, and the UN in Sri Lanka's Early Warning System (EWS).[31] Feeding into the EWS, the CRD provides the data, trend lines and geo-spatial tagging for incidents in several key issue areas such as hate speech, human rights violations, and land issues from verified data collections.[32] The Early Warning System utilizes this data to provide analytic reports through various formats, including regular monthly and quarterly reports as well as issue snapshots and issue papers. The data reflected in these mechanisms has been key to shaping strategic priorities on the ground and has contributed to the evidence base for the UN's periodic Common Country Assessments on which strategic plans are formed. While the sensitive nature of some topics within these mechanisms has compelled the UN in Sri Lanka to maintain a policy of strict confidentiality that prevents users from downloading or sharing the data, specific reports and analyses have been shared on a case-by-case basis with the CoP. The EWS, in particular, has been cited by UN staff and development partners as a valuable resource of verified data that has enabled them to advocate with government and their own offices or organizations for evidence-based policy initiatives and prioritization that responds to emergent problems. Development partners, however, noted that easing restrictions on sharing the EWS data would enable them to raise alerts more effectively with internal and external colleagues and promote coordinated action based on a common understanding of the context.

The evaluation team assesses that data collection efforts within Outcome 1 have indeed generated important evidence that can animate such early warning mechanisms, and that without access to

[31] The CRD, however, does not utilize data collected by CPA, which will be discussed in more detail in below sections on Effectiveness and Conflict Sensitivity.

[32] While the evaluation team acknowledges the CRD as a valuable resource, it also notes some limitations to its effectiveness. Specifically, it is difficult to identify sources for the data on which the trends are based and the cite does not appear to allow for filtering based on key demographic features like sex or age to yield a more nuanced and granular understanding of the trends.

updated data, mechanisms such as the EWS and CRD quickly become irrelevant. That said, the team was only able to verify how data from Hashtag Generation's monitoring of hate speech has been utilized within the EWS, despite substantial investment in the efforts of others. Project partners such as CPA and NCEASL that produce incident data on which the systems operate, moreover, reported that they do not have information about how their data is used once it is delivered to the UN, a shortcoming they view as limiting since they are unsure whether organizations are responding to the social problems they are raising. **Taken together with limitations in sharing the EWS reports and data, the majority of project and development partners reached through this evaluation urged greater transparency as these systems evolve and mature.**

While the project took direct aim at reducing or responding to specific instances of hate speech, **it also adopted a systemic approach by supporting the development of a voluntary Code of Practice, the first such code in the Asia-Pacific region,** to be adopted by internet-based technology companies. The Code, developed through a participatory process that began in March 2022 through a multi-stakeholder group of Sri Lankan nationals facilitated by project partner Factum, outlines a set of principles and concrete measures which signatories are expected to adopt.[33] As of February 2024, all members of the Asia Internet Coalition, such as Apple, Facebook, and Twitter (now known as X), have endorsed the Code, with Google as the notable exception. While the Code may not reduce how much hate speech and other harmful content is uploaded to social media platforms by users, if successful, it should lead to faster and more effective response from the companies. It is, however, too early to assess the true effectiveness of the Code's application. The only available data on tech companies' responsiveness is reporting by Hashtag Generation which produced a report for 2023. That report helpfully provides analysis of what percent of harmful speech has been responded to by category of hate speech for Facebook, TikTok, and YouTube, but does not include change over time to enable the evaluation team to see whether endorsement of the Code has improved companies' response rates.

UNDP also sought to strengthen systems by establishing a Community of Practice (CoP) to provide a platform through which like-minded civil society and possibly development partners could coordinate action to combat hate speech. The CoP began in the early months of project implementation, with its first meeting in June 2022. By November 2022 the CoP had already met three times, despite disruptions to work plans and re-ordered priorities caused by the political and economic crises that delayed other areas of programming. The evaluation revealed evidence that **after two years of close support by UNDP, the CoP, currently sitting at 22 active members, has become a dynamic network for sharing data, analysis and coordinating advocacy and action.** Concretely, CoP members have held consultations on the dynamics of cyber GBV that have helped secure monitoring within the Code of Practice for the tech industry.

Finding 5. Engaging religious leaders, provincial journalists, social media content producers and young people in and out of school yielded changes largely within the individuals who participated in project activities, with some signs of broader reach.

The project deployed a range of very different approaches with its diverse set of stakeholders. Some of these approaches, like those with social media content producers, were by definition designed to foster individual-level change that could produce content or actions that should reach a wider audience. Others, like training young people to identify hate speech, were less clear on how change at the individual level

[33] Sri Lanka Code of Practice for Online Safety and Responsible Content; the Code includes reporting across six themes, including: hate speech, incitement to violence, disinformation, misinformation, cyber bullying and sexual abuse and exploitation.

was meant to concatenate into broader change. Because the approaches and stakeholder groups were so different, this section will address outcomes for each group in order to capture the differences in approach and outcome.

Religious leaders. Targeting religious leaders within the project was meant to help sensitize a key group of trusted actors to how online and offline hate speech undermines social cohesion. Like work with students and young people, detailed below, delays in rolling out initiatives with religious leaders meant that many project activities took place within the last six months of project implementation. In addition to the economic and political crisis that presented barriers to timely implementation throughout much of the project, other delays are specific to the challenges of working with religious leaders. Three of the four main religions targeted by the project have comparatively strict hierarchies, requiring advocacy, agreement, and awareness raising among the leadership prior to engagement with religious leaders in target communities. Generating consent across all programming commitments such as the inclusion of female religious leaders, moreover, proved difficult. Not until January 2024, less than two months before the project closure, did UNICEF manage to mobilize the participation of 75 female Islamic religious leaders in a meeting held in Colombo.

Despite delays, however, by November 2023 UNICEF reported that it had provided training or awareness raising to 121 (male) religious leaders through four-district level forums. In KIIs conducted by the evaluation team, **all faith leaders reported that their own awareness about the negative impact of hate speech had been raised**. Faith leaders most frequently cited that the workshops they attended made them more apt to fact check what they read and to reflect more deeply on how their own posts or social media shares may be perceived. Nearly half of those interviewed, moreover, reported weaving positive narratives into their sermons and their interactions with young people in their communities.

Perhaps equally important as their heightened sensitivity to intolerance and hate speech, religious leaders report that inter-religious meetings conducted under the auspices of the project broke down long held stereotypes and increased their understanding and cooperation with others. Improved relationships have already translated into several concrete interventions, such as a Hindu leader eliciting support from a Buddhist monk to intervene with authorities on behalf of a parishioner and an Anglican priest rallying support for Hindu community members made homeless by a fire. While these, admittedly, constitute anecdotal accounts of individual level change, these instances of inter-religious assistance have sprung from the relationships that have been established through the workshops and bode well for further, broader engagement if continued support is provided.

Social and Traditional Media content producers, and gamification. The project produced a wealth of interesting and compelling content, delivered through a range of online and offline producers, that has been largely well-received. 4K, NCEASL and CEJ were all tapped to support online content producers to develop social media content and campaigns to raise awareness about the negative effects of hate speech, cyber GBV and mis/disinformation. Implementing partners and content producers report having generated broad online engagement with their content. For example, two young content producers working under a partnership with 4K recorded a song written for the campaign that paired with compelling illustrations to probe family and societal resistance to interfaith marriage. The video garnered more than 16,000 views, representing a vastly larger audience than either content producer typically attracts, with universally positive reactions. The two report that this was the first time that either one had intentionally created content on a social problem, but their positive experience has encouraged them to develop future messages. NCEASL, meanwhile, supported 50 university students equally divided between young women and men to produce positive content on social media. Divided into teams, they completed seven social

media campaigns through a range of approaches, including online quizzes to combat hate speech, short video competitions, messages on religious diversity and a short video on cyber GBV. Finally, as noted above, CEJ developed a network of 30 young women and those who identify as LGBTQI+ who underwent extensive training throughout the project lifecycle, including a 20-day residential program that brought participants together for face-to-face learning and networking. Their anti-cyber GBV campaigns reached approximately 200,000 viewers on TikTok and Instagram from March to April 2023. In contrast to other content produced through the project, these content producers experienced substantial backlash and cyber-attacks for their posts, a phenomenon that may indicate that their content was reaching audiences beyond their immediate circle of like-minded social media users, as much as it also reflects deeply entrenched gender-based hate speech on Sri Lankan social media. Each of these discrete efforts was effective at producing visually compelling and engaging material that, in one way or another, resonated with audiences and provoked reactions. As one implementing partner noted, however, **each of these campaigns was launched separately and without connection or coordination with other efforts.** Given that many touched upon common topics, the evaluation team concludes that UNDP missed opportunities to place content creators or implementing partners who supported them in contact with each other. Having done so could have further amplified the reach and effectiveness of the individual messages and posts, enabling content producers to repost each other's content and create more of a buzz to attract additional viewers and raise awareness.

In addition to online content, the project supported content development through television media, with implementing partners Project 72, which is producing a docuseries confronting historical inaccuracies that drive social division, and Time and Space, which produced a series of television spots to raise awareness on a number of topics. Six video commercials produced in both Sinhalese and Tamil by Time and Space aired each day for three months across three Sri Lankan television channels and touched upon harmful speech, social cohesion, misinformation on the economic crisis, non-violent communication, gender, and ethical journalism. In addition to these commercials, Time and Space ran 18 half-hour interviews on primetime on similar topics.[34] **While reaching a broad audience, Time and Space indicated that the commercials and interviews primarily resonated with academics and intellectuals, who may have already been receptive to the messages**. Project 72, on the other hand, experienced delays in delivering a set of three one- to one-and-a-half-hour television spots confronting historical inaccuracies. In lieu of full-length videos, Project 72 produced three 10-minute shorts, although these do not appear to have been aired as of the drafting of this report. In fact, both Project 72 and Time and Space reported a lack of clarity about whether or how their video contributions have been used or may be used in the future. Given the high production values these videos represent, the evaluation believes there are opportunities to capitalize on these compelling pieces and repurpose them to support future training and educational efforts on hate speech.

Finally, as a third approach to raising public awareness on hate speech, UNDP launched an innovative online game to test and hone the skill of users to identify hate speech and mis/disinformation. Implemented by Moonshot through a four-month contract, the approach modified an existing game to the Sri Lankan context before launching it on Facebook and Instagram for one month, from September to October 2023. Despite the short run time, 365,844 of the reported 3 million Sri Lankans who saw online announcements of the launch clicked through to the actual game. Despite its reasonably good audience capture, learning outcomes were mixed. In pre- and post-tests administered to gamers, users' pre-test scores were marginally higher than their post-test scores, raising questions about the effectiveness of the

[34] One commercial, which promotes gender equality and social inclusion of people in the LGBTQI+ community, won a national award, an indication of the high production value of these commercials.

game. Moonshot has offered plausible reasons for the dip, including the possibility that the post-test questionnaire was too difficult or that users who were already familiar with the topics opted out of the game before completing it. Given its short 4-month contract, Moonshot reported that it was unable to validate questions prior to launch or establish a control group to capture what was learned by whom. This is a missed opportunity to learn whether games such as this are effective learning tools. Signaling a clear demand for games of this kind, 40% of students and young people engaged in UNICEF-supported training who took part in FGDs for this evaluation specifically cited a desire for games and other engaging pedagogical tools. Given this interest, the evaluation team concludes that games can be effective at teaching new information or imparting skills but urges UNDP and UNICEF to grant longer timelines for implementing partners to be able to establish robust monitoring systems to test the games' efficacy.

Journalists and media administrators. As part of the project, a training program for the professional development of provincial journalists was conducted by the Ministry of Mass Media in collaboration with UNDP from December 2023 to February 2024. **A needs assessment was conducted by Verite research at the outset of the project to generate a comprehensive view of the knowledge gaps and capacity requirements of provincial journalists, which provided a solid evidence base and baseline to frame the training.** Based on the needs assessment, Factum developed a set of training modules covering four thematic areas: crime reporting, hate speech in reporting, information disorder, and concerns of privacy in reporting. In the end, 373 provincial journalists participated in a series of two-day training programs that took place across all nine provinces. Before the training for provincial journalists, the project also conducted two discussions covering similar topics with the heads of print and electronic media institutions in August 2023.

Key informant interviews with the Ministry confirmed that the project was seen as an opportunity to work with provincial journalists. The need to work with provincial journalists had been a long-standing priority, but lack of funding had prevented the Ministry from supporting their professional development. **Training through this project represented the first such program for many provincial journalists, especially younger journalists who lack academic background in journalism or prior exposure to formal training programs.** The training workshops were effective in encouraging provincial journalists to identify the ways in which their reporting may contribute to hate speech. The starkest example occurred when one participant saw his own reporting presented as an example of hate speech within the training materials, prompting him to reflect that he had been unaware of his ethical obligations or of the possible legal consequences of submitting inflammatory or unsubstantiated accounts. During data collection for this evaluation, journalists who participated in the training programs confirmed these lessons, underscoring that the training programs helped to promote the work of journalists who report ethically and provided opportunities to learn from each other and from experts about their ethical and legal obligations. Participants in the trainings across the provinces also highlighted improvements in terms of their understanding of professionalism, the importance of critical self-reflection and how these have produced personal attitudinal change. For those without formal training in journalism, exposure to concepts like "the 5Ws and 1H," was also an important learning outcome.

Despite these positive individual effects, nearly all respondents indicated that **the length of training was too short to be as effective as it could have been**. Journalists rightly pointed out in KIIs that practices that have been in place for decades cannot be radically changed with a two-day program. Instead, **continuous engagement and follow-up to the training programs would have been more beneficial in ensuring that information to which journalists were exposed – some for the first time in their careers – can be applied**

in their daily professional lives. Journalists also voiced a variety of practical constraints that hinder their work, including threats to personal security, lack of official press cards issued by the Ministry, excessively low wages, and a lack of equipment necessary to do their work, many of which were identified by the needs assessment carried out by Verite.

Aside from additional support at a personal level, provincial journalists also flagged that improving their level of professionalism is only part of the problem. Responding to demand for inflammatory or salacious news from reader and viewership, newsroom editors and owners of media companies frequently direct journalists to develop sensational stories or edit submissions in ways that play up social flashpoints and chauvinisms.

Recognizing this problem. the project organized two workshops – one with news administrators and another with newsroom editors – under the auspices of the Ministry of Mass Media. Through these workshops, participants from both groups reached consensus on principles of self-regulation on several topics, including a commitment to avoid hate speech, mis- and disinformation, and how to ethically report on crime and protect individuals' privacy. The intention was for the Ministry to reach administrators and editors with the consensus document through the Media Broadcasters' Guild and the Editors' Guild, but as of the drafting of this evaluation, the Ministry reports that they have received no response from either Guild. Reflecting on industry actors' non-compliance, Ministry colleagues cited the lack of incentive for newsroom editors or media administrators to put aside the prospect for financial gain or political favor in order to promote ethical reporting. **Given the incommensurate goals of maximizing profit versus producing more nuanced reports, unless demand for news that generates hate or inflames tensions decreases, journalists will face difficulty in applying many of the lessons from the training or face losing their jobs in the longer term.** Demand for such content originates with the public but is driven by media owners and editors seeking to maximize their outlet's profitability by responding to public fascination with inflammatory content, issues that the project did not fully address but will need to if an across-the-board improvement in print and broadcast media is to be realized.

Students, Children's and Youth Club members. Young people were reached through the project in secondary schools, as well as in Children's and Youth Clubs that are present throughout the country.[35] By embedding the training of young people within established institutions such as these, UNICEF aimed to support already existing structures that are mandated to shape young minds and guide them to be educated and responsible adults. While the aims of working within these two institutional settings were the same, the programmatic entry points and approach to training was different across institutions and geographic locations, with correspondingly different outcomes for learning.

Organization of the training

[35] While another avenue of UNICEF work accompanied curriculum reforms at the national level, these were being finalized only as the evaluation was underway. As such, reform efforts are addressed in sections on Relevance, Coherence and Sustainability, but will not be examined under Effectiveness, which will focus on concrete changes of attitudes, behaviors or practices among project stakeholders as a result of implementation that was conducted during the project.

In North and Eastern Provinces, where work with Children's and Youth Clubs was rolled out through a 10-month long[36] implementing partnership that started in February 2023 with Sarvodaya, a well-established Sri Lankan civil society organization, training within Clubs took place over three or seven days spread over several months and occasionally was conducted in a cross-communal setting that was welcomed by participants. To conduct the training, Sarvodaya enlisted the help of Hashtag Generation, whose training content and style was also widely appreciated by both young women and men who participated in FGDs in Batticaloa and Jaffna. Sarvodaya also reached CRPOs with training and awareness raising on online hate speech, but in recognition of how unfamiliar Officers were with social media and hate speech, opted for experts from Hashtag to conduct the training with young people.[37] In fact, such was the lack of skill and knowledge of the officers prior to training, the Sarvodaya team went as far as to say that if the project had ***only*** produced training outcomes for CRPOs it could have been counted as a success.

In Central and Uva Provinces, UNICEF's approach reached young people in secondary schools as well as Children's and Youth Clubs through an initiative led primarily by the Central Province Department of Education. In contrast to Sarvodaya's training approach that brought in subject matter experts from Hashtag Generation, provincial education officials undertook a process to adapt learning modules developed by Hashtag to ensure that content was tailored to their contexts and to drive local ownership. Representatives from the education sector, District Secretariats, the Youth Service Council, the National Child Protection Authority, Probation and Youth Services and civil society partners were brought together from June to August of 2023 to develop five learning modules that would form the basis for three layers of cascading training: a three-day initiative for 12 Master Trainers per district from August to September 2023[38]; two-day district-level workshops conducted by Master Trainers for 105 resource people from schools and Clubs from October to November 2023; and training lasting between two hours to one day of upper level school prefects, teachers and parents as well as Club members by resource people from November through early 2024.[39] Provincial Departments of Education oversaw implementation of training in schools, while T-Field was brought in to implement and oversee training of and by CRPOs and YSOs. Starting so late in the project lifecycle, and accounting for a participatory process of module development, the time dedicated to actual training efforts with students, young people and resource staff was limited. In total, the Central Province Department of Education where field work took place reported that approximately 3000 students were trained by March 2024.[40]

As work got underway in secondary schools in Central Province, T-Field turned to launching its work with local CRPOs and YSOs but quickly discovered that they needed to pause direct work with local officers to

[36] During the project lifecycle, Sarvodaya received two contracts, an initial six-month contract that launched in February 2023 followed by a second contract of four months that ended on 31 March 2024.

[37] Sarvodaya staff, as well, reportedly raised their own awareness and skill in countering hate speech by attending training meant for others, which was seen as an unintended benefit or outcome.

[38] T-Field and Central Province Department of Education noted that additional training workshops were added as it became clear that Master trainers would need more support to effectively roll out the training. In the end, Master trainers received three two-day workshops, which also afforded important feedback on the content of the training modules before launching training programs with young people.

[39] While training was apparently planned to end in December 2023, some students reported that they had participated in training within the last month of the project, with one training occurring the week before the evaluation team visited the field site.

[40] UNICEF's training approach differed between larger and smaller schools, with larger schools pursuing a more focused targeting by training school prefects and class monitors, while smaller schools sought a class-wide approach for grades 10 and 11.

gain support from district and provincial level authorities to whom the officers report. While the evaluation team acknowledges that UNICEF field staff engaged provincial level officials, those efforts appear to have been inadequate to fully gain local officials' endorsement, thus compelling T-Field to also engage in advocacy and sensitization. The domino effect of needing to take time for unplanned advocacy with higher level authorities ultimately took precious time away from engagement with CRPO and YSO officers – and by extension Club-affiliated young people, who were unable to benefit from the longer period of engagement offered to their Northern and Eastern Province counterparts. The time bind that T-Field found itself in was exacerbated by their short contract that began only in April 2023 and left T-Field only six months to plan for module development with provincial officials, gain the support of CRPO and YSO hierarchies, conduct training with resource staff and Club members and implement small scale social impact projects through Club members.[41]

Learning Outcomes

By and large, project stakeholders, including secondary students, Children's and Youth Club members, CRPOs and YSOs report greater awareness about the negative impact of hate speech and can demonstrate knowledge about how to address it. Youth stakeholders across all organizational types typically reported that training sessions heightened their awareness of hate speech, whereas CRPOs and YSOs noted that training made them aware of problems they did not know existed prior to the training.

Focus Group Discussions with young people probed several angles with respect to hate speech, including young people's definitions of what constitutes hate speech, examples of it from real life, as well as testing whether they viewed various statements as examples of hate speech and what to do if one encounters it on- or offline. Discussions revealed that young women and men from all parts of the country have witnessed, been targets of, or commits acts of hate speech related to a wide variety of issues, including ethno-religious hate speech, jokes based on stereotypes, derogatory name calling, and cyber GBV and body shaming. Young people associated with Clubs that received more training were also seen to be more forthcoming and thoughtful about their experiences, compared to the relatively less engaged participation of students, many of whom received only several hours of training. And while young people from all institutional settings generally were able to identify statements of hate speech against religious and ethnic groups, respondents were divided about cards that called for adherence to traditional gender roles and subjugated women or sexual minorities, an outcome that was observed as much among young women as it was among young men.

The most widely cited learning outcome reported by young people across all institutional settings was that they are now more prone to fact check posts before sharing them, which was raised by 29% of FGD respondents and equally spread between young women and men. Among the next most common behavior change is a greater tendency to actively counter negative posts, a change exhibited in 60% more young women than young men.[42] Other prevalent learning outcomes include greater sensitivity and empathy toward others, blocking or reporting harmful speech, and ceasing online bullying, swearing or body shaming. All of these outcomes point to an awareness among young people of their own, individual responsibility to not hurt others and how to respond to others that do that has been fostered by their

[41] T-Field was later granted a 5-month extension, bringing their total contract length to nine months. Our critique of T-Field's short contract, however, is not offset by this extension, since T-Field was compelled from the outset to draft and launch a multi-phased work plan with only six months.

[42] Responses such as this were part of a trend observed by the evaluation team of young women declaring they would defend others. We address this further in the section Conflict Sensitivity and Risk.

exposure to training. The degree to which these learning effects have extended beyond the deeply personal level, however, is questionable and will be examined more closely in the section on Impact.

One group of duty bearers that plays an important role in students' lives are teachers and school administrators, who were targeted as resource people tapped to train secondary students on hate speech and digital citizenship. Given the pivotal role these duty bearers play in learning outcomes for students, inconsistent access to or uptake by resource faculty led to scattered outcomes on the ground. Participating schools were required to designate two educators within their faculties to attend training designed for school resource people, based upon which they would provide training to students, other teachers, and parents. The evaluation team learned of only two instances of teachers training other duty bearers: one example of a two-hour seminar with teachers, prefects and parents of prefects and another instance of a faculty resource person extending training to teachers in four neighboring schools that were not targeted by the project. Reports on the effectiveness of the training given to resource people, moreover, were mixed. In one school, where teachers from relevant subjects were tapped as school trainers, participants credited the training with providing them with the knowledge and tools to impart lessons to students. In schools that sent faculty from less relevant subjects to be trained, resource people were less able to clearly articulate the importance of combating hate speech and reported that the training they received was too brief and that they lacked instruction on how they should disseminate the training to others outside of the prefects targeted by the project. The importance of ensuring that teachers are sensitized to the problem of hate speech was underscored in a KII with a Human Resource Officer from Jaffna, who noted that he self-initiated training with teachers after observing teachers using derogatory terms with students in the classroom. The impactful role of teacher behavior was also raised by 29% of the young people who participated in FDGs, with some citing positive cases where a teacher's understanding had helped ease tensions and bullying, while others indicating incidents of teachers using harmful speech in the classroom. In Jaffna, for example, Children's Club members recalled the refusal of teachers to attend classes with Muslim students following the Easter Sunday attacks until students held a meeting with teachers to allay their fears and induce them back to the classroom. Stark examples such as this reveal the explicit and overt discrimination and hate speech that occurs even in institutional spaces that should be protective of students and highlights the need for greater efforts to reach teachers even as training continues with students. Addressing behavior and knowledge gaps such as this will become even more imperative with the nation-wide roll out of the Ministry of Education's curriculum reforms over the coming year.

In addition to training, young men and women in targeted Children's and Youth Clubs were also provided with modest resources to implement small-scale social impact projects. As noted in the section on Relevance, small scale projects were rolled out within Clubs that had completed training. In the end, vastly more social impact projects were launched than what had been anticipated – despite the results framework setting a target of 10 projects across all programming locations, 120 projects were implemented through support by T-Field, while an additional five were implemented through support from Sarvodaya. While the criteria for project selection focused on a tangible output that would benefit the whole community, the projects were otherwise disassociated with other aspects of this project. Given this disassociation, their effectiveness at consolidating learning or translating lessons learned about hate speech and social cohesion into practice seem minimal. This conclusion is supported by FGDs with Club members. While 100% of Children's and Youth Club members reported satisfaction or pride with having delivered a resource to their communities, none of the members were able to articulate how projects related to efforts to combat hate speech or support social cohesion and instead focused on their personal growth in planning and executing a project. The projects, moreover, received insufficient amounts of funding to be fully implemented without external support, raising questions about whether project

resources were spread too thinly over too many stakeholders. Attempts to link the projects with messages learned through the training were superficial at best, with hate speech messaging tacked onto most project deliverables in incongruous ways – writing "stop hate speech" on a packet of seeds or taping similar slogans on a library's bookshelf or Club walls. Considering the rush to implement training during the last six months of programming and reported deficiencies in implementing partners' budgets to reach remote Club locations, the evaluation believes that scarce resources of time and money could have been better directed away from small social impact projects and toward additional training on hate speech with CRPOs, YSO and Club members.

With respect to CRPOs and YSOs, the evaluation team found reported evidence of heightened awareness in all three project locations visited by the team. 100% of Officers report increased personal awareness and more thoughtfulness in their online actions, similar to the individual level and deeply personal changes noted among young people. Implementing partners Sarvodaya and T-Field, however, both observed that they lacked the monitoring tools to be able to capture specific learning or changed attitudes and behaviors of Officers, a shortcoming that limits the evaluability of this project.

III.B.2. Impact and Sustainability

Q13. If the project effectively achieved its main outcomes related to combating hate speech, did that success contribute to greater social cohesion?

Q14. What are the most significant changes achieved through this project for youth?

Q15. Have durable, long-term processes, structures, institutions, and capacities for peacebuilding been created?

Q16. How resilient are the project's outcomes – including partnerships and relationships – and will they sustain beyond the project lifespan? Has the intervention addressed the role of "spoilers," i.e. those who benefit from on-going conflict?

Finding 6. Embedding work within systems – Codes of Practice, journalism professionalization plans within the Ministry of Mass Media, the ongoing curriculum reforms undertaken by the Ministry of Education and the existing national network of CRPOs and YSOs – offers the best opportunity to realize nation-wide, sustained change. While the Effectiveness section highlights that many of the project's effects were limited to individual learning and personal growth, project initiatives that showed the greatest potential for more impactful and sustained change are those that worked through systems. While the evaluation team is referring to formal and informal institutional systems present in Sri Lanka, the benefits of "systems thinking" for generating meaningful change has been widely noted within the peacebuilding and development fields.[43] Taking such an approach enables initiatives to think more holistically about the various factors contributing to a problem and the actors that produce and reproduce them. The evaluation team observed evidence that actions within the project that targeted drivers of hate speech in this more holistic way show early signs of generating meaningful change.

The dedication of the Ministry of Mass Media, through its partnership with UNDP, Factum and CIJ, to support greater professionalism within broadcast and print media, displays promising outcomes for longer term, industry-wide change. Together with the training program for provincial journalists conducted through this project, the Ministry's support for an independent, voluntary National Media

[43] See, for example, Designing Strategic Initiatives to Impact Conflict Systems: Systems Approaches to Peacebuilding. A Resource manual. CDA Collaborative Learning, or Robert Ricigliano's "A Systems Approach to Peacebuilding," Conciliation Resources.

Policy – if adopted – promises to safeguard freedom of expression while offering guardrails for curbing incendiary reporting. While the policy applies to all media professionals, including company owners and managers, as noted by provincial journalists with whom we spoke, current market incentive structures will limit the policy's impact unless priorities can be reordered and public appreciation for ethical, fact-based journalism takes root. Training that was begun with provincial journalists through this project, moreover, should be viewed as an introduction and not the end of a learning process. Learning gains need to be reinforced through access to comprehensive, regular training.

The CoP and engagement of its members to produce a Code of Practice for social media companies as well as engagement with the Ministry of Mass Media have played important roles in securing gender-responsive commitments to combat online hate speech and advancing professionalism within traditional print and broadcast media. This work has been aided by the various data collection and analysis efforts funded by this project, which have filled important information gaps and shaped UN strategic priorities and commitment to ensure the UN's continued engagement through 2027. **Specifically, regarding the Code of Practice for social media companies, while it is too soon to note specific changes at the impact level generated by the Code, the project has done well to advance these commitments in the relatively limited time of the project lifecycle.** Ultimately, close monitoring of social media companies' application of the principles and actions within the Code will demonstrate whether it has been an effective lever to reduce online hate speech or if other approaches are needed.

It should be recognized that **initiatives like the Code of Practice were made possible through the coordinated efforts of the Community of Practice managed by UNDP.** A large part of the CoP's success has been due to its institutionalization-with-a-light-touch through regular meetings, meeting summaries, reference documents outlining various initiatives led by group members, and a document repository. These measures have grounded the work of the CoP and enabled coordination among actors to promote evidence-based solutions. Responsibility for ensuring the smooth functioning of these measures, however, has fallen to UNDP and has, at times, presented a substantial burden. While the CoP has been impactful in advocating for better moderation of online hate speech and the evaluation team believes it should be continued, we conclude that it does not appear to be sustainable in its current form and offer a recommendation below for launching a pathway for CoP self-sufficiency.

UNICEF is to be commended for having seized the opportunity to influence the structure and content of curriculum reforms undertaken by the Ministry of Education and the National Education Institute. Such partnerships ensure that curricula that will be in place for years to come will reflect the underlying commitments of the project related to good digital citizenship, tolerance, and critical thinking skills. While recognizing that this approach also resulted in slower implementation and less control over the approved content, by investing in partnerships and working to effect systemic change, UNICEF has secured its best chance at locking in mandatory curriculum on digital citizenship while influencing the way that critical thinking and topics related to tolerance and inclusivity are mainstreamed. If successful, these changes will reach all students in the formal education system for years to come – an effect that goes well beyond what this project could have otherwise managed. As of the drafting of this evaluation report, the reform efforts are still in progress, rendering it too early to tell just how many of these goals UNICEF will be able to achieve in the end. As with implementation of the Code of Practice for social media companies, UNICEF is urged to work with the Ministry of Education to ensure ongoing and proper training of educators on the new curricula and to substantively monitor its impact, findings that are echoed in the forthcoming Review of the National Action Plan on Education for Social Cohesion. To do so, UNICEF will need to not only track how many students have been exposed to the new content, but whether and how they are

applying their new knowledge and skills, as well as extend the focus of their work to a range of duty bearers – such as school administrators, teachers, parents and community leaders – that influence them.

Finding 7. In contrast to systems approaches noted above, efforts to produce interesting content through print, broadcast, and social media – while compelling and relevant in their own right – could have been more impactful if coordinated. The different outputs of the project produced a range of interesting, professional, and relevant outputs. By our count, these include:

- 20 stories for both print and broadcast media and six in depth investigative stories that were published in papers and online through work with Center for investigative Journalism,
- Social media campaigns by five social media influencers supported by CEJ that reached an audience of about 200000,
- Six public awareness television commercials and 18 interviews for primetime produced through Time and Space,
- Approximately 12 story pathways on social cohesion to combat online hate speech through 4K,
- Three 10-minute video clips confronting historical inaccuracies about Black July, the early adoption of the Sinhalese-only language policy, and the southern rebellion by Project 72, and
- 10 social media campaigns launched by university students supported by NCEASL.

These campaigns, articles, interviews, and commercials ran according to their own timelines, without connecting to the efforts of others. As a result, while some managed to reach broad audiences, the project as a whole was limited in generating and sustaining wider public awareness. This is a missed opportunity by the project that could have been avoided if the various journalists, television producers and social media content producers had been brought together.

Finding 8. Small scale social impact projects need to be planned in such a way that they clearly consolidate learning or skills building developed through the project and advance, or at least are coherent with, broader project aims. Within this project, 125 small scale social impact projects were awarded to Children's and Youth Clubs that had completed anti-hate speech training, each one with its own goals being implemented in isolation from others. The target of 10 small scale projects noted in the results framework raises questions whether the original purpose and ambition of these initiatives was to have drawn young people from diverse backgrounds together to jointly implement community-based actions. Having planned fewer social actions that brought Club members from diverse backgrounds and parts of the country together could have better underscored lessons in cross-cultural understanding and tolerance that were the cornerstones of the training to which they were exposed. It also would have addressed a clear interest on the part of young people to engage others from other communities, a curiosity and openness that should be nurtured. To be impactful, small-scale initiatives such as this require a clear framework, with overarching goals to which projects should contribute and must meaningfully connect to outcomes achieved in other parts of the project. Without such a framework, small scale initiatives will not realize their potential and instead will be merely one-off, nice experiences that do not clearly advance project aims.

III.B.3 Catalytic effects

Although not a part of the original list of evaluation questions, the evaluation team added examination of the project's catalytic effects to this evaluation, given the importance of this criterion for the Peacebuilding Fund. For the PBF, an investment is considered to have a catalytic effect when "it creates momentum for the removal of barriers that block longer-term/sustained peacebuilding support or for

the engagement of additional stakeholders in existing peacebuilding efforts. In addition, a PBF investment is also considered to have a catalytic effect when it facilitates new initiatives that bring in new resources to peacebuilding priorities in the country."[44] Given this definition, the team has examined two aspects of the project's catalytic nature: financial and programmatic effects.

Financially, **the project catalyzed $207,500 in additional external funding**, including $16,500 in 2022 from OHCHR to add monitoring of discrimination of minority groups to the project's existing social media monitoring themes, as well as $191,000 from the United States Department of State, also related to augmented data collection and monitoring, from 2022 to 2024. While a number of proposals for follow-on funding are currently being finalized or under review by development partners, as of the drafting of this report, the above reflects funding that has been concretely catalyzed.

Programmatically, the project has contributed to an opening of two areas that had not been anticipated by the core aims of the project. Some of these entry points have shaped broad strategic frameworks or policy change, while others appear to be more discreetly programmatic.

Specifically, by highlighting the importance of addressing on- and offline hate speech, the **project is said to have shaped the priorities in two important UN frameworks, the UN SDCF and UNDP's SDG 16 Flagship Program.** With respect to the UN SDCF, addressing hate speech is among the key assumptions driving the theory of change for Outcome 5 on social cohesion, which aspires to see Sri Lankans living in a society that is more cohesive and strengthened by transformative and rights-based processes and mechanisms. The approach to realizing these aims commits the UN in Sri Lanka to continue investing in early warning mechanisms and the data on social and traditional media necessary to animate it, outputs directly related to the project. Similarly, within UNDP's SDG 16 Flagship program, the social cohesion pillar specifically identifies investing in early warning systems and promoting information integrity and media governance among its four approaches. Reporting in the project's Annual Reports of 2022 highlighted the role the project played in securing these commitments and was reaffirmed in interviews with the former National PDA and UNDP program staff.

Regarding the education sector, the evaluation team observed one potential catalytic effect that was not foreseen by but is a direct outcome of the project. In Central Province, one of two pilot provinces where modules on good citizenship and combating hate speech were developed and piloted, the **provincial Department of Education has plans to utilize its own resources to partner with the University of Peradeniya to refine the curricula and pilot a new, more intensive course with provincial ninth graders in select schools, including closer accompaniment by pedagogical researchers associated with the university**.[45] This pilot is expected to reach 3,000 students with regular coursework over a six to seven month span and builds directly off of the modules developed under this project. Both the Deputy Provincial Director and Deputy Director within the Central Province Department of Education attributed the Department's initial awareness and continuing commitment to addressing hate speech to the project.[46] The evaluation team cautions, however, that the Central Province pilot is still in the planning stage, meaning the evaluation team can be only cautiously optimistic about this finding.

---

[44] Bautista Logioco, PBF Catalytic Effects Guidelines, Peacebuilding Support Office (March 2024)

[45] Once fully formulated with a cost plan, the Central Province pilot will be an additional contribution to the project's catalyzed finances. As of the drafting of this report, however, the Department did not have a budget estimate for the pilot nor a dedicated funding stream.

[46] While the evaluation team notes that this is a potential catalytic effect, we are also concerned about the potential for conflicts between the Provincial Department of Education's efforts to develop and deploy training

III.C. MEANS

III.C. 1. Efficiency

Q10. Have financial and human resources been allocated sufficiently and strategically to achieve project outcomes?

Q11. Have project outputs been delivered in a timely manner?

Q12. How has the project ensured value for money through the use of the most cost-efficient methodologies to achieve planned results? Were implementing partnership arrangements the most efficient means of implementing the activities?

Finding 9. The high volume and short length of partnership, individual and service contracts contributed to inefficiencies.

As noted elsewhere, the project was marked by a large number of various kinds of implementing partners and contractors. Table 1 below includes a summary of these different types of partners and the roles or deliverables they were tasked with.

Type of Implementing Partner	Role/Type of Collaboration
Divisional Secretariats	Led division-level steering committees and rolled out prevention programs related to Children's and Youth Clubs
National Institute of Education	Implemented "digital citizenship skills" with secondary students through strengthening media subjects within schools.
Ministry of Mass Media	Supported the development of training materials and media guidelines for provincial journalists; supported policy changes for media companies.
Ministry of Education	Developed new curricula, including modules on digital citizenship and critical thinking related to new mandatory media studies course.
Provincial Departments of Education	Led the development of training materials and coordinated targeted schools in Uva and Central Province.
Sarvodaya and T-Field (CSOs)	Provided training/orientation to young people in Children's and Youth Clubs as well as religious leaders (T-Field)
Centre for Equality and Justice and Search for Common Ground	Provided training to women's organizations and female social media influencers on combatting hate speech and networking.
Hashtag Generation, Deloitte, National Christian Evangelical Alliance, Sri Lanka, CPA, Sec Dev, Factum	Collected and analysed data on hate speech across a range of platforms and targets/themes; produced regular reports at varying intervals; led

modules and those that will be rolled out by the Ministry of Education's curriculum reforms. These concerns are addressed in more detail in the section on Coherence.

	training and policy development initiatives (Hashtag Generation, SecDev and Factum).
Broadcast, print and social media companies and social media influencers	Provided training to social media influencers, journalists, media company owners and editors and produced positive content on television, print and social media.

Within the project documentation, there is some confusion about the number of implementing partners, with the November 2023 PBF annual report indicating 13 implementing partners, which appears to omit an additional three for which the evaluation team has full contract information, and which had appeared in the November 2022 PBF annual report. In addition, the evaluation team received contracts for six others that do not appear in PBF reports as partners. Four other organizations produced a study or other deliverable but neither appeared in PBF reports nor were contracts made available. Confusion over the number of implementing partners is perhaps a result of the sheer number of them: the evaluation team counts 26 contracts of varying type,[47] that were executed in relation to this $3 million, 27-month long project.

Of the 26 contracts, 19 contracts of differing type issued by UNDP and UNICEF to support the project were made available to the evaluation team, included as Annex E to this report. A full budget analysis of the project is also included in Annex F. Based on these 19 contracts, the average contract length was 7.2 months, with only one contract lasting 12 months, which is the longest contract, and the shortest being 1.5 months. According to UNDP, several organizations were granted short term contracts because they were piloting new approaches or they were new partners and so the six-month time frame was a way to test their fitness and a means for UNDP to manage risks.[48] This explanation, while plausible for some partners, is not entirely satisfactory for several reasons. First, some partners such as Hashtag Generation and Search for Common Ground, both of which received eight-month contracts, were known to UNDP prior to the project and have established track records that should have induced confidence in their technical and financial responsibility. Second, since 2005, UNICEF, UNFPA and UNDP have been implementing the Harmonized Approach to Cash Transfers (HACT), a risk management system for reviewing organizational technical and financial strength, based on which they set funding limits and transfer modalities for each implementing partner to ensure that funds have the greatest chance of being used in an appropriate and effective manner. According to UNDP policy, while a HACT review is not required of partners receiving less than $150,000, it is strongly encouraged.[49]

The high volume of relatively short-term contracts was cited as a constraint by 50% of implementing partners reached by the evaluation team. The most common frustration from partners was that their work had only just started to yield meaningful interaction or interesting deliverables when it was time to close the project. As one KII participant put it, “We’ve just finished building the foundation and have laid only the first brick.” While short contracts were cited across all different types of partners, those producing content or news articles – which need time to develop – were most vocal about time constraints.

[47] This number does not take into account potential contracts with Rata/lk, Dreamspace Academy, MultiTech and Marga Institute, which appears to have completed work on this project but for which the team did not receive contracts, as well as an additional seven individual contracts, including the two under which the evaluation team was hired.

[48] Online interview with Fathima Rajap and Charudaththa Ekanaya, March 28, 2024.

[49] Financial Management and Implementation Modalities: Harmonized Approach to Cash Transfers, UNDP Guidance

Compounding the rush to deliver within short timeframes, review processes by UNDP to green light some of the broadcasting or social media content were seen to be barriers to timely implementation. The evaluation team recognizes the need for UNDP's internal review processes, especially given the sensitive nature of countering hate speech and promoting positive narratives. The delay of two or three weeks to support such review and approval processes, however, had a disproportionate effect on timely implementation for contracts lasting only six to eight months. In these cases, longer contracts could have better tolerated longer content approval processes.

Finding 10. The project's lack of cohesion, given the independent nature of the project's two outcomes, reduced opportunities for UNDP and UNICEF to maximize efficiency.

As indicated in the sections on Cohesion and Project Management, the project missed opportunities to draw its two outcomes into a unified approach and for implementing partners to learn from and take advantage of resources and knowledge produced by others connected to the project. The evaluation team believe this is perhaps due to the combined effect of the large number of contracts UNDP and UNICEF set themselves up to manage, and the comparatively lower level of engagement of UNICEF within coordination meetings. Lack of coordination and sharing of systems, skills and information across partners caused partners to be unaware of each other.[50] T-Field, which received a six-month long contract, for example, reported that before launching their work, they spent several weeks conducting a "background study" to familiarize themselves with the issue of hate speech, since this was a new topic for their organization. They discovered the work of Hashtag Generation on their own, after searching on the internet, instead of being introduced to their work and that of others engaged in delivering on Outcome 1. Missed opportunities to share knowledge and resources created inefficiencies in delivery.

The team also notes that both UNDP and UNICEF received similar amounts of project funding, $80,000 and $70,000 respectively, to help religious leaders build their skills in identifying and responding to hate speech. UNDP, however, is not indicated in the description of Outcome 2 or this specific output as having an organizational advantage with respect to work with religious leaders and never appears to have spent any of the $80,000 allocated to it for this activity. In fact, of the $150,000 allocated for work on this output, only $45,000 was expended by the November 2023 report, all of it by UNICEF through its work with T-Field and Sarvodaya.

Finding 11. Relatively low amounts of the budget were reserved for substantive implementation of the project's two outcomes.

Examining the budget lines within the project reveals that non-programmatic budget lines represent 42.6% of the total budget, resulting in a substantially reduced budget available for programming. Of UNICEF's $1 million allocation, only $573,000, or 57% of the total allocation, went to programming, with over $145,000 dedicated to monitoring, documentation and visibility. While the various innovations the project was seeking to make in terms of evidence generation and its approach to reducing hate speech perhaps warranted a sizable portion of funds to capture key evidence about project effectiveness, as will be discussed in findings on Project Management and Monitoring, the evaluation team did not benefit from improved project monitoring of higher order effects related to project activities. Implementing partners, moreover, cited how the lack of funds due to low value contracts led to funding shortfalls in the implementation of core project activities such as training and field data collection, funding gaps they

[50] Cite how many times this came up in interviews.

ultimately filled through their own limited resources. UNICEF programming team members explained that much of this budget line supported spot checks and field visits by program team members, oversight activities that became increasingly expensive as the economic crisis triggered price hikes for transportation and food costs. While staffing, oversight and visits for the purpose of project management are reasonable project expenditures, the evaluation team notes the disproportionate distribution of funds across the project budget and the negative impact this had on project partners in the field.

Finding 12. UNDP and UNICEF's project teams were alert to some of the emergent contextual challenges that were causing delays, while others failed to trigger timely adaptation.

The project's approved budget of $3 million, was transferred in two tranches: the first on 3 December 2021, with $1,400,000 going to UNDP and $700,000 to UNICEF. By the summer 2023 when programming teams requested a No Cost Extension (NCE), UNDP had expended $1,184,473.19, or 85% of its first tranche. UNICEF, on the other hand, had expended only $397,414, or $56% of its much smaller first tranche, citing delays with changes among government partners and competing priorities among government and civil society partners who were responding to humanitarian needs triggered by the economic crisis. The MPTFO transferred the second tranche on 22 August 2023, at the same time that an exceptional 9-month NCE was granted, extending the project end date from 1 June 2023 to 1 March 2024. While the evaluation team appreciates the real obstacles to implementation posed by the context and notes UNICEF's reinvigorated implementation plan starting around March and April 2023, with a number of new contracts issued to advance work at the provincial level, UNICEF's description within reporting and the NCE of how it planned to get back on track was vague and does not address how compressing 18 months' worth of work into nine months might affect training and awareness raising outcomes on the ground.

III.C.2. Gender Responsiveness and Human Rights

Q18. To what extent did the project actively identify and include marginalized and excluded groups and ensure disability inclusion during the design and implementation of activities?

Q19. To what extent did program targeting reproduce or overcome traditional norms or practices that exclude key stakeholders such as women, young people, persons with disabilities, and stakeholders from a wide range of gender identities? Were special measures taken to accommodate specific groups to ensure their participation?

Q20. Were commitments made to gender equality and women's empowerment in the proposal, i.e. Gender Marker 2, realized in implementation?

Finding 13. Evidence largely supports the substantive inclusion of many traditionally excluded groups within the implementation of activities and deliverables, with some results capturing individual change while others have managed to extend beyond the individual to the community or institutional level. Nonetheless, there are notable areas for improvement.

While the overall aim of the project is to strengthen social cohesion, the means through which social cohesion is to be bolstered is through reducing publicly available narratives that frequently fuel human rights infractions, including mis- and disinformation, hate speech, and incitement to violence for reasons related to ethnic, religious and sexual identities. Specific project activities related to professionalization of traditional media, encouraging a Code of Practice for social media companies, and training programs in schools, Children's and Youth Clubs, and with religious leaders all have elements of reducing intolerance of others' differences, which should have a net positive result for excluded groups' abilities to exercise

basic human rights. OHCHR's contribution of additional funding to expand the topics being monitored, as noted in the section on Catalytic Effects, also speaks to the ways in which the project's aims and approach address human rights.

While it is difficult to capture the aggregate effect of the project's results on gender equality and women's empowerment through project implementation specifically, there are a number of positive results that show promising indications for broader change. For example, UNDP, with support from the RCO, committed to generating an evidence base to support national stakeholders to respond to hate speech, including gendered hate speech, and to advocate for appropriate policy and legal frameworks within Outcome 1. To implement this outcome, UNDP contracted with five organizations to monitor social media platforms for different kinds of hate speech and produce reports of varying periodicity. A random sample of these organizations' reports indicates that all but one provided clear and at times powerful reporting on hate speech against women and the LGBTQI+ community that enabled stakeholders to take action. Reporting by Hashtag Generation, Factum and Deloitte, for example, revealed that women were targeted in 53% of hate speech picked up by the monitoring mechanisms between January and April 2023, a perceptible rise over the 25% noted from 2021 in the project document. In response, national stakeholders engaged in the CoP and development partners have been able to advocate for gender provisions within a revised Code of Practice for media and tech companies developed through this project, which – if fully adopted – could have far reaching effects.

At a more personal level, as noted above, training provided through the Center for Equality and Justice supported 30 young social media content producers, including those identifying as women and as LGBTQI+, to recognize cyber sexual and gender-based violence and counter it. The training resulted in five online campaigns that ran from March to April 2023 and reached over 200,000 viewers, most of whom liked or commented positively on the content. The experience of going through the training together and learning about the difficulties or attacks faced by others subsequently inspired participants to create their own support network via WhatsApp, through which they collaborate and mobilize to counter disinformation and cyber GBV, an encouraging sign of the training's sustainability and relevance, despite the fact that CEJ's contract lasted only eight months and a had relatively low budget.[51]

One line of work where the project fell short of expectations is within output 2.1, for which $150,000 was allocated to support religious leaders to understand and effectively respond to hate speech. The results framework indicates that the project team intended that female religious leaders would constitute between 20 to 30% of those targeted. Reporting on this output, however, not only shows that execution rates were well below expectations, with only $45,000 expended as of November 2023, but that both outreach in the project approach and monitoring of the activities was gender blind. The evaluation team notes this shortcoming in detail because the lower budget execution rate has contributed to the downgrading of the project from GM2 to GM1. For more details on implementation of the budget, see Finding 2, below.

Another area where project deliverables and results could have been more gender-sensitive is in the needs assessment of provincial journalists done by Verite through output 2.2, which is gender blind

[51] While reporting in the PBF periodic progress reports asserted that 50 young women were trained, in an in-depth interview, CEJ confirmed this number to be 30. Interview with Suhanya Joseph de Saram, project officer and Shamala Gomex, executive director, 28 March 2024.

despite notable barriers faced by female journalists.[52] A review of the final report shows that participation of women was not indicated within the KIIs or FGDs conducted to inform the report nor was any recommendation tailored to the specific needs of these groups. This gender blindness carried over into reporting on the training provided to provincial journalists by the Ministry of Mass Media and within the periodic progress reports provided to PBF.

Evaluation questions 18 and 19 called for an assessment of the degree of inclusion of a broad list of often excluded segments of the population, including People with Disabilities (PWD). The evaluation team notes that PWD were not identified in the project document nor captured as an intentional group with respect target setting for indicators within the results framework or included in the many data collection exercises that constituted key outputs within Outcome 1. Despite that, reporting against indicator 2.3.1 in Outcome 2 on training and awareness raising with young people engaged in Children's and Youth Clubs did reflect a breakdown of the number of young people with disabilities who received an orientation to the project (28 out of 2,765 total participants). Although this number is small relative to the total number of youth reached, the evaluation team does not view this as a shortcoming within the project. Within peacebuilding projects, programmatic approaches – including the selection of stakeholder populations – derives in large part from a context or conflict analysis. Through this analysis, key actors and factors for conflict and peace are identified and project pathways for achieving overall goals related to a peacebuilding aim are articulated through a theory of change. In this sense, peacebuilding programming must be assessed through a slightly different lens than evaluations of development programming, which would plausibly key results more intentionally to a poverty reduction or "Leave No One Behind" strategy. Instead, actors who are tagged for inclusion within a peacebuilding project should have a particular role, need or comparative advantage with respect to the conflict factor being addressed. When considering combatting online hate speech and supporting social cohesion in Sri Lanka, the evaluation team is not of the opinion that PWD are a more relevant constituency or lever for effecting positive change than other stakeholder groups included in the project design. The already large number of stakeholder groups, moreover – provincial journalists, social media influencers, students, youth in clubs, religious leaders – within an 18-month long, $3 million project argues for reducing the number of stakeholders not expanding it.

Finding 14. While the project document is clear and detailed about the rationale for its commitments to ensuring an approach grounded in gender equality and women's empowerment, its results framework and reporting are less clear and at times stretch definitions, while financial execution rates fall below Gender Marker 2 thresholds.

The PBF has issued long-standing guidance on how to apply its gender marker (GM) to be able to substantively and financially track how commitments made within the project document are realized in practice.[53] From the outset, the project under evaluation was scored as GM 2, which obligates it to dedicate at least 30% of its total budget to activities that contribute to GEWE, and, among other things, to track how GEWE is built into implementation through a mix of indicators that are sex-disaggregated and/or measuring GEWE-specific change.

The project's results framework consists of 23 outcome and output indicators, of which only one, output indicator 1.1.2, is dedicated to measuring a GEWE-specific change. Of the remaining 22 indicators, sex-disaggregation in the baseline and target is explicitly noted in 13 indicators but realized in practice by only

[52] Such risks included security risks to nighttime reporting and misogynistic reporting and editing of content by newsrooms.

[53] PBF Guidance Note on Applying the Gender Marker (2019)

nine. Many of the indicators included within the nine that are sex-disaggregated, however, suffer from being vague or stretching definitions in order to meet a target. To illustrate overly vague language, outcome indicator 1b commits the project team to securing the regular tracking of three new sex-disaggregated indicators on social cohesion by national partners. Reporting on this outcome notes only the number of new indicators without describing the indicators or signaling into which framework(s) they have been added or which national partners are tracking it. Elsewhere within the results framework, indicators that are meant to highlight how many women advocates or women led CSOs have conducted a given advocacy or anti-hate speech action are reported in ways that stretch the definition. For example, indicator 1c in the results framework measures the number, diversity and type of national stakeholders using evidence to address hate speech or support social cohesion. Reporting on this indicator, however, highlights only that two of the project's implementing partners were contracted to train women and women's organizations to engage in advocacy or advance positive narratives.[54] To substantively report on this indicator, the project would need to measure the effects of that training on the female national stakeholders or women-led organizations who have begun utilizing evidence to counter hate speech and promote social cohesion.

In terms of budget expenditure, as of the November 2023 annual report, which is the last date for which verified financial expenditure is available, the amount dedicated to GEWE fell to only 26% of the total budget. Preliminary reporting in June 2024 indicates an even lower final expenditure of only 15% of the total budget.[55] If this budget distribution persists, the project will be downgraded to GM1 due to its failure to meet the financial threshold for GM2, which would be a shame since, as noted above, the project has managed to produce a number of advances or positive effects with respect to GEWE.

III.C.3. Conflict Sensitivity and Risk Management

Q23. To what extent were the risks associated with conflict dynamics factored into the project's approach? Were risks continuously monitored and responded to throughout implementation?

Q24. Was the risk management approach sensitive to the differential risks of women, men, girls, boys and marginalized or hard to reach populations?

Finding 15. The risk management matrix developed at the proposal stage focused on risks to implementation at the expense of considering contextual political or security risks or specific risks to stakeholders through their participation in the project. Additionally, monitoring of these risks appears to have been uneven during implementation.

The project document identified six risks associated with the project. All but two of these highlighted risks to timely project implementation, including risks that the government's response to the COVID-19 pandemic would limit movement and convening capability and that key duty bearers may not have sufficient buy-in. While risks to timely implementation are certainly valid considerations to be tracked,

[54] Even with this definition, however, progress reporting still comes up short by dubiously labeling project partners as national women-led organizations. The first one, Search for Common Ground is an international NGO dedicated to youth empowerment and social cohesion that is labeled in project reports as a national women-led organization. The second organization, the National Christian Evangelical Alliance of Sri Lanka, is listed as a women led CSO in the project document.

[55] Within the project document, the amount dedicated to GEWE was $896,017, whereas reporting in the June 2024 PBF reporting indicates actual expenditure of only $450,888.47.

given that this project was meant to tackle highly sensitive matters that touch upon longstanding social divisions and sources of conflict, the risk matrix could have gone further to unpack how these may present risks to the achievement of project outcomes as well as participants. Some of the specific types of risk we refer to are detailed below in Findings 17 and 18.

In terms of how well even those risks that were identified were monitored and offset during implementation, the team finds uneven performance that maps directly onto the differences in implementing timeliness between UNDP and UNICEF. For example, the second risk noted within the project's risk matrix is that the COVI-19 pandemic will prevent the physical gathering of members of Children's and Youth Clubs, work led by UNICEF. UNICEF's risk management approach was to have been the early development of training tools that could be deployed through online formats in order to keep to the project timeline. UNICEF concluded their assessment of this risk with "This will not require large gatherings of people or children and can move forward without delay." In practice, however, the hiring of implementing partners and development of training materials did not occur until month 16 of implementation, two months before the project's original end date. Thus, the evaluation team assesses that at least for project activities targeting students and young people, more effective monitoring and risk mitigation could have prevented some of the worst of implementation delays.

Finding 16. Project implementation was largely sensitive to inter-communal differences and context factors that trigger or perpetuate tension and risks for stakeholders on the ground.

16.a. Caste. While not identified in the project document itself, within field data collection, caste emerged strongly as a deep division in society that contributes to the spread of hate speech. For example, caste was raised as a source of social division directly by young people in four FGDs and indirectly in an additional two. Caste is an issue that requires careful handling due to its deep entrenchment in cultural practices and religious beliefs. Challenging caste, then, may be seen as similar to challenging the tenets of a religion. In Jaffna, UNICEF's implementing partner Sarvodaya has been cautious not to mix children from different caste groups during training programs, a practice based on requests from CRPOs, in order not to introduce tension and division among participants. This approach can be seen as responsive to conflict dynamics as it recognizes that addressing caste issues requires long-term engagement, especially with duty-bearers such as parents, teachers and sports coaches, as well as strategic planning. In sum, ignoring these dynamics in project implementation could have worsened already existing divisions among children rather than nurturing unity and understanding.

Similarly, interviews with government officials indicate that the decision to not introduce topics about caste discrimination was an intentional and practical decision. At present, caste is not easily or openly discussed, even among colleagues, and they fear that raising the topic could provoke tensions. According to government officials, however, discussions on caste should be part of longer-term interventions, and perhaps working with children can produce more promising results than engaging with adults.

The evaluation team also observed evidence that this kind of sensitive approach has also been taken into account within the curriculum reforms and specifically within the development of learning modules on digital citizenship and hate speech. According to officials from the National Institute for Education interviewed for this evaluation, care was taken to avoid using names and occupations that could reinforce caste roles and potentially make some students uncomfortable in caste-conscious communities. Instead,

modules emphasize the importance of refraining from hate speech and harming "other individuals or groups".

16.b. Religion. The project's implementation partners have taken several steps to address the specific concerns of religious groups during the project period, which have been seen as crucial from the point of view of project participants. For example, the National Institute for Education conducted an awareness-raising and feedback workshop for teachers from selected schools in all nine provinces to socialize the draft of the digital citizenship curriculum. During this workshop, teachers identified issues sensitive to their religions and changes were subsequently incorporated into the curriculum revision process, which the evaluation team sees as important given the prominent role of teachers in delivering the training and new lessons. In other project activities, youth participants in FGDs noted with appreciation that implementing partner Sarvodaya adjusted the location of their programming to accommodate the needs of Muslim children during the fasting season. In addition, in their training programs with religious leaders in Jaffna, Sarvodayd invited a Buddhist monk from Naga Viharaya, despite the monk not having a Buddhist community in the area with which to disseminate the messaging. While his presence may have seemed counterintuitive, given the ongoing tension between Tamils and the Sinhala Buddhists in the Northern Province, ensuring the presence of monks within the training and – by extension – the community, rather than excluding them, was seen as important.

16.c. Language – room for improvement. Since independence, language differences have been recognized as one of the root causes of conflicts in Sri Lanka. Consequently, people are highly sensitive to language matters. Lack of mutual intelligibility between Sri Lanka's two main languages – Sinhala and Tamil – furthermore, present practical challenges in day-to-day communication. This issue was raised within KIIs with provincial journalists, who highlighted challenges due to the need for translation within training, as well as FGDs with young people. By way of illustration, in two FGDs in Kandy and Jaffna, young people who had participated in training conducted within Children's and Youth Clubs highlighted instances when sessions were not conducted in their mother tongue. When queried, CRPO officers explained that the youth were presumed to understand the language of instruction, even if it was not their mother tongue. Decisions such as this reflect not only lack of understanding about how to conduct effective training, but illustrate a lack of sensitivity about how language choice can have a chilling effect on participants and prompt them to feel unwelcome.

On the positive side, in addition to the choice of language, the evaluation team observed evidence of conflict sensitivity on the part of both UNDP and UNICEF with respect to the use of certain terms or phrases or of tackling issues more indirectly instead of head on. UNICEF, for example, stressed the importance of introducing the notion of hate speech among Ministry of Education and NIE colleagues using the lever of concepts that already existed within the curriculum, such as good citizenship and critical thinking. UNICEF reports that over time, education sector colleagues have come to accept the more direct term "hate speech" and the role of educational institutions in imparting lessons on how to curb it. Similarly, UNDP implementing partners tasked with producing various kinds of social and traditional media content highlighted guidance by UNDP on Do No Harm. Time and Space, which produced a series of six "infomercials" on hate speech that ran over a three-month period, noted that they revised their approach to the film clips' content from the more direct and explicit approach they had envisioned to one that delivered a clear message but in a more indirect way to avoid triggering communal tensions. Implementing partners 4K and Project 72, which supported anti-hate speech social media campaigns additionally noted that UNDP's decision to not include UN branding within the content produced by the

project helped the creators maintain their independence and legitimacy, which made their messaging more effective. Fact checking and vetting by UNDP of content produced through Project 72, which sought to confront disinformation about historical narratives, was seen as particularly welcome.

Finding 17. Success of several aspects of project implementation on hate speech and addressing localized conflict has prompted participants to take action which has exposed them to potential risks.

The aim of working with certain groups of stakeholders is that they take certain actions or refrain from taking action in public fora, including online spaces. These include social media content producers, local peacebuilders connected to the grassroots incident reporting with CPA, as well as students and young people engaged in participating Clubs. Across all of these stakeholders, the evaluation team has noted evidence of programmatic effectiveness, with registered increases in awareness of the negative repercussions of hate speech and a personal sense of social responsibility when acting online and within communities. As project stakeholders act upon learning outcomes or contribute to successful project outputs, their actions will place them at risk for backlash, physical attacks and cyber bullying or GBV. For social media content producers, who have intentionally placed themselves in the public discourse and actively elicit responses from their followers, these risks are calculated risks, as influencers and content producers by and large are aware of and accustomed to such threats and negative comments. In fact, the WhatsApp group developed by content producers through their work with CEJ, for example, arose in part as a result of a sense of solidarity with others experiencing backlash and bullying. For others in the project, however, most notably the vast majority of young women who received training through secondary schools and Clubs, their participation has instilled in them a sense of personal responsibility that they feel compels them to counter hate speech or harmful posts. Even when probed within FGDs about the potential for online backlash or offline encounters, these young women display a sense of bravado that is both admirable and worrying. Although aspects of the training touch upon online safety and provincial level Department of Education officials stressed to us that students and Club members are not encouraged to push back online, it is somewhat predictable that once alerted to the injustice or harm done by such posts, young people will want to react. It is imperative, then, that if future initiatives are to continue or extend this work – for example, the six-month pilot in Central Province planned by the provincial Department of Education – clearer messaging about the risks young people may face as well as a more robust risk monitoring system is warranted.

The evaluation team encountered another example of how success in producing a project output has triggered a negative reaction on the ground in its data collection with CPA and its network of peacebuilders. Over the 10 months of project implementation, local peacebuilders have collected and verified information on instances and sources of communal conflicts, efforts that have yielded a valuable database of conflict events at the local level. Collecting this data has raised expectations among the peacebuilders and community members alike, and the lack of a mechanism to respond to such incidents has generated frustration and disillusionment. In some instances, peacebuilders have taken it upon themselves to intervene in local conflicts, which may not only place them in harm's way but could also jeopardize their data collection efforts in the long term.

Finding 18. The participation of state-based actors as trainers or within networks of information sharing has produced a perception of risk among project participants.

The evaluation team found evidence that in some instances, partnering with governmental actors to conduct training on sensitive topics can generate mistrust that undermines training efforts. Journalists

who participated in a training co-facilitated by the government in Batticaloa, for example, perceived the training as a state intervention aimed at ensuring compliance with the Online Safety Bill and other laws. Instead of mere awareness raising on the legal framework, they interpreted it as a warning about the potential repercussions journalists could face if they failed to comply with a law that is seen by many as punitive. For some, this was perceived as a threat from a state institution. Additionally, unlike training programs facilitated by civil society organizations, given the prominent role of government actors, journalists did not feel free to fully express themselves, as they questioned the confidentiality of the sessions. While the evaluation team recognizes the importance for the UN to conduct programs that are endorsed by and even delivered with government institutions, more attention should be paid to the perception of participants, especially when engaging minority communities that may have a complicated history with formal state institutions.

These issues were further underscored in KIIs in Jaffna with provincial journalists, who themselves have experienced hate speech as journalists from ethnic minority communities, going as far as to have been labeled terrorists during the government's war with the LTTE. Given these deeply troubling personal experiences, some provincial journalists took issue with a training designed to curb their use of hate speech and inflammatory reporting and instead took the messages themselves as examples of hate speech. While the evaluation team does not perceive the general effort of training provincial journalists as a high-risk action, such training and its delivery – including the organizational affiliations of the trainers – can have a negative impact on the training's effectiveness and the perception of harm experienced by participants.

Citing another similar area of risk, a peacebuilder associated with the grassroots data collection network established by UNDP's implementing partner, CPA, raised concerns over confidentiality within a WhatsApp group established for reporting incidents of inter-communal violence and hate speech. The presence of individuals connected to government authorities or affiliated with the police within the WhatsApp group, specifically, was a concern and prompted members of the group not affiliated with the government to refrain from sharing incidents that could raise risks to themselves or to those involved in the incident. During a KII, CPA confirmed to the team that they do not restrict the participation of individuals such as officers from the groups, out of concern that excluding them could potentially worsen risks.

III.D. MANAGEMENT AND OVERSIGHT

III.D.1. Coherence

Q8. How well did UNDP and UNICEF coordinate, jointly-plan and jointly implement the project?
Q9. Did the project complement, duplicate, or undermine the work of others?

Finding 19. Lack of operational and substantive interaction between Outcome 1 and Outcome 2 reduced programmatic coherence. Our analysis of the project's design and implementation indicates that the two project outcomes can be considered independent of each other, and as such, fall short of the true spirit of joint programming. Implementation delays on the part of UNICEF provide one piece of evidence to support this conclusion. As of the first 11 months of this 18-month project, UNICEF had

managed to execute only 12.5% of its total budget;[56] by June 2023, the original project end date, that figure rose to only 36%.[57] Despite these challenges, the work in Outcome 1, which was led by UNDP, was able to be executed largely on time without being negatively impacted by the lack of implementation in most of Outcome 2.

Project implementing partners, moreover, expressed a lack of knowledge or awareness about project components implemented by others. As noted in findings on Efficiency, one of UNICEF's implementing partners, T-Field, which was engaged to help train religious leaders and youth participating in Children's and Youth Clubs, reported that at the outset of their contract, they were unaware of data and analysis on hate speech produced in Outcome 1 and instead needed to spend the first several weeks of a brief six-month contract conducting their own research to familiarize themselves with the topic.[58] Indeed, over 90% of project partners interviewed for this evaluation were unaware of what other implementing partners were doing or that the project included outputs and outcomes outside of their narrow focus, a pattern that was present among both UNDP and UNICEF partners.[59] In fact, only project partners engaged in the Community of Practice (CoP) under Outcome 1 were aware of data and analysis produced by others within the project. CEJ, CPA and NCEASL, for example, which are members of the CoP, referenced the work of Hashtag Generation and others producing data to inform their approaches. Even these more integrated partners within the CoP, however, were unaware of project components under Outcome 2.

The lack of integration and joint implementation led to missed opportunities for the project's various stakeholders to amplify the messages of others or benefit from some of the creative and innovative outputs produced by others engaged in the project. One implementing partner, for example, lamented not having been put in touch with other social media content producers, given that they could have coordinated their campaigns and helped to post and repost the content of others, facilitating broader audience reach. Another partner, Time and Space, indicated that they would like to see video clips they produced on combating prejudice and stereotypes used within education systems, seemingly unaware of the project's workstream with secondary schools.

Finding 20. The work related to data generation and analysis and the CoP complemented efforts by others.

As noted above in findings on Effectiveness, as a central part of the project's work in Outcome 1, UNDP committed to establishing an inclusive and participatory Community of Practice (CoP) to mobilize anti-hate speech action and advocate for policy change. At the time the project was approved by the PBF, however, another civil society-based CoP was already in place, having been established earlier through the support of the German Deutsche Gesellschaft für Internationale Zusammenarbeit (GiZ), raising the potential for overlap or competition between the two efforts. UNDP became aware of this CoP in April 2022 at a partnership meeting to launch this PBF-funded project. Competition between the two never materialized, however, as the GiZ effort was scheduled to wind down and outreach by UNDP ensured continuity between the two groups. Over time, through labor-intensive nurturing, the CoP has become a dynamic mechanism for sharing data, analysis and coordinating advocacy and action. An indication of

[56] PBF Periodic Report on IRF-427, November 2022.

[57] PBF Periodic Report on IRF-427, June 2023.

[58] Interview with T-Field staff, 14 March 2024, Nuwara Eliya.

[59] Notable exceptions are coordination between Sarvodaya and T-Field on their work with Children's and Youth Clubs, as well as sharing of information by Factum with 4K to inform their work with content producers.

how it has been extended beyond the earlier GiZ configuration is the growth in membership, from 14 to 22 active members at the time of this evaluation. Part of the CoP's success is due to the greater degree of institutionalization of the group, with regular meetings, summaries of meetings, reference documents outlining various initiatives led by group members, and a document repository, all of which is facilitated by UNDP. While the sustainability of such efforts is a question, UNDP's efforts to ensure complementarity and value-added to the CoP has driven greater coherence with external actors.

Apart from the CoP, development partners note that access to data produced by the project through social media monitoring and incident monitoring and available through the Early Warning System, as well as associated analytic reports, have helped to sensitize them to emergent problems and have supported their communications with their Home Offices regarding policy, diplomacy and prioritization of funding. Nonetheless, restrictions to accessing certain datasets have limited their usefulness as an advocacy tool, an issue probed more deeply in findings on Effectiveness.

III.D.2. Project Management and Monitoring

Q21. Were the project's monitoring and evaluation strategies adequate to inform decision-making?
Q22. How appropriate was the project's results framework to measure results?

Finding 21. The project's results framework falls short of articulating the higher order changes envisioned by the project's overall description and amplifies the separate nature of the workstreams in Outcomes 1 and 2 that negatively affected project cohesion.

We have included the project's results framework, with the project team's latest reporting on results, as Annex D. Based on the description of the initiative in the project document, the evaluation team's field data collection, and analysis of the results framework, we conclude that the results framework suffers from three types of weakness that have impeded the ability to adequately monitor the effects of the project and reduced overall evaluability.

First, as noted in the section on Cohesion, the project's two outcomes function independently from one another. This could have been overcome if a strategic outcome or impact statement had been set at a level above the outcome statements that identified the collective effect of these otherwise separate workstreams. Such an overall goal would make clear the project's aim of reducing hate speech to improve social cohesion and enable the project team – and the evaluators – to track the project's contribution to that change. Without a unifying outcome, each of the project's two outcome statements remain as independent interventions and set at too low a level to measure meaningful contribution to change in the social, institutional or societal spheres.

To illustrate our point, the underlying aim of Outcome 1 is to foster social cohesion by reducing hate speech through greater social and traditional media professionalism and self-regulation. The cascading theory of change assumes that if decision makers (media owners, government officials, social media companies) and content producers were presented with reliable data on the prevalence and impact of hate speech on social cohesion, they would take measures to reduce it. The problem, in essence, is a lack of knowledge about the social ills that hate speech combined with a lack of pressure from civil society and development partners to change it. It follows that if less divisive or harmful narratives or reporting were circulating in the public domain, it would be harder for those fomenting divisions to mobilize supporters or marginalize those targeted by hate speech. The inputs for this vein of work are to generate the needed data by monitoring incidence of hate speech on social media and to use this data to populate an early

warning mechanism. Armed with this early warning mechanism and evidence-based analysis, government actors, development partners and civil society partners should be able to mobilize to combat mis- or disinformation and hate speech and support institutional measures to professionalize traditional and social media industries, including through the establishment of new Codes of Practice and capacity building for traditional and social media content producers.

To test the efficacy of this approach, the project should have measured whether better information about the incidence of hate speech has prompted changes in governmental or tech and media industry policies or practice and whether these changes have reduced hate speech.[60] If so, the project should have further monitored whether reduced hate speech has supported a greater sense of social cohesion. It is crucial to take this extra step to measure social cohesion in order to test the assumption that hate speech that appears in traditional and social media is a cause or amplifier of weakened social cohesion. Only by delineating the relationship between these two higher level changes can future efforts have confidence in the overall approach of addressing hate speech to bolster social cohesion.

Instead, Outcome 1 is articulated as "Systematic research and evidence generation supporting the countering of online hate speech for effective advocacy and preventive action." As such, this language is more appropriate to an output, as it would potentially enable the project team to track the extent to which research and evidence was enabling advocacy for the prevention of hate speech. This is the second type of weakness identified by the evaluation team. The three indicators deployed to measure Outcome 1 further exacerbate this problem by tracking the number of things: the number of social media companies "committed to, and undertaking actions to strengthen hate speech moderation," the number of new indicators on social cohesion that are tracked, and the number of national stakeholders using evidence to address hate speech. None of these measures helps the project team understand whether undertaking these actions has either helped reduce hate speech or improved social cohesion, a shortcoming that constitutes the third weakness in the results framework, which is also present in Outcome 2.

Based on its analysis, the evaluation team understands that Outcome 2 aims to foster social cohesion by reducing the receptivity of hate speech among key constituencies, including young people, duty bearers who engage with young people (such as teachers and CRPOs and YSOs), religious leaders, and provincial journalists, and building their skills in combating it. A plausible theory of change would assume that if these key actors have knowledge about how harmful hate speech is and are given tools and skills to counter it, they will not (re)produce hate speech and perhaps even effectively deploy positive narratives to counter it. To test the efficacy of this approach, the project should measure the change in key actors' engagement online and offline and, if they are resisting and/or combatting hate speech, whether their more positive engagement has translated into a sense of greater social solidarity or at least reduced social division. The project's outcome statement reads, "National stakeholders actively engage in and lead early action to prevent and address hate speech and to support social cohesion." Although it could have been better formulated to focus on the change of reduced hate speech for improved social cohesion – and to remove the vague reference to "national actors actively engaged in and lead early action…" – this outcome statement offers the possibility of identifying indicators to measure the effectiveness of stakeholders' anti-hate speech efforts and whether this has contributed to social cohesion. Instead, outcome indicator 2a reads: "% and diversity of national key stakeholders targeted by the project reporting actively engaging

[60] The evaluation team recognizes that Hashtag has produced one report on social media companies' implementation of the policy, but as noted above, as of the drafting of this report, we are unable to track change over time.

in preventing and addressing hate speech and supporting social cohesion," which is not so much an indicator as it is a weak rephrasing of the outcome statement. Monitoring teams will still need to determine what is meant by "actively engaging in preventing and addressing hate speech and supporting social cohesion. Indicators like this, moreover, that create a chain of necessary actions for the respondent to meet the threshold are simply not helpful. Beyond this weakness, indicator 2b, which reads "% of sample stakeholders (disaggregated by gender, age, location etc.) demonstrate behavioral shifts in likelihood to share hate speech, appears to be measuring the same behavior as indicator 2a. The difficulty of trying to measure these indicators, moreover, is evident from the project's semi-annual and annual reports, where indicator progress for indicator 2a, for example, refers to the number of consultations with faith leaders rather than an actual measure of the stated indicator. Finally, the third outcome indicator, 2c, which reads "# and types of platforms/forums engaged in inter-ethnic dialogue and exchanges on hate speech dynamics and trends and response of participants engaged," measures output level change and is not appropriate for an outcome, as it does not shed light on whether dialogues have any effect on the attitude or behavior of participants, let alone at a broader societal level.

Taken in the aggregate, none of the project's outcome statements or indicators adequately capture the kind of change to which the project had hoped to contribute. This is a tremendous loss, as several aspects of the project display hopeful indications of positive effects or creative approaches to addressing hate speech, as is described more fully in this report's section on Effectiveness. The project did, however, produce a number of research reports and datasets that track some of these higher-level changes, such as the incidence of hate speech, offering the project team at least insight into increasing or decreasing levels of hate speech appearing in social media. The absence of a clear indicator within the project's results framework, however, renders the results framework less effective as a project management tool than it could have been.

Additionally, the results framework does not allow for consideration of changes within higher levels of change like incidence of hate speech and whether reductions can plausibly be attributed to the project's actions. The evaluation team considers this an important shortcoming, given that many of the project's actions aimed at influencing public engagement or individuals' on- and offline behavior did not occur until the last six to nine months of the project lifecycle, with some training workshops occurring as late as February 2024. [61] This timing, coming so late in the implementation timeframe, compels analysts to proceed with caution when trying to attribute even positive behavior change among stakeholders to the project, let alone any changes in higher level change.

Of course, the extent to which a two-year project can contribute to meaningfully change in social cohesion or even behavior change needs to be reflected in selecting the right indicators and setting modest targets. Despite the challenge of short project timelines, however, projects that seek to positively influence social cohesion or contribute to other higher-level changes cannot shy away from examining these higher order changes because doing so obscures the project's effectiveness at contributing to the intended change. The project must be oriented to a higher order change to keep its sights on the ultimate goal of fostering social cohesion and to understand whether the approach it has selected is appropriate.

IV. RECOMMENDATIONS

[61] These training workshops, moreover, conducted with youth engaged in Children's and Youth Clubs as well as in target schools, reportedly lasted two hours in many cases, further raising questions about the degree to which the project can said to have contributed to positive change. These points are explored more deeply in the section on Effectiveness.

Recommendation 1. In the medium to long term, UN Country Team members designing future projects on combating hate speech should consider a more prominent focus on gender-based hate speech, including cyber GBV, or more fully integrate gender responsiveness into other approaches. While the project's focus on all forms of hate speech enabled it to remain relevant as certain topics or foci of attack waxed and waned over the project lifecycle, the consistent and overwhelming incidence of hate speech against women and people identifying as LGBTQI+ auger for a more focused and dedicated gender-responsive approach. The project produced several nascent positive results upon which to build, such as the network of young content producers reached by CEJ, but these will likely – and unfortunately – dissipate if left unsupported. These results, however, delivered with very modest budgets and through very short-term contracts, illustrate that this is a productive area of work that is in demand from national stakeholders and highly relevant to the context. Project design teams should investigate how to harness the power of partnerships not just with women and the LGBTQI+ community, but of allied men.

Recommendation 2. In the medium to long term, UN Country Team members working to advance the UN SDCF's pillar on social cohesion should consider designing initiatives that are dedicated to working with and through religious leaders. Religious leaders emerged through our evaluation process as a key stakeholder group for influencing communities – among the most trusted actors according to opinion polling. As important as they are to reaching communities, they are not a simple, monolithic group. They have different organizational structures and even different pathways into a life of service. Buddhist monks and Catholic priests, for example, do not have families and do not receive compensation for their work; Muslim imams have families but do not receive compensation, and Hindu leaders both have families and receive compensation – these are all differences that can cause misunderstanding and tension even between the religious leaders at the local level. They also have different levels of hierarchy within their organizations and different cultural-religious tenets when it comes to women's inclusion and leadership, among other issues. Language differences, moreover, are as much a barrier for them as they are for the general population. As a result, trying to work across religions is difficult and, as revealed by the time that UNICEF needed to dedicate to gain enough traction with each religion's upper hierarchies, it needs its own dedicated workstream. When engaging them on capacity building themes, religious leaders are not accustomed to accepting training, which means that capacity building approaches that may work for others will have a harder time connecting with religious leaders. Getting all groups to include female religious leaders, moreover, further illustrates the challenges of working through religious leaders. Given their importance as influencers at the local level, however, and the regular referencing of religious differences as the basis for social tension and violence, the team recommends further work with religious leaders to strengthen nascent intra-religious bonds established through the project and advance social cohesion through a dedicated, intensive focus on this group, not as one among five or more additional stakeholder groups.

Recommendation 3. In the short term, the CRD and EWS are innovative and powerful tools for capturing much needed information on drivers of conflict that could be better utilized if UNDP and the RCO enabled a broader set of activists and policy actors to have access, as envisioned by the project document. Broadening access to a wider array of stakeholders, moreover, may also provide a greater diversity of voices and opinions that could help these platforms remain sensitive to emergent conflict drivers in order to be not just relevant but essential. Sensitivities about the data are justified, given the potential for misuse in the wrong hands (or social media accounts). The evaluation team recommends that if it is not possible to make a portion of the existing platforms more widely available, then the RCO and UNDP should explore creating an external-facing platform that utilizes select data within these excellent resources and makes them available to more actors. Given that the CRD and EWS have been in

place for several years, moreover, and considering the cost of maintaining up-to-date data, UNDP should invest in systems to better track how this information is used so that it can make informed decisions about its continued usefulness. The evaluation team is confident that with more transparency and inclusion, combined with the ability to rigorously demonstrate the mechanisms' importance, financial support to sustain the systems will be easier to attract.

Recommendation 4. In the short term, the CoP, with support from UNDP, should develop a one- to two-year plan to stand up the group to be able to operate on its own or with minimal assistance. Current levels of support by UNDP are not sustainable. More importantly, however, if the CoP is seen as a valued platform by national actors, they should take the reins and determine for themselves what the CoP's role should be moving forward. Once this collective vision is articulated, UNDP can work with CoP members to plan for a gradual reduction of support, offering capacity building if needed, but establishing clear benchmarks for tracking progress toward the CoP's independence.

Recommendation 5. In the short and medium terms, UNDP and UNICEF should capitalize on and repurpose the compelling commercials produced by Time and Space, as well as social media content, games and other creative outputs that can help community members, journalists, students or other stakeholders get excited about learning. As it stands, much of the excellent creative content produced for this project risks being shelved as the project winds down, despite commitment on the part of various actors to continue training and education efforts with a variety of stakeholders. UNICEF, for example, would do well to work with UNDP to identify specific outputs such as Moonshot's game, Time and Space's commercials or social media content produced by NCEASL's university students that could be tapped to invigorate training with students and young people in Children and Youth Clubs. These stakeholders are already demanding such creative learning tools. To not provide them when they have already been developed would be a missed opportunity to demonstrate both value-for-money as well as responsiveness to stakeholders' explicit learning preferences.

Recommendation 6. **In the short term, as UNICEF anticipates the rolling out of the NIE and Ministry of Education's curriculum reforms in the coming months and years it should encourage that provincial-level education initiatives designed to address hate speech are brought into the planning and rollout process, and that learning from these are incorporated into the national plan.** The planned pilot in Central Province offers both an opportunity and a risk. It is illustrative of the dedication and ownership felt by provincial level stakeholders, which is unmitigatedly positive. Their intention to proceed on a parallel track from the national reforms as well as module development on hate speech connected to the reforms, however, may place Central Province at odds with the national effort down the road. While the national reforms may be compulsory, and Central Province may be forced to adopt them, if not managed well, Central Province stakeholders may feel alienated and lose their sense of deep local ownership and drive, which has been such a force for good until now. To avoid this, and to capitalize on the excellent opportunities for informing the national rollout, UNICEF should plan now for key moments of reflection where national-level actors can jointly learn with and from Central Province actors what is working and what should be revised to better reach students with effective pedagogical approaches and content. These moments should coincide with programmatic inflection points in the Central Province pilot with ninth graders as well as monitoring and revision opportunities during the early phases of the implementation of the new national curriculum.

Recommendation 7. Across short to long terms, training efforts across the board need a more sustained approach both for master trainers as well as for participants of training, and effective monitoring of their outcomes. Efforts to combat hate speech and promote social cohesion are fighting against deeply

entrenched, long standing social separation, bias and othering. One or two 3-hour sessions will not change attitudes, behaviors or practices in a sustained manner. Conversely, a full weeklong session tends to overwhelm and sometimes bore participants – and some, like religious leaders, cannot be away from their duties for that long. UNDP and UNICEF should plan a series of training and other learning opportunities over the course of months or years to intentionally build toward a clear objective. The identity of trainers, moreover, especially those affiliated with the state, must be taken into account if training is to be effective. While this recommendation applies to all aspects of training witnessed by the evaluation, that provided to teachers, CRPOs and YSOs deserves special attention. **If UNICEF plans to continue working on hate speech and social cohesion through teachers, CRPOs and YSOs, it must provide access to more frequent training and structured opportunities to share experiences among not only the Master Trainers, but among school and Club resource people.** Evidence from the field clearly demonstrates that where training was more widely available over a longer time, trainers were more effective at delivering their messages. This should not be surprising. Planning in advance not just for initial training but periodic remedial sessions will help ensure consistency and identify challenges faced by individual trainers.

Recommendation 8. In the short to medium term, UNDP and the Ministry of Mass Media should continue to offer training and other professionalization opportunities to journalists, including provincial journalists to build upon learning begun through this project. While taking forward future training, **greater attention to the specific needs and barriers of female journalists should be factored into the training content as well as the organization of workshops.** Complementing efforts to promote professionalism of journalists, UNDP and the Ministry should consider secondary and tertiary programmatic strands to tackle continuing industry incentives that compel industry leaders and the public to demand inflammatory and hateful content. This three-pronged approach would help elevate media standards by advancing a more holistic approach with a broader number of relevant actors.

Recommendation 9. In the medium to long term, to more effectively link training programs like combating hate speech to support for social cohesion, UNICEF should look for opportunities to promote cross-school or cross-club sharing and launch joint events to reduce the social separation currently experienced by many young people with whom we spoke. UNICEF could consider twinning schools or Clubs from different ethno-religious communities or parts of the country. On the basis of these cross-communal exchanges, UNICEF could consider study visits or sporting events that bring diverse young people together offline to break down social barriers. This will need to be planned well – young people with whom the evaluation team spoke expressed curiosity and openness toward others but also exhibited underlying biases and stereotypes that reflect the formal and informal institutions in which they live and learn. And to achieve deeper effects, such efforts must be embedded in broader initiatives that support young people through a medium- to longer-term process of personal growth and relationship development. One cross-communal event or small-scale social impact project will not be enough.

V. LESSONS LEARNED

Lesson Learned: a more limited number of stakeholders and implementing partners would have enabled longer-term engagement and, likely, deeper effects. While each of the project stakeholders was relevant to the issue of combating hate speech and supporting social cohesion, attempting to work with all of them within an 18-month project spread financial and human resources too thinly. Given that the project never brought its work with secondary students, Club members, journalists or religious leaders together in any meaningful way, any one or even two of these stakeholder groups could have been omitted from the project without affecting outcomes for the remaining stakeholders. Not having done that yielded an unwieldy number of implementing partners operating under tight timeframes with very

modest resources, with the result that cross-learning and coordination of efforts was lost. In the future, UNDP and UNICEF should consider focusing on a smaller number of stakeholders, implementing partners and workstreams, and maximizing opportunities for connections where work from one stream could complement and amplify work in another.

Lesson learned. The lack of a unifying outcome, and outcome indicators, that clearly articulated how work with various stakeholders on hate speech would contribute to social cohesion led the project team to focus too narrowly on individual project outputs. This is a recurring shortcoming, having already been identified in an evaluation from 2021 of an earlier project with similar components. If the project had been designed with one layer of change in between the current outcomes and the overall goal of improving social cohesion, it would have been better able to link its activities to the desired change. The project, then, would have stood a better chance at having been monitored on the basis of testing the underlying assumptions of the theories of change noted in Section I of this report.

Lesson Learned. Advancing Gender Equality and Women's Empowerment needs an explicit, dedicated approach, otherwise, it is difficult to keep track of whether the project is maintaining the commitments it made in the proposal, and opportunities to leverage work in one part of the project can be lost on other aspects of programming that could have benefited from the experience or resources of others. There were scattered positive effects within the project, but identifying them was not easy, and they seem to be disconnected from each other – Code of Practice including gender-responsive indicators, but journalist work fairly gender blind; good solidarity network and content from social media content producers within one initiative, but opportunity lost to broaden the network across all social media content work; barriers to reaching female religious leaders meant that 20 to 30% of the focus of that workstream meant to advance GEWE went unrealized. The sum total of underperformance across various outputs yielded a final project budget implementation rate below the 30% minimum threshold for GM2.

Lesson Learned. Differences in language use between Sinhala and Tamil speakers is a real barrier to greater interaction, effective training, and ultimately, social cohesion. Training needs to be provided fully in the language with which participants are most comfortable. Not only will that improve their understanding of the subjects, but it will signal to the participants that the training was designed for them, and that they weren't included as an afterthought. Young people in schools and Clubs expressed that even when they have encountered peers from another language group in UN-sponsored events, they are only able to have the most basic interactions and the friendship never goes anywhere. Social media monitoring by SecDev, as well, noted distinct differences in which platforms are utilized by which group, signaling that as much as social media may connect people, through the way that algorithms tailor content to individual users, it also enables deeper divisions where those from one community or perspective never need to encounter the other. Language barriers further deepen this divide. No easy fix for this, but it is a strong lesson learned that will need to be built into the design of any future social cohesion efforts.

Lesson Learned. Working with journalists will only get you so far, if editors and owners have every incentive to publish inflammatory articles. While the project touched upon the importance of working with newsroom editors and administrators, the voluntary nature of adhering to fact-based and ethical reporting place media outlets in a bind, needing to choose between profitability and ethics. All respondents in KIIs for this evaluation, however, noted that Codes of Practice, consensus documents and media policies are a necessary first step in provoking a positive change, even if their voluntary nature means that adoption will be evolutionary rather than revolutionary. UNDP and the Ministry of Mass Media should be encouraged to deepen work in this area and to simultaneously seek to reduce public demand for inflammatory reporting.

VI. CONCLUSION

The evaluation of the project Combating Hate Speech and Promoting Social Cohesion in Sri Lanka has initiated significant changes that lay a path for future directions. Overall, the project has made commendable strides in generating needed data on factors that contribute to hate speech and social division. It has effectively mobilized various stakeholders, including media professionals, religious leaders, and youth, to foster a more cohesive society. The project's systemic approach, particularly through partnerships with government ministries and the establishment of a Community of Practice (CoP), has shown promising early signs of generating sustained, impactful change. However, the evaluation also highlighted areas for improvement, especially in coordinating efforts across different project components, effectively monitoring change at a meaningful level, and ensuring long-term sustainability.

The project's emphasis on embedding activities within existing systems, such as the Ministry of Education's curriculum reforms and the Ministry of Mass Media's professionalization initiatives, offers the best opportunity for nationwide, enduring change. These approaches ensures that the project's impacts will extend beyond individual learning to effect broader institutional and societal transformations. Nonetheless, the evaluation noted that the absence of a unifying outcome and clearly articulated indicators hindered the project's ability to measure its overall impact on social cohesion effectively. Future initiatives should aim to integrate these elements more cohesively to enhance their evaluability and overall effectiveness.

One of the specific successes of the project was UNDP's work in generating needed data on hate speech, which played a crucial role in shaping national strategies and policies. The development and promotion of a voluntary Code of Practice for social media companies is a notable achievement, representing the first such code in the Asia-Pacific region. This initiative, coupled with professionalization programs for traditional media conducted in partnership with the Ministry of Mass Media, aims to elevate media standards and reduce the prevalence of incendiary reporting. Despite these advancements, the project must continue to address market incentives that drive media outlets to publish inflammatory content, ensuring that ethical journalism takes root and flourishes.

UNICEF's collaboration with the Ministry of Education on curriculum reform has been another highlight of the project. By integrating concepts of digital citizenship, critical thinking, and inter-group tolerance into mandatory coursework, the project has laid a strong foundation for fostering a more inclusive and respectful society. The ongoing reforms promise to impact all students in the formal education system, significantly extending the reach of the project's aims. However, the successful implementation of these curricula will require sustained efforts to train educators and other influential duty bearers properly and monitor the impact of the new content on students' attitudes and behaviors.

The project also demonstrated innovative approaches in working with young social media content producers and traditional media journalists. Training programs equipped these stakeholders with the skills to recognize and counter hate speech, resulting in a range of impactful media content and campaigns. For example, young content producers developed social media campaigns that reached broad audiences, promoting messages of tolerance and countering cyber gender-based violence (GBV). These efforts have shown positive early results, although continued support and training are necessary to sustain and build on these achievements. Better coordination and connectivity between social media campaigns, moreover,

could have further amplified the new contents' reach and solidified supportive working relationships among up-and-coming content producers.

Looking ahead, the evaluation recommends a more focused approach to tackling gender-based hate speech, including cyber GBV. The consistent and overwhelming incidence of hate speech against women and the LGBTQI+ community suggests a need for dedicated gender-responsive strategies. Additionally, future projects should consider a more intensive focus on engaging religious leaders, who emerged as key influencers capable of fostering social cohesion within their communities. Strengthening intra-religious bonds and promoting inter-religious understanding through targeted initiatives can further advance the project's goals.

In conclusion, while the project has made substantial progress in addressing hate speech and promoting social cohesion, it must continue to refine its strategies and enhance coordination among various components. By embedding activities within existing systems, focusing on gender responsiveness, and leveraging the influence of key stakeholders like religious leaders, future initiatives can build on the foundations laid by this project to achieve more profound and lasting impacts.