

FINAL

REPORT

RIGHT TO PLAY

MA'AN MID-TERM EVALUATION

MARCH 2023



proteknôn



مركز دراسات التنمية
Center for Development Studies



BIRZEIT UNIVERSITY

Table of Contents

EXECUTIVE SUMMARY	3
OVERVIEW.....	9
METHODOLOGY AND DESIGN.....	10
SAMPLING AND PARTICIPANT SELECTION	10
SURVEY DESIGN AND TOOL DEVELOPMENT	19
DATA COLLECTION.....	21
ANALYSIS.....	21
LIMITATIONS	26
FINDINGS.....	29
EXTENT PROJECT RESPONDS TO THE CURRENT AND FUTURE NEEDS AND PRIORITIES OF THE AFFECTED POPULATION	29
EXTENT PROJECT IS ON TRACK TO ACHIEVE ITS INTENDED OUTCOMES	34
EXTENT PROJECT CONTRIBUTED TO THE MAIN OUTCOMES	46
EXTENT PROJECT IS GENDER RESPONSIVE	52
EXTENT THAT ANY OUTCOME CHANGES ARE LIKELY TO CONTINUE BEYOND THE END OF THE PROJECT	56
RECOMMENDATIONS.....	60
RECOMMENDATIONS FOR FINAL YEAR OF PROJECT	61
RECOMMENDATIONS FOR FUTURE ITERATIONS OF THE PROJECT.....	65

Executive Summary

Overview

MA'AN is a five-year project (2019–2024) of Right to Play, financed by the Government of Canada through Global Affairs Canada, that works towards improved, safe, supportive, and gender-responsive education for girls and boys in grades one through nine enrolled in every United Nations Relief and Works Agency for Palestinian Refugees in the Near East (UNRWA) schools in the West Bank and Gaza Strip.

Methodology and Design

This mid-term evaluation used a **quasi-experimental mixed methods design** to assess the project's relevance, effectiveness, impact, and sustainability, and to generate actionable learning and recommendations.

The evaluation **surveyed 1,959 total participants** (59 percent female and 41 percent male) across the West Bank and Gaza to compare project attainment from baseline in 2019 to midline in 2022 and between treatment and control schools. The survey sample was drawn by randomly selecting and interviewing primary teachers, PE teachers, and counsellors and then interviewing a randomly selected subset of these staff's students and the students' parents. Control schools, which were selected during the baseline evaluation, were comprised only of primary-level schools in Gaza, and at mid-term were compared with treatment schools, which were other primary-level schools in Gaza that began their participation in year three of the project. The evaluation **triangulated survey data** with **110 total participants** by conducting **interviews** (70 percent female and 30 percent male) and **focus group discussions (FGDs)** (85 percent female and 15 percent male).¹

Data collection began on 3 December 2022. It concluded in Gaza on 29 December 2022. However, local worker strikes, escalating violence, and sampling revisions caused a five-week delay with data collection in the West Bank, which concluded on 2 March 2023. The resumption of the local worker strikes, which were ongoing as this evaluation report was being written, meant that only 16 out of the planned 23 West Bank interviews and FGDs could be completed.

¹ Semi-structured interviews (SSIs) were held with educational supervisors, primary teachers, and counsellors. Key informant interviews (KIIs) were held with principals. FGDs were held with parents. Time and resource constraints prevented necessary child protection and ethical reviews to be able to include children in the qualitative part of this evaluation.

Survey data were **analysed** for all outcome-level indicators and according to the MA'AN performance measurement framework (PMF). Interview data were thematically analysed with each of the main evaluation questions and criteria linked to a set of questions across the different qualitative methods. All data were disaggregated by gender, age-cohort, location, and respondent category. Initial findings were **validated** in a hybrid workshop held on 28 January 2023 with MA'AN and Right to Play staff. Key **limitations** to the evaluation included: revisions to the tools, which prevented some comparability of baseline and midline outcomes; control schools were limited to primary schools in only one territory (Gaza) and some spill-over effects of the project might have constituted a contamination of the treatment group; time and budget constraints prevented including students in qualitative interviews; and lastly, the project was implemented during remarkably adverse conditions, which included a pandemic, increasing violence in both territories, increasing economic insecurity, and several local workers strikes.

Findings

The evaluation found the MA'AN project to be relevant, effective, and impactful. It responded to felt needs in the community, was widely appreciated, helped to build the capacity of school staff to implement Gender-Responsive Play-Based Learning (GR-PBL), and positively impacted on students' social-emotional well-being.

Key Outcomes

In both **treatment** and **control**, and **baseline** and **midline comparisons**, MA'AN attained **statistically significant positive outcomes** at the student, teacher, and parent level. Students reported having an **overall improvement** in their **psychosocial well-being**, having a “positive” **perception of their teacher**, and exhibiting **context-appropriate social/emotional skills**. Staff were more likely to report using the **referral mechanisms** to address **gender equality and gender-based violence (GBV)**. Teachers were more likely to deliver **GR-PBL**. **Baseline and midline comparisons** also detected positive effects along several other indicators, which included **counsellors** and **PE teachers** demonstrating **play-based learning methodologies** in their psychosocial support activities for students, **classrooms** meeting **criteria** for **emotional safety/emotional well-being**, **students** with feelings of **social connectedness**, **teachers** knowing **strategies** to **include girls** in educational activities, and **caregivers** having knowledge of **psychosocial support services** and where and how to formally **report GBV**.

Relevance to Stakeholders

MA'AN was viewed as an appropriate and **relevant** intervention to address concerns identified by UNRWA and other stakeholders in the targeted schools and communities, which included **low**

school performance, psychosocial challenges, and gender equality. The project's **relevance increased** as the pandemic and school closures, the political situation, and worsening economic conditions increased stress, anxiety, and isolation. MA'AN's relevance was reflected in school staff's commitment to implementing and adapting the training, high attendance at parents' meetings, and stakeholders' unanimous calls to expand the project.

Staff Capacity

Interviewed research participants further emphasised the impact and effectiveness of the project. For teachers, it was reported that MA'AN **built their capacity and confidence** and practically helped them improve their teaching practices. Counsellors reported that MA'AN helped them **improve** their **day-to-day** work with students and **overall performance**. It was reported that MA'AN helped to create a **more collaborative staff environment**. Education supervisors reported an increase in teachers wanting to **support underachieving and marginalised students**.

School Environment

Interviewed research participants also reported that the inclusivity of the GR-PBL approach made classroom environments more **cooperative** and **inclusive** and helped to **instil values** and skills such as respect, accepting others, forgiveness, and imagination. GR-PBL was a **more holistic intervention** and helped students gain new **skills** in **leadership, cooperation, empathy, and self-esteem**. The GR-PBL activities helped to build trust among girls and boys, better include children with disabilities, and increase a sense of belonging, respect for school rules, and paying attention in class. Most parents reported that they saw a **direct impact of the parents' meetings** on the way they deal with their children.

Attitudes Towards Gender Equality

Findings suggest project effects on outcomes related to gender equality. Both baseline/midline comparison as well as treatment/control comparison suggests that teachers feel **more supported in delivering GR-PBL and know how to use referral mechanisms**, among other mechanisms, to address GBV. This was echoed by interviews where counsellors reported that they are now using these referral mechanisms and that there is an increased awareness about them among parents. Interviews suggest that the MA'AN training helped counsellors implement **gender-responsive approaches** with students and that there was a noticeable decline in **gender stereotyping**.

However, there were decreases from baseline to midline, most notably among students' attitudes towards **ending GBV** as well as caregivers' **awareness of the gender-based education needs of girls and boys**. One explanation for these decreases might be related to awareness of

the issue with only approximately 60 percent of parents and students indicating that they know what GBV means. Another explanation is that these decreases may not be due to the project intervention, but instead have largely resulted from contextual factors such as COVID-19, school closures, and increasing violence in the region. Treatment/control data suggests MA'AN schools are still performing better than school communities not served by the project; that they experienced less of a decrease suggests that the wider contextual factors are influencing attitudes towards GBV and that the project helped to cushion some of their detrimental effects.

Parents overall reported high support for gender equality, but also disclosed information that reinforced gender stereotypes. Research participants felt that individual efforts and beliefs for greater gender equality were still overshadowed by dominant societal views.

Gender Equity

While the project largely seems to have helped both girls and boys to the same extent, in some cases girls seem to have benefited to a larger extent than boys—demonstrating **higher levels of improvements in their well-being**, more **appropriate social/emotional skills**, and higher use of **school counselling mechanisms and GBV services**. Further, while differences in gender outcomes were not commonly found in the evaluation, on certain indicators where there were differences, both boys and girls at **single-sex schools** demonstrated **higher levels of attainment** than boys and girls at **co-ed schools**. For example, boys and girls at **single-sex schools** were more likely than boys and girls at co-ed schools to see improvement in their psychosocial well-being (1000.2), have positive perceptions of their teachers (1100.3), and exhibit higher social/emotional skills (1100.4/1200.5).

Sustainability

Research participants were highly interested in the future **sustainability** of the MA'AN project with ideas and suggestions for what could be improved and expanded and how to sustain it. All participants thought the **project should be expanded** by increasing the number of staff trained and involved in future phases of the programme, extending MA'AN to other schools, including non-UNRWA schools, and supporting schools with more toolkits and resources. Participants highlighted that the role of counsellors was very important, but that more were needed as their capacity was very strained. Staff suggested that **further training** would help staff better implement the activities. The **communities of practice (CoPs)** were viewed as crucial in helping teachers to implement what they have learned in the training and transfer knowledge to other teachers, even those who did not participate in the MA'AN training. **Parent meetings** were viewed as **extremely positive**, and it was recommended to increase their frequency, target more parents to attend, and possibly expand to events such as open days at the school.

Challenges

Research participants did however reveal many challenges in implementing the project. Teachers at single-sex schools had to make frequent adaptations to the activities to fit their context. There were resource constraints, which included limited time, space, and sports tools/supplies to conduct activities as planned, feelings of teachers already being over capacity and high caseloads among counsellors and education supervisors. Left unaddressed, these challenges will limit the impact of the project in the long term.

Recommendations

Across all research participants, the evaluation found consistently high levels of trust and appreciation for the project, which it can leverage in its **final year** of activities to implement strategic changes to ensure **sustainability**, adapt **tools and materials**, and improve **GBV interventions**.

Sustainability: MA'AN's primary focus in its final year should be on building systems to support the longevity of the project. Focus should shift from what MA'AN staff are doing, to how MA'AN staff are equipping other key stakeholders to carry the work forward. Conversations and planning with UNRWA should be initiated to set the foundation for long-term sustainability measures, which could involve forming multi-stakeholder leadership teams in communities and territories to oversee future activities, creating a train the trainer programme to equip UNRWA staff to lead and run trainings currently run by MA'AN staff, and leveraging the role of educational supervisors to help ensure there are regular times and a schedule for trainings and CoPs.

Tools and Materials: MA'AN should work to equip education supervisors and principals with tools they need to help assess the adoption of the project and continue to promote GR-PBL. MA'AN should also work to reinforce the definition of GR-PBL by continuing to develop very clear examples of what is meant by GR-PBL, to ensure that certain practices are not conflated as part of the methodology when they are not. One way to do this would be to develop signage that can be used in classrooms and school common spaces to communicate key GR-PBL messages. Based on feedback from school staff, MA'AN should make final and key revisions to the training materials.

Gender-based Violence: MA'AN should work to increase promotion of the GBV referral mechanisms and ensure that all schools are holding regular parents' meetings and increasing the frequency as capacity allows. Further, MA'AN should work to increase activities in the training as well as support specifically for teachers in co-ed schools.

Project design: In any **future iterations** of the project, MA'AN should work to increase the alignment of the targeted outcomes on gender equality with the project interventions, decide on the most strategic use of staff roles, and support UNRWA to identify key resource needs for them to carry the project forward.

Student outcomes: MA'AN should consider making a longer-term investment to understand the project's impact on student outcomes, by possibly increasing the intensity of the project in selected schools to evaluate any difference and/or conduct a longitudinal study. In the final evaluation, MA'AN should budget time and resources to conduct qualitative participatory research to capture the project's impact more holistically in the students' own words.

Gender equality: It will be important moving forward for MA'AN to clarify the project's gender equality goals and ensure that the intervention is designed to be able to accomplish these. To advance gender equality, MA'AN should develop more targeted strategies for single-sex schools and a class- and school-level gender monitoring framework. MA'AN could work to strengthen and expand the project's gender components by more actively leveraging parents' and students' ideas and leadership on the issue. MA'AN should also plan for increased integration with community-based initiatives to change attitudes towards gender equality.

Mid-Term Evaluation Final Report

Overview

MA'AN: Towards Gender-Responsive and Empowering Education for Girls and Boys in the West Bank and Gaza is a five-year project (2019–2024) of Right to Play, financed by the Government of Canada through Global Affairs Canada. The project is aimed at addressing barriers to quality education for Palestinian refugee children in grades 1–9 enrolled in every United Nations Relief and Works Agency for Palestinian Refugees in the Near East (UNRWA) schools in the West Bank (95 schools) and Gaza (275 schools).²

The ultimate outcome of the project is the improved quality of education for children, especially girls, enrolled in UNRWA primary schools in the West Bank and Gaza. This is being measured and reported through a set of mixed-method indicators (qualitative and quantitative) on feelings of social connectedness, improvements in psychosocial well-being, and emotional safety/emotional well-being in classrooms.

The three key intermediate outcomes of the MA'AN project focus on:

1. Increased integration of GR-PBL by primary teachers into teaching practice.
2. Improved resilience of girls and boys to overcome conflict and violence, including GBV.
3. Improved response of school officials, UNRWA staff, mothers, fathers, and community members to the gender-based education needs of girls and boys and GBV.

This evaluation, conducted by the Proteknôn Foundation and Birzeit University consortium, has evaluated the **relevance, effectiveness, impact, and sustainability** of the project to understand its performance and to generate actionable learning and recommendations to improve it. As this evaluation has been conducted in year four of a five-year project, it has assessed progress on outcomes and identified components of the project design or implementation that could be strengthened in its final year. It has also generated actionable learning relevant to the MA'AN project and to improving similar future Right to Play projects based on what has been learned from delivering MA'AN. This has been done by responding to the following evaluation questions set out in the Terms of Reference.³ These evaluation questions have been used to structure this evaluation final report.

² Right to Play, MA'AN PIP Final, 26 June 2019.

³ As presented in the Terms of Reference (June 2022).

- To what extent does the project respond to the current and future needs and priorities of the affected population?
- To what extent is the project on track to achieve its intended outcomes?
- To what extent has the project contributed to the main outcomes?
- To what extent is the project gender responsive?
- To what extent are any outcome changes likely to continue beyond the end of the project?
- How could the project be adapted or strengthened to improve performance against the intended outcomes?

The evaluation used a quasi-experimental mixed methods (quantitative and qualitative) design with treatment and control schools to assess MA'AN's performance at midline compared with baseline results, outcome targets, and the comparison/control group to generate findings and learnings to inform this and future projects.

Report Structure

This report begins by detailing the **methodology and design** of the evaluation, describing the sampling framework, tool development, data collection, and data analyses for both the quantitative and qualitative components of the study. Key **limitations** of the evaluation are then discussed, along with implications of how to interpret the findings considering these limitations. Next, the **findings** are reported to answer each of the evaluation questions above. Each section of the report details the evaluation question, criteria, and data sources. The report concludes with providing **recommendations** for the final year of the project and its future iterations.

Methodology and Design

Sampling and Participant Selection

The evaluation reached **1,959 participants** through school-level surveys (59 percent female and 41 percent male), and an additional **110** participants through interviews (70 percent female and 30 percent male) and FGDs (85 percent female and 15 percent male) for a **total sample of 2,069 individuals**.⁴

⁴ The participants of the FGDs were selected in collaboration with the counsellors who reported that women tend to participate more in the parents' meetings organised by counsellors.

QUANTITATIVE SAMPLING FRAMEWORK

SAMPLING OVERVIEW

The quantitative sampling framework was designed to achieve the following two goals:

- **Sampling Goal One:** assess the difference between treatment and control schools (referred to as “treatment/control sample”)
- **Sampling Goal Two:** assess progress from baseline to midline against targeted outcome indicators (referred to as “logframe sample”)

The approach to achieving each goal, as well as considerations that led to the approach, are described below.

The treatment/control sample

At baseline, a group of schools in Gaza were determined to constitute the control group. This group eventually became smaller, since most of these schools were gradually included in the project. By the time of the evaluation, only five primary-level schools remained outside the project and therefore teachers at these schools had to be used as the control group for the evaluation. While the control sample was already drawn at baseline, the treatment sample was not determined at baseline and had to be drawn for this evaluation. Drawing the treatment sample was influenced by a few factors: First, treatment schools within Gaza varied in terms of their school type (e.g., co-ed schools versus single-sex schools) and size. Second, within treatment schools it varied as to when teachers first participated in the project (i.e., year one, year two, year three). To respond to this, it was decided to pull the sample from primary teachers who began their participation in year three of the project and were therefore the most recently engaged cohort in the training activities. The **final treatment/control sample** was only drawn from **primary-level schools in Gaza** and was comprised of **primary teachers, primary students, and parents of primary students**.

The logframe sample for the baseline/midline comparison

Unlike the treatment/control sample, the **logframe sample** was comprised of **all participant groups** from **both Gaza and West Bank**: primary teachers, primary students, and parents of primary students, as well as principals, PE teachers, counsellors, secondary students, and parents of secondary students. The logframe sample selected primary teachers who began their participation in the project in years one, two, and four (because the treatment/control sample

only selected teachers who participated in year three of the project) and selected counsellors and PE teachers who began their participation across all project years.

SAMPLE CONSIDERATIONS

While ideally, there would have been one sampling design to accomplish both goals and allow for a joint comparison of the midline sample with both the baseline and the control sample, this was not possible for the following reasons. First, given that the project is implemented in both Gaza and West Bank, the baseline sample covered both territories. By contrast, the control sample included teachers from five schools in Gaza only. Second, given that the project targeted school staff and students at both primary as well as secondary level, the baseline sample consisted of both grades one through four (primary) as well as grades five through nine (secondary). By contrast, control schools included teachers and students from the primary level only.

To ensure that sufficient data is available within both types of comparisons, the sample required for the treatment/control comparison was determined first and then the sample for the baseline/midline comparison (logframe). To keep logistical burdens to a minimum, the treatment sample was integrated into the sample used for the baseline/midline comparison. By contrast, control schools were not integrated into the baseline/midline comparison. Given that participants interviewed in the context of the baseline/midline comparison formed the basis to determine project attainment against all indicators, this sample was called the “logframe” sample.

SAMPLE SELECTION

In both the treatment/control and logframe sample, the sampling used **teachers as the unit of selection**. Primary teachers (as well as PE teachers and counsellors) were randomly selected out of the participant lists provided by Right to Play. After this, for a subset of these primary teachers, PE teachers, and counsellors, students were interviewed. Identifying students for the selected teachers required different strategies given the differences in teacher role/scope. Whereas primary teachers largely only work with one set class of students, PE teachers and counsellors work with students across multiple classes. Thus, the strategy for selecting the class for each school staff type was as follows: For selected primary teachers, six students out of their class were interviewed. For selected PE teacher and counsellor, four classes were randomly selected, and then within these four classes, six students were randomly selected.

In each class, the following steps were used to select students:

1. Consent forms: Field teams used consent forms as the starting point. Prior to data collection, consent forms were collected from students. Thus, by the time field teams entered schools, it was already known which students would participate in the surveys.
2. Class-specific groups: Consent forms were kept with the administration of each school and clustered by class. When field teams arrived, they were given the completed consent forms for the class selected for the survey.
3. Random student selection: Out of each class group of consent forms, field teams selected students using a random number generator programmed within the Kobo Toolbox, which was used for data collection. Within co-ed schools, the class-specific sample was stratified to select three female and three male students for each class.

SAMPLE SIZE

Multi-stage sample size calculation guided the overall scope of the sample. It was partly constrained by the fact that only a set number of teachers and counsellors were available to form the sample frame. A challenge with having teachers as the unit of selection for sampling is that it required drawing from a large number of schools; the initial sample would have required data collection across 193 different schools (141 schools in Gaza and 53 schools in West Bank), which posed significant challenges to feasibility. To respond to this, in-person visits to schools were only made for the primary teachers, PE teachers, and counsellors that also had students selected for the sample; all others were contacted by phone. Out of the 389 primary teachers, PE teachers, and counsellors selected for the study, 179 (46 percent) were interviewed in person. Several t-tests were performed on the indicator data and no significant differences were found between those contacted by phone and those visited in person.

Table 1, Table 2, and Table 3 below outline the details of the actual sample realised. Table 1 outlines the sample by each sampling goal—treatment/control and logframe. Table 2 outlines the sample by territory. Table 3 outlines the number of primary and secondary classes represented in the sample.

Table 1 also provides indication of the extent to which each sub-sample was completed. The corresponding percentages are presented in brackets: 100% indicates the entire sample has been completed as planned; any figure below 100% indicates underachievement; any figure above 100% indicates overachievement. There are two sample targets with significant variation from the goal—primary student targets were exceeded, while secondary school targets were not fully attained. The reasons for each are related. The evaluation design assumed that PE teachers and

counsellors were working only with secondary students. However, during data collection it was discovered they were partly working with primary students as well. This resulted from changes and rotations by UNWRA of staff, particularly for counsellors and PE teachers. Since PE teachers and counsellors were “units of selection” as well, on-the-spot randomisation procedures led at times to selecting students who were both secondary and primary students. For parents, the actual interview numbers were also below target. In general, the evaluation aimed at reaching the parents of 30 percent of selected students. There are two factors that might have contributed to this under achieving of the target: First, parents were interviewed via telephone, which made it more difficult to reach them and, even when reached, to have them engage in the survey; second, for each student interviewed, a random number generator determined whether their parent was to be interviewed as well.

Table 1 also provides a gender breakdown of the sample. Around 60 percent of the entire sample consisted of female participants. This ratio was somewhat consistent across the sub-samples of logframe, treatment, and control. The proportion of females among participants, however, varied across sub-groups. Whereas the majority of primary students were male, more than 85 percent of primary teachers were female. In fact, all primary teachers within the control group were female.

TABLE 1: QUANTITATIVE SAMPLE: TOTAL PARTICIPANTS SAMPLED AND ACHIEVEMENT TOWARDS SAMPLING GOAL

Participants	Sub-samples									Total		
	Treatment			Control			Logframe					
	All	F	M	All	F	M	All	F	M	All	F	m
Primary teachers	65 (92%)	51	14	72 (100%)	72	0	205 (188%)	170	35	277 (110%)	242	35
PE teachers	N/A	N/A	N/A	N/A	N/A	N/A	83 (130%)	40	43	83 (130%)	40	43
Counsellors	N/A	N/A	N/A	N/A	N/A	N/A	67 (102%)	39	28	67 (102%)	39	28
Primary students	181 (79%)	87	94	252 (100%)	143	109	673 (374%)	362	311	925 (140%)	420	505
Secondary students	N/A	N/A	N/A	N/A	N/A	N/A	283 (55%)	192	91	283 (55%)	192	91
Parents	61 (57%)	39	22	84 (70%)	56	28	240 (68%)	166	73	324 (69%)	222	102
Total	307 (103%)	177	130	408 (126%)	271	137	1551 (120%)	969	581	1959 (103%)	1155	804

Note: The treatment sample was included in the logframe sample. For each sub-group, the total sample is equal to the sum of the logframe sample plus the treatment sample. Figures in brackets indicate the percentage of the (sub-)sample achieved. 100% indicates the entire sample has been completed as planned. Any figure below 100% indicates underachievement while any figure above 100% indicates overachievement.

TABLE 2: LOGFRAME SAMPLE BY TERRITORY

Participants	Territories						Total		
	Gaza			West Bank					
	ALL	F	M	ALL	F	M	ALL	F	M
Primary teachers	233	207	26	44	35	9	277	242	35
PE teachers	66	32	34	17	8	9	83	40	43
Counsellors	40	22	18	27	17	10	67	39	28
Primary students	775	402	373	150	103	47	925	420	505
Secondary students	168	84	84	115	108	7	283	192	91
Parents	294	198	95	30	24	6	324	222	102
Total	1576	945	630	383	295	88	1959	1155	804

TABLE 3: CLASS BREAK DOWN BY SAMPLING GOAL AND TERRITORY

	Primary Classes	Secondary Classes
Treatment	32	N/A
Control	43	N/A
Logframe	167	47
Gaza	160	28
West Bank	50	20

QUALITATIVE SAMPLING FRAMEWORK

The qualitative sample used semi-structured interviews (SSIs), key informant interviews (KIIs), and FGDs to capture the experience of different stakeholders in both Gaza and West Bank. The qualitative aspects of the evaluation were designed to complement the quantitative aspects, explain the data generated by the quantitative methods, and triangulate the findings. Table 4 outlines the type of qualitative method used with each participant.

Interviews and FGDs were only used in schools that were part of the treatment or logframe sample; they were not conducted in any of the control schools. The participants involved in qualitative data collection were selected through a purposeful random selection that gave sufficient weight to the two territories of intervention (Gaza and West Bank), the geographical distribution in each of the territories (five governorates in Gaza⁵ and three areas in the West Bank⁶), the gender of participants, and the type of the staff involved (counsellors, teachers, principals, parents, etc.). It also gave weight to the type of school (co-education, boys-only, or girls only).

Students only participated in the quantitative part of the evaluation as time and resource constraints prevented them from being included in the qualitative part of the evaluation. Meaningful involvement of students would have necessitated a much larger number of student interviews and FGDs to be conducted in comparison to the number of staff and parents. Including students, especially of younger ages, would have required the design of several additional tools and careful attention to other variables, including age, disability, school performance, gender, etc., as well as adherence to robust safeguarding and research ethics. See the recommendation for quality participatory research to be undertaken with students as a key part of any final evaluation of the MA'AN project.

Counsellors, teachers, principals, and education supervisors were pre-selected to represent the categories described above. Attempts were made to select principals, teachers, and other staff from the same schools wherever possible. Parents were selected in collaboration with the school counsellors based on an already established relationship and trust, since parents were usually invited to participate in regular meetings organised by the counsellors. This approach was also adopted to increase the response rate of participants and to avoid sharing parents' contact details with data collectors without the consent of parents.

⁵ This includes North Gaza, Gaza, Middle Area, Khan Younis, and Rafah governorates.

⁶ This includes North, South, and Jerusalem governorates.

TABLE 4: QUALITATIVE SAMPLE

Target participant	Gaza	West Bank	Total	Objectives
Semi-Structured Interviews				
Education Supervisors	6	1	7 5 F 2 M	Understand the change in teachers' knowledge, attitudes, and practices in relation to gender equality in school environment as a result of the R2P intervention and reflect on the approach used in the project (training, coaching, and the CoPs).
Teachers and Counsellors	8	8	16 11 F 5 M	Understand their experience of being part of the project, the perceived impact of the project, its strengths/shortcomings, and future recommendations.
Key Informant Interviews				
Principals	5	5	10 7 F 3M	Understand overall experience of the project, with a specific focus on GR-PBL at a school level and on school environment, its impact (based on feedback from parents and school staff), and differences in implementation across phases/years.
Focus Group Discussions				
Parents/ Caregivers	51 44 F 7 M (5 FGDs)	26 26 F 0 M (2 FGDS)	77 70 F 7 M (7 FGDs)	Explore the change in parents' attitudes on gender equality, their awareness of referral mechanisms, and the impact of awareness-raising workshops on tackling gender-related issues.

Note: FGDs included five female-only groups and two mixed groups.

Survey Design and Tool Development

A review of the quantitative and qualitative survey tools used at baseline demonstrated several limitations. In response to MA'AN staff's request to improve the tools, all of them were significantly updated and then validated at the training for field researchers. The revisions included adding gender-sensitive language, clear interview scripts, and skip patterns, and rewording questions to clarify the intent. The revisions contributed to delays in getting necessary approvals from UNRWA to begin data collection and affected the ability to draw comparisons with quantitative data from the baseline. The revisions and their implications are outlined below.

QUANTITATIVE TOOLS

In total, seven different survey tools were employed to collect quantitative data from each of the relevant sub-groups (see Table 5 below). To ensure comparability between baseline and midline, the aim was initially to rely on the tools employed at baseline, with minimal revisions. However, there ended up being key divergences.

Some baseline data collection relied on in-class observations. However, these tools were not used in the evaluation. In addition to being unfeasible and beyond the scope of what was agreed upon for the evaluation, there were methodological issues with in-class observations as there was no documentation of the observation protocols employed at baseline. The efficacy of observations is largely thought to depend on protocols and principles through which they are organised (e.g., the length and frequency of student observations in classrooms) and thus any attempt to develop a new observation protocol might have resulted in different results than what was reported at baseline. Indicators collected at baseline through observation tools were replaced with self-reporting tools. This might have eased logistical constraints at midline but may have reduced the comparability of baseline and midline, at least to some extent.

Further, a review of the quantitative tools employed at baseline also led to the decision to revise them. While quantitative baseline tools included basic question and answer options, the entire flow and skip pattern were missing. Further, questions were not worded with gender-responsive language nor was there a script that enumerators could use. All tools were carefully reviewed and revised to have a stronger gender-responsive lens and gender-sensitive language. To ensure maximum standardisation of quantitative data collection, the baseline tools used at midline were further expanded into detailed data collection tools with explicit instructions and survey logic (e.g., skip patterns, etc.). While these changes greatly strengthened the tools, it may have affected the comparability of the baseline and midline results. The final midline tools were translated into Arabic, cross checked, and submitted to Right to Play and UNRWA for feedback and eventual approval. UNRWA provided additional and invaluable feedback on form, structure,

and content of the survey tools. These feedback points were incorporated and then all survey tools were programmed into the Kobo Toolbox and administered by enumerators in the field using tablets.

TABLE 5: QUANTITATIVE SURVEY TOOLS

Tool Name	Completed With
Child survey 1–4 (Included Well-being Survey)	Primary students in grades 1–4
Child survey 5–9 (Included Student Learning in Emergency Checklist)	Secondary students in grades 5–9
Parent survey	Parents of both primary and secondary students
Teacher survey	Teachers of primary students in grades 1–4
Counsellor survey	Counsellors working with both primary and secondary students
PE teacher survey	PE teachers working with both primary and secondary students
School Assessment Form	Principals

QUALITATIVE TOOLS

The qualitative tools were also significantly updated from the tools used in the baseline, to reflect the objective and methodology of the evaluation and priorities of MA'AN staff. The main objective of the qualitative methods employed in the evaluation were to provide clarification, contextualisation, and interpretation of the quantitative data found at midline. Thus, the qualitative data at midline are not compared with the qualitative data collected at baseline.

Once approved by Right to Play and UNRWA, minor revisions were introduced (only in the wording of some questions), based on the feedback provided by the field researchers as part of the training and validation process.

TABLE 6: QUALITATIVE SURVEY TOOLS

Tool Name	Completed With
Key Informant Interview Guide	Principals
Education Supervisors Semi-structured Interview Guide	Education Supervisors
Teachers' Semi-structured Interview Guide	Teachers
Counsellors' Semi-structured Interview Guide	Counsellors
Parents' Focus Group Discussion Guide	Parents

Data Collection

All data collection was managed and conducted by the Centre for Development Studies at Birzeit University. Data collection in Gaza began on 3 December 2022 and concluded on 29 December 2022. Data collection in the West Bank began on 3 December 2022 and concluded on 2 March 2023. Coordinating data collection with UNRWA schools, specifically in Gaza, went very smoothly. Schools were cooperative and facilitated data collection. Training sessions prior to data collection instructed enumerators on the overarching project, the specific tools and approach that would be used, and on child safeguarding and research ethics. Quality assurance checks demonstrated that enumerators complied well with instructions and survey and interview administration.

However, despite these positive aspects of data collection, there were considerable challenges in both locations. There was a UNRWA local workers strike in the West Bank from 5 December 2022 through 8 February 2023 and again from 5 March 2023 onwards, which prevented access to schools. The political and security situation was escalating throughout the evaluation in both locations, but particularly in the West Bank. Further, as outlined above, the design and implementation of the MA'AN project necessitated a complex sampling process to begin with. This was then further compounded by multiple revisions to the sample once data collection had begun—both in response to requests from MA'AN staff and in response to the situational challenges described above. These challenges caused a five-week delay in data collection.

Analysis

QUANTITATIVE ANALYSIS

Each of the sub-groups (i.e., principals, primary teachers, PE teachers, counsellors, primary and secondary students, and parents) were surveyed with different tools (see Table 5 above for list of tools).

The survey tools centred around the different scales developed to operationalise the latent constructs implied by the project's outcome indicators. Latent constructs constitute phenomena that are not directly observable (e.g., gender attitudes). The scales used to operationalise the latent constructs were largely the ones developed at baseline. So as not to further jeopardise the comparability between baseline and midline, it was decided to rely on the baseline scales to the extent possible. The survey tools employed were programmed into Kobo and used on tablets during data collection. Data was then uploaded daily into Kobo Toolbox's secure online server. A Power BI dashboard was developed to monitor progress of the data collection.

Due to the delays in data collection, only a minimum amount of time was left to carry out the statistical analysis. Upon retrieving and cleaning the data, the statistical analysis focused on comparing key sub-groups: Gaza with West Bank, female with male participants, treatment with control, baseline with midline, and the different school types. This was done using a Pearson's Chi-squared test with Yates' continuity correction.⁷ There are a few caveats with this that should be noted: Student data may be hierarchical. Since teachers function as the unit of analysis, students were recruited via the classes of their teachers. Thus, students are grouped at the teacher level. As a result, similarities between students being taught by the same teacher/interacting with the same counsellor would be expected. In technical terms, this would manifest itself through clustered standard errors. Variance component models allow testing the level of dependency within the data. While hierarchical models allow robust estimations of effects when the data is hierarchical, given time constraints, this was not possible within the current analysis. Thus, the estimated effects reported below may constitute upper bounds at most. All data analysis was performed using R script.

To compare baseline with midline, baseline values were first pulled from the baseline evaluation report. For each of the indicators, the baseline report provided baseline figures. The Pearson's Chi-squared test, a non-parametric test, was performed. This test does not require any standard errors and instead contingency tables were constructed. To obtain count data for the baseline (e.g., how many respondents that exhibited behaviours implied by a given indicator), the proportions observed at baseline were multiplied with the size of the corresponding sub-sample (both of which were reported in the baseline report).

Performance Measurement Framework Survey data were analysed for all outcome-level indicators and according to the MA'AN PMF (see Table 7 and Annex 1). Outcome-level indicators fall into immediate and intermediate indicators. The assumption here is that immediate outcomes may represent avenues through which intermediate outcomes may materialise.

⁷ Contingency tables were analysed that always compared an independent categorical variable (e.g., baseline versus midline) with a dependent categorical variables (e.g., teacher (not) demonstrating GR-PBL methodologies).

Within the PMF, the project’s immediate and intermediate outcomes are divided into different thematic domains. For each of the indicators, the PMF also specifies the data sources through which progress against any of the indicators is determined. Not all indicators are measured through the evaluation as several rely on monitoring data that the quantitative analysis reported here did not have access to. The thematic domains associated with the immediate and intermediate outcomes as well as their data sources are outlined in Table 7. Annex 1 presents the operationalisation of the indicators through the different survey questions. The operationalisation of each indicator was derived from the baseline survey. To determine operationalisation, the baseline sets were analysed. Thus, the way each indicator was measured in the context of the midline was a continuation of the way each indicator was measured in the context of the baseline.

TABLE 7: PROJECT OUTCOME DOMAINS AND INDICATOR, BY DATA SOURCE

Outcome Domains/Indicators		Data Source
Ultimate Outcome		
1000 Improved quality of education children, especially girls, enrolled in UNRWA schools in the West Bank and Gaza Strip		
1000.1: %/total girls and boys with feelings of social connectedness (m/f)		Student survey (1–4)
		Student survey (5–9)
1000.2: %/total girls and boys that report an overall improvement in their psychosocial well-being (m/f)		Student survey (1–4)
		Student survey (5–9)
Intermediate Outcomes		
1000.3 %/total of classrooms that meet criteria for emotional safety/emotional well-being		Student survey (1–4)
1100 increased integration of GR-PBL by primary teachers into teaching practice		
1100.1 %/total of teachers demonstrating GR-PBL methodologies (m/f)		Teacher survey
1100.2 %/total girls and boys who feel empowered to participate in classroom activities (m/f)		Student survey (1–4)
1100.3 %/total of girls and boys who have a “positive” perception of their teacher (m/f)		Student survey (1–4)/Student survey (5–9)
1100.4 %/total of girls and boys that exhibit context-appropriate social/emotional skills (m/f)		Student survey (1–4)

Outcome Domains/Indicators		Data Source
1200 Improved resilience of girls to respond to and overcome conflict and violence, including GBV (intermediate outcome)		
1200.1 %/total of counsellors demonstrating play-based learning methodologies in their psychosocial support activities for students (m/f)		Counsellor's survey
1200.2 %/total of PE teachers demonstrating play-based learning methodologies in their psychosocial support activities for students		PE teacher survey
1200.3 % total of girls and boys who use school counselling mechanisms and GBV services (m/f)		Student survey (5–9)
1200.4 total of girls and boys reporting positive attitudes towards ending GBV (m/f)		Student survey (5–9)
1200.5 %/total of girls and boys that exhibit context-appropriate social/emotional skills (m/f)		Student survey (1–4)
1300 Improved response of school officials, UNRWA staff, mothers, fathers, and community members to the gender-based education needs of girls and boys, and GBV		
1300.1 %/total of UNRWA staff who efficiently use existing referral mechanisms to address gender equality and monitor the safety of girls and report on GBV (m/f) Quantitative, and triangulated through KIIs and FGDs		Teacher survey
1300.2 %/total mothers, fathers, caregivers, and community members responding to gender-based education needs of girls and boys (m/f)		Parent survey
1300.3 %/total of girls and boys who receive focused psychosocial and GBV services (m/f)		School assessment form
1300.4 % total of teachers who receive support from education supervisors to deliver gender-responsive learning		Teacher survey
Immediate Outcomes		
1110 Improved skills of teachers to implement GR-PBL into teaching practice (m/f)		
1110.2 %/total of teachers that know the strategies to include girls in educational activities (m/f)		Teacher survey

Outcome Domains/Indicators		Data Source
1310 Improved capacity of school principals, education supervisors, and UNRWA staff to ensure gender-responsive learning environment in schools (immediate outcomes)		
1310.1 %/total of education supervisors and school principals with positive attitudes towards ensuring gender-responsive learning environments in schools (m/f)		School Assessment Form
1320 Increased awareness of mothers, fathers, and caregivers on how to support gender equality and psychosocial well-being of girls and boys, and respond to GBV (immediate outcomes)		
1320.1 % total of mothers, fathers, and caregivers with knowledge of where psychosocial support services are provided and where and how to formally report GBV (m/f)		Parent survey
1320.2 % total of mothers, fathers, and caregivers with positive perception of gender equality (m/f)		Parent survey

QUALITATIVE DATA

Qualitative data were thematically analysed. Each of the main evaluation questions and criteria were linked to a set of questions across the different qualitative methods. The answers to questions were grouped and disaggregated by gender, age-cohort, location, and respondent category and then coded and analysed using the MAXQDA package. FGDs, KIIs, and SSIs were audio recorded where possible. In some instances, especially in Gaza, school staff did not agree to audio record the interviews, and therefore an amendment was made to ensure that comprehensive data was collected. This involved working in teams of two field researchers, whereby one facilitated the interview while the other took extensive notes.

Full transcripts/detailed summaries were then generated in Arabic by the field researchers and uploaded to a secure cloud. The lead qualitative researcher (native Arabic speaker) coded the data and analysed it using the MAXQDA software. The data was first coded based on the questions in each of the five qualitative tools, and these codes were linked to the evaluation criteria. Then, data from different stakeholders were read together in relation to each of the evaluation criteria. Data was also analysed by gender, territory, and education level. The analysis was then reviewed by team members to identify inconsistencies or areas that needed further exploration and clarification.

The delays in collecting qualitative data, due to the escalating political situation and violence and the UNRWA local workers strike, impacted the process of data analysis with the bulk of data from Gaza analysed in January 2023 and West Bank data analysed in March 2023. This may have resulted in interpreting different datasets in a slightly different way.

VALIDATION WORKSHOP

Initial findings were validated in a hybrid workshop held on 28 January 2023 with MA'AN and Right to Play staff as part of the MA'AN annual planning meeting. The validation workshop served as a platform to review the **preliminary** results from both the quantitative and qualitative data collected in Gaza **only** with the UNRWA local workers strikes precluding any data being available from the West Bank. The workshop was further used to propose initial recommendations and possible next steps for the MA'AN project and similar Right to Play projects more generally. Findings were discussed and clarified, and possible next steps were identified.

Limitations

Before detailing the findings from the evaluation, the following outlines key considerations of the evaluation's methodology and design and how to interpret results in light of this.

STUDENT PARTICIPATION

Although the Proteknôn-Birzeit consortium proposed qualitative research with students, the time and budget constraints on the evaluation limited their participation to the survey only, as it prevented the necessary child protection and ethical reviews to ensure their meaningful and safe participation. However, since students are the group that the intervention is designed to ultimately impact, not having their participation limits the evaluation's understanding of its impact on students.

BASELINE TO MIDLINE COMPARISONS

To assess the project's effectiveness and the extent to which it is on track to achieve its intended outcomes, the evaluation compares attainment at midline to baseline. However, there are limitations to this comparison: First, changes over time can generally **not** be attributed to the project. The observed effects are always the product of the intervention interacting with the contexts in which they are administered. The best example of this is the COVID-19 pandemic, which affected implementation. Thus, any changes (or lack thereof) between baseline and midline cannot solely be attributed to the project but also in part explained by the consequences

of the pandemic. Second, the revisions made to the survey tools changed the measurement of some indicators between baseline and midline, which reduces the comparability of results.

As outlined above, baseline values were compared with midline values using Pearson's Chi-squared test with Yates' continuity correction. Within this non-parametric test, total samples were focused on without disaggregating by gender. These tests did, however, identify significant results, which will be discussed below in the findings.

ASSESSING PROJECT IMPACT BETWEEN TREATMENT AND CONTROL

To assess the extent to which the project has contributed to the main outcomes, the evaluation compared attainment of treatment schools with control schools. Within impact evaluations, a treatment/control comparison is considered the gold standard. However, there were some limitations in how the control and treatment groups in the MA'AN project were formed.

1. Even though *staff* (primary teachers, PE teachers, and counsellors) were chosen as the unit of selection and analysis, the treatment/control comparison had to rely on the *schools* that were earmarked as control schools at baseline. Despite constituting the project's control sample, some teachers within control schools had already been subject to some project interventions. This may imply some risks for spill-over effects in that those teachers who made up the control group may have already taken away some learning from their colleagues who had already participated in or been exposed to at least some of the project interventions. If this is the case, then the results of the treatment/control comparisons may constitute the lower bounds of the "to be expected" project results. It was beyond the scope of this evaluation to run baseline diagnostics to test whether the random assignment at baseline was successful in ensuring that there were no systematic differences between the groups.
2. The treatment/control comparison concerns schools in Gaza only. Thus, the counterfactual design of this study does not allow any inferences about schools in the West Bank.
3. The treatment/control evaluation design involved the primary school level only. Thus, conclusions about the effects of the project on outcomes at the secondary level cannot be drawn.
4. The evaluation design of the project has not been set up in a way that allows attributing effects observed to specific project interventions. Thus, the results discussed below cast light on whether treatment status is generally associated with (positive) effects. However, they do not cast light on what aspect of the project may have led to that effect.

QUALITATIVE INTERVIEWS

The qualitative findings of the evaluation should be read as complementary to the quantitative evidence provided.

The qualitative data and findings are based on the *perceptions* of those who have been involved in the project. The findings were generally consistent, and this points to a high level of reliability. The reliability of these perceptions depends on many factors, including the experience of the stakeholders, the trust established as part of the data collection process, and the general school internal culture, including the extent to which it allows for openness and transparency. The evaluation team estimated that these findings are generally reliable, especially when balancing the positive and negative perceptions about their experiences, which were shared by the research participants.

Although the selection of participants in the qualitative evaluation draws on the differences in the targeted participants (in terms of geographical location, demographic characteristics, type of schools, etc.), the qualitative component of the evaluation cannot claim to be fully representative of the views of all involved in this project. This is in part because of the relatively small *sample* of participants (32 interviews and seven FGDs of 7–10 participants⁸ each) in relation to the overall number of beneficiaries.⁹ However, inherent to most qualitative research, the approach in this evaluation included, is that it does not aim for generalisation and full representation. The purpose of the qualitative evaluation fulfilled its intended goal to analyse the reported perceptions and general attitudes of the different stakeholders to explore key themes in more depth and complement and expand upon the quantitative data.

Because students were not included in the qualitative part of the study, qualitative findings about project outcomes in relation to them are reported based on adult stakeholders who took part in the interviews and FGDs (teachers, principals, education supervisors, counsellors, and parents).

There are no major differences found in qualitative responses based on the territory (Gaza or West Bank), apart from the contextual differences related to the political and economic situation, with Gaza experiencing harsher living conditions in general. Thus, unless specifically noted, findings represent both territories.

⁸ One of the FGDs was organised with 17 parents although the plan was to include 7–10 participants in each FGD. This change happened to respond to the time limitation on data collection, especially given that data collectors attempted to collect as much data as they could before the planned local workers strikes in the West Bank.

⁹ The project targeted 370 UNRWA-run school, more than 1,830 school staff, and more than 123,750 students.

THE CONTEXT

Lastly, the project and this evaluation were implemented under incredibly challenging circumstances. Project roll-out occurred in the context of a) a pandemic, b) increasing violence in both territories, c) increasing economic insecurity, and d) several local workers' strikes. While the impact of these cannot be fully captured, all achievements described below are particularly noteworthy.

Findings

The evaluation found the MA'AN project to be relevant, effective, and impactful. It responded to felt needs in the community and was widely appreciated and well received. The project helped to build the capacity of teachers, counsellors, and education supervisors to implement GR-PBL. The evaluation demonstrated positive impacts on students' social-emotional well-being and improvements in the learning environment. The findings of the evaluation are discussed below in response to each evaluation question.

To what extent does the project respond to the current and future needs and priorities of the affected population?

Note: This section uses qualitative data from FGS, KIIs, and SSIs to describe how participants view the relevance of the project.

RELEVANCE TO THE NEEDS OF THE POPULATION

MA'AN responded to relevant concerns identified by both UNRWA as well as other stakeholders, including school staff and parents in the targeted schools and communities. MA'AN's relevance was particularly noted regarding concerns about the quality of education, psychosocial well-being, and gender equality. The below section discusses MA'AN's relevance to each of these areas as reported by these stakeholders.

SCHOOL PERFORMANCE

"Low" school performance was reported as a challenge by all interviewed stakeholders with higher prevalence among boys than girls. Low school performance can be attributed to many factors, including the school environment and quality of teaching as well as the social, economic, and political conditions under which children and their families live. These conditions have

resulted in school drop-out, early marriage of girls,¹⁰ and psychological stress and anxiety. All of these have a direct impact on the quality of education in schools. Other factors play a role, including high unemployment rates (in Gaza these are higher than in the West Bank), which have contributed to devaluing school education.

However, despite these conditions and factors, research participants think that education is important, and school is the right place for children to be. The value placed on education was specifically prominent among participants in rural and camp settings. Education in these settings is particularly relevant due to the vulnerabilities and experience of displacement and is perceived as a means to improve living conditions through the social status and employment opportunities that are linked to education.

PSYCHOSOCIAL CHALLENGES

Parents and counsellors reported that they regularly interact with children experiencing psychosocial and psychological challenges such as anxiety, stress, panic attacks, phobias, and depression. They attributed these to the presence of the Israeli occupation and the persistent uncertainty in their lives. These concerns were more frequently reported in Gaza than in the West Bank. This is in line with UNRWA's reporting in its Strategic Plan 2023–2028. According to UNRWA, “over a quarter of students entering Grade 1 in Gaza [display] symptoms of severe trauma”, and increasing psychosocial services was outlined as one of the needs for achieving objective 3: Palestine refugees complete inclusive and equitable quality basic education.¹¹ This is also reiterated in the UNRWA commitment to “support student well-being and inclusion, through its school health strategy, continued efforts to address violence in schools and strengthened PSS services.”¹²

In Gaza, these concerns were linked to frequent military operations carried out in the last 10 years, and in particular, the last one in May 2022. In the West Bank, similar concerns were reported more frequently in camp settings, which are usually more exposed to night raids and incursions. That the MA'AN project took place during this critical period makes it even more relevant.

¹⁰ Student dropout (including girls) and its link to early marriage was reported in at least seven interviews and FGDs.

¹¹ UNRWA (2023). Strategic Plan 2023–2028. p.34-35.

https://www.unrwa.org/sites/default/files/content/resources/strategic_plan_2023-2028.pdf

¹² Ibid. p.36

GENDER EQUALITY

Palestinian society is usually described as having conservative and traditional views on gender and gender roles. Stakeholders shared views that both countered and supported this. Most participants affirmed that both boys and girls are equally entitled to the right to education and affirmed their support for children's equal access to education, including higher education and other educational opportunities. However, participants, especially parents, also acknowledged that social norms place barriers on girls' education and employment opportunities.¹³ At times, their own expressed views reinforced these norms. For

example, mothers and fathers expressed support for their daughters' education *as long as* it does not take place abroad. They did not express similar concerns about their sons' education.

EXAMPLE: "Our society does not put mercy [does not help] when a girl is working in hospital, and they consider this shameful, but when a woman gets sick, she asked a female doctor to treat her [but they don't accept that this doctor is their daughter]." - 33-year-old mother, Focus Group Discussion, Rafah, the GS

Parents generally had a negative attitude toward mixed-gender (co-ed) schools, although they were more accepting of them for the early school years. The majority thought that separate gender schools were safer and more suitable for their daughters, especially given the dominant perception that boys tend to be more violent and thus represent a risk to girls. This represents a structural barrier against girls' participation in mixed schools but also in other spaces that are perceived to be male dominated (including the street and marketplaces, etc.). Further, there was a dominant perception that not all jobs are equally suitable for boys and girls. Participants in the qualitative evaluation, especially parents in FGDs, reported jobs that are typically suitable for girls include teaching, translation, and IT jobs, but not jobs that are physically demanding or those that require working for extended hours or during the nights like doctors, nurses, police officers, drivers, etc. Participants reported several reasons for this, including preserving a girl's honour and respectability and ensuring societal acceptance, which are considered crucial for a girl's reputation and marriageability. There were no major differences in these attitudes shared in the FGDs in both the territories, possibly because students who go to UNRWA schools usually live with their parents in the camps and it is likely that a similar culture is developed in the camp setting. This finding was contrary to the assumption that attitudes in Gaza tend to be more

¹³ Although rates of enrolment at school are generally high in Palestine (97 percent according to PCBS, 2022), especially for the early years of school, learning opportunities become more limited as children grow up (79 percent in 2021/2022 school year). This is particularly relevant to girls whose navigation of the public space becomes restricted. So even when girls have access to school, their access to other extra-curricular and cultural activities in other institutions, which is part of the overall learning experience, is limited.

conservative and stricter. Many mothers in FGDs reported that they dreamed of working in the police force, being firefighters, or studying and working abroad, but this was not always considered acceptable by their families and/or by society in general. Although many, especially the more educated parents and school staff, acknowledged that this is a misconception and that women and men should have equal access to education and employment, they admitted that societal norms and tradition can prevail and prevent these values from being actualised.

Although the society is generally conservative and traditional, it was highlighted that changes are taking place. Discussions in the FGDs and interviews with staff pointed to the fact that restrictions on women's work have been eased and there are more girls at universities than ever before. Many attributed this change to civil society activities, including awareness raising and training. Others explained that the new generation is being given more opportunities as their parents (mostly educated) failed to achieve their dreams. For these parents, supporting their children in their choices is a way to see their own dreams being realised. A parent gave an example of how her parents had changed their attitudes on which disciplines women can or should study at university.

She reported that she wanted to study laboratory medicine. Her father disagreed, and his reasons were that women and girls should not be dealing with tests and examinations (possibly the ones related to reproductive and sexual health). After her marriage, she needed to undertake many examinations for her reproductive health, and she was examined by lots of male doctors. The father found out about this and was annoyed, and therefore subsequently allowed her sister to study laboratory medicine.

EXAMPLE “Amera (pseudonym) is a 41-year-old counsellor in one of the schools in Gaza. She reported that she met an anxious child who was ashamed and reported with anger that another student mentioned his mother’s name. The counsellors explained to the child that the use of his mother’s name, or any women or girls’ names, should not be shaming, neither to the student, his mother, or anyone else.”
- *Alshejaea Camp, Gaza City, Palestine*

Overall, concepts of honour, shame, and pride were central to discussions around equal opportunities to education. It was interesting to see that religion was not stated as one of the key factors behind these conservative views. Words like “haram” or “halal”, which refer to what is allowable and what is not within Islamic traditions, were not mentioned at all in any of the interviews and FGDs’ transcripts.

RELEVANCE OF THE INTERVENTION

The MA'AN project is seen as an appropriate and relevant intervention to address the expressed needs of stakeholders. Many reported that the pandemic, the changing political situation, and economic hardship—which increased stress, anxiety, and isolation, and limited school attendance time—made the MA'AN project even more relevant. Its relevance was reflected in many aspects of the project, including:

- Attendance and engagement of parents in meetings organised by the school as part of MA'AN.
- Commitment of school staff, especially teachers and counsellors, to the activities and objectives of the project and their willingness to try and adapt practices to their local settings and culture.
- Alignment of the MA'AN intervention to existing practices in UNRWA schools, including curricula for human rights, referral mechanisms, and the move to adopt play-based activities in the classrooms.
- Participants' calls to increase and expand the project.

These aspects are explored further in the findings below relating to project effectiveness, impact, and sustainability. The project also aligns with UNRWA's commitment to "make its programmes more gender sensitive so that services it provides are equitable, ensuring that all Palestine refugees [including] girls and boys can fulfil their needs and enjoy their rights."¹⁴ This is also reflected in considering gender as one of the key thematic imperatives that should be reflected in its systems, structures, processes, and programmes.¹⁵

The evaluation team concluded that the intervention of MA'AN aligns both with UNRWA's strategic goals, and the needs of the community in relation to quality of education, psychological challenges, and gender equality. The different components of the project were seen relevant to existing challenges in the schools and communities, and these were priorities for educators and UNRWA administration alike.

¹⁴ UNRWA (2023). Strategic Plan 2023–2028. p.5.

https://www.unrwa.org/sites/default/files/content/resources/strategic_plan_2023-2028.pdf

¹⁵ Ibid. p.6

To what extent is the project on track to achieve its intended outcomes?

Note: This section focuses on understanding the **project's effectiveness**. To do that:

- It uses **quantitative data** from **surveys** to detail the **progress** beneficiaries have made on targeted indicators **from before the project began** (baseline evaluation) **until now** (midline evaluation).
- **This data** is based on the “logframe sample” described earlier, which was a **random sample of participants** from **primary** and **secondary schools** in **Gaza** and **West Bank**.
- It also uses **qualitative data** from FGS, KIIs, and SSIs to describe **how participants view** the **effectiveness** of the project.

OVERVIEW

The evaluation suggests that MA'AN is **improving in most outcome areas**. In survey results, out of all indicators where tests were possible, **12 exhibited significant improvements between baseline and midline**. Interviews with research participants demonstrated that MA'AN is perceived to be effective in building the capacity of **teachers, counsellors, and education supervisors** to implement GR-PBL to improve outcomes and gender equality for **students**. This was reflected in discussions about the project's core objectives and in recommendations for future interventions.

The evaluation however also found signs of challenges to MA'AN's effectiveness. Survey results demonstrated three indicators for which there was no discernible differences, and four indicators where there was a decrease between baseline and midline. Further, while research participants reported that the project was largely effective, they also noted challenges limiting its effectiveness, including the pandemic, military escalation, and traditional views regarding gender, play, and inclusion. The following sections discuss in more detail the effectiveness of MA'AN from baseline to midline as well as the perceptions of effectiveness among different stakeholders.

IMPROVEMENTS FROM BASELINE TO MIDLINE

Table 8 below presents the findings of baseline to midline progress.

In comparing baseline with midline values, the evaluation finds **statistically significant results in almost all outcome areas**. This is true for the perceived improved quality of education of children

(1000), increased integration of GR-PBL by primary teachers (1100), improved resilience of girls (1200), improved responses to gender-based education needs (1300), and increased awareness towards gender equality and psychosocial well-being (1320). This in turn suggests that the **project has attained broad-based progress despite the different contextual challenges in which it operated.**

The most significant improvements were seen in outcomes related to school staff. The percentage of counsellors that demonstrated play-based learning methodologies in their psychosocial support activities for students (1200.1) increased by 72 percentage points, and UNRWA staff who efficiently use referral mechanisms to address gender equality and GBV (1300.1) increased by 65 percentage points. The percentage of teachers knowing strategies to include girls in education activities (1110.2), the percentage of teachers who receive support from education supervisors to deliver gender-responsive learning, and the percentage of PE teachers who demonstrated play-based learning methodologies in their psychosocial support activities for students all increased around 40 percentage points. Since staff were the direct recipients of the MA'AN project, this points towards high effectiveness.

SETBACKS FROM BASELINE TO MIDLINE

There were four indicators that demonstrated **statistically significant decreases** in outcomes between baseline and midline (i.e., 1200.3, 1200.4, 1310.1, and 1320.2; all four are indicators related to gender. On three indicators, no discernible differences between baseline and midline were found [1100.2, 1100.4, and 1300.2]).

The decrease observed on some indicators may potentially be explained through contextual factors that go beyond the project. Parents exhibiting positive perceptions of gender equality (1320.2) dropped from 61 percent at baseline to around 40 at midline. The level of students that reported positive attitudes towards ending GBV dropped by nearly the same amount as their parents—decreasing from around 77 percent at baseline to 58 percent at midline (i.e., 1200.4). At the same time, however, the treatment/control comparison that will be discussed in the next section revealed that parents linked to schools associated with the treatment communities were more likely to express positive perceptions around gender equality (see Table 10 in section below). This in turn suggests that school communities not served by the project (i.e., control schools) are doing worse than those communities served by the project. Taken together, both findings may imply that contextual factors (e.g., the COVID-19 pandemic and the associated school closures) may have led to a general deterioration in attitudes towards gender equality within both treatment and control schools. However, the contextual factors may have had a bigger (i.e., more detrimental) impact on school communities not served by the project, which in turn may explain the positive treatment effects. The example of indicator 1320.2 also highlights

why treatment/control comparisons are considered to be the gold standard of impact evaluations. By using a counterfactual, they are able to control for factors a simple pre/post assessment often cannot do.

The decrease found in attitudes of education supervisors and principals was not supported in the qualitative work. As explained below, the contradictory finding might have been due to differences in methodology.

TABLE 8: CHANGES FROM BASELINE TO MIDLINE, WHERE STATISTICAL TESTING IS POSSIBLE

Time	Sub-group	Proportion (in %)	Pearson's Chi-squared test with Yates' continuity correction
1000.1) Girls and boys with feelings of social connectedness			
Baseline	All	31.00	Midline significantly higher at p<0.00
Midline	All	50.41	
1000.2) Girls and boys that report an overall improvement in their psychosocial well-being			
Baseline	All	14.00	Midline significantly higher at p<0.00
Midline	All	28.06	
1000.3) Classrooms that meet criteria for emotional safety/emotional well-being			
Baseline	All	32.00	Midline significantly higher at p<0.00
Midline	All	58.27	
1100.1) Teachers that demonstrate GR-PBL methodologies			
Baseline	All	37.00	Midline significantly higher at p<0.00
Midline	All	61.01	
1100.2) Girls and boys who feel empowered to participate in classroom activities			
Baseline	All	81.00	No discernible difference
Midline	All	87.14	
1100.3) Girls and boys who have a “positive” perception of their teacher			
Baseline	All	47.00	Midline significantly higher at p<0.00
Midline	All	60.86	
1100.4) Girls and boys that exhibit context-appropriate social/emotional skills			
Baseline	All	29.00	No discernible difference
Midline	All	26.49	

Time	Sub-group	Proportion (in %)	Pearson's Chi-squared test with Yates' continuity correction
1110.2) Teachers that know the strategies to include girls in educational activities			
Baseline	All	20.00	Midline significantly higher at p<0.00
Midline	All	62.45	
1200.1) Counsellors that demonstrate play-based learning methodologies in their psychosocial support activities for students			
Baseline	All	22.00	Midline significantly higher at p<0.00
Midline	All	94.03	
1200.2) PE teachers that demonstrate play-based learning methodologies in their psychosocial support activities for students			
Baseline	All	16.00	Midline significantly higher at p<0.00
Midline	All	55.42	
1200.3) Girls and boys that use school counselling mechanisms and GBV services			
Baseline	All	57.00	Midline significantly lower at p<0.00
Midline	All	24.73	
1200.4) Girls and boys that report positive attitudes towards ending GBV			
Baseline	All	77.00	Midline significantly lower at p<0.00
Midline	All	58.30	
1200.5) Girls and boys that exhibit context-appropriate social/emotional skills			
Baseline	All	20.00	Midline significantly higher at p<0.00
Midline	All	38.05	
1300.1) UNRWA staff who efficiently use referral mechanisms to address gender equality and GBV			
Baseline	All	20.00	Midline significantly higher at p<0.00
Midline	All	85.61	
1300.2) Caregivers and community members that respond to gender-based education needs of girls and boys			
Baseline	All	25.00	No discernible difference
Midline	All	19.08	
1300.4) Teachers who receive support from education supervisors to deliver GR-PBL			

Time	Sub-group	Proportion (in %)	Pearson's Chi-squared test with Yates' continuity correction
Baseline	All	31.00	Midline significantly higher at p<0.00
x	All	70.04	
1310.1) Education supervisors/school principals with positive attitudes towards GR-PBL			
Baseline	All	61.00	Midline significantly lower at p<0.00
Midline	All	34.29	
1320.1) Caregivers with knowledge of where psychosocial support services and where and how to formally report GBV			
Baseline	All	42.00	Midline significantly higher at p<0.00
Midline	All	53.85	
1320.2) Mothers, fathers, and caregivers with positive perception of gender equality			
Baseline	All	61.00	Midline significantly lower at p<0.00
Midline	All	40.00	

Note: Controlling for gender within the statistical tests to compare baseline and midline would have resulted in three-way tables that are often perceived to be difficult to interpret. However, Annex 3 presents the baseline/midline proportions broken down by gender. Results in this table seem to largely follow the ones reported in the table above. Midline levels are generally higher for both male and female.

TERRITORIAL VARIATIONS

A comparison between Gaza and the West Bank is limited to the midline data only. Table 9 below presents the results of the statistical tests performed to compare both territories. By and large, both territories were comparable in terms of progress observed on the different indicators. In only six of the 19 indicators were some differences observed between Gaza and West Bank that were statistically significant. While these results demonstrate differences between the two territories, they do not imply that the project has performed better in the West Bank compared to Gaza. It appears that at baseline the West Bank started off with higher levels in various domains of the project activities than Gaza. Therefore, what we are now observing may just be a continuation of what was already observed at baseline.

TABLE 9: DIFFERENCES AT MIDLINE BETWEEN GAZA AND WEST BANK

Time	Sub-group	Proportion (in %)	Pearson's Chi-squared test with Yates' continuity correction
1000.1) Girls and boys with feelings of social connectedness			
Gaza	All	48.25	Levels significantly higher in West Bank at p<0.00 at p<0.00
West Bank	All	58.11	
1000.2) Girls and boys that report an overall improvement in their psychosocial well-being			
Gaza	All	22.80	Levels significantly higher in West Bank at p<0.00 at p<0.00
West Bank	All	46.79	
1000.3) Classrooms that meet criteria for emotional safety/emotional well-being			
Gaza	All	59.23	No discernible difference
West Bank	All	53.33	
1100.1) Teachers that demonstrate GR-PBL methodologies			
Gaza	All	61.80	No discernible difference
West Bank	All	56.82	
1100.2) Girls and boys who feel empowered to participate in classroom activities			
Gaza	All	87.87	No discernible difference
West Bank	All	83.33	
1100.3) Girls and boys who have a “positive” perception of their teacher			
Gaza	All	60.26	No discernible difference
West Bank	All	64.00	
1100.4) Girls and boys that exhibit context-appropriate social/emotional skills			
Gaza	All	22.32	Levels significantly higher in West Bank at p<0.00 at p<0.00
West Bank	All	48.00	
1110.2) Teachers that know the strategies to include girls in educational activities			
Gaza	All	62.66	No discernible difference
West Bank	All	61.36	
1200.1) Counsellors that demonstrate play-based learning methodologies in their psychosocial support activities for students			

Time	Sub-group	Proportion (in %)	Pearson's Chi-squared test with Yates' continuity correction
Gaza	All	97.50	No discernible difference
West Bank	All	88.89	
1200.2) PE teachers that demonstrate play-based learning methodologies in their psychosocial support activities for students			
Gaza	All	51.52	No discernible difference
West Bank	All	70.59	
1200.3) Girls and boys that use school counselling mechanisms and GBV services			
Gaza	All	10.71	Levels significantly higher in West Bank at p<0.00 at p<0.00
West Bank	All	45.22	
1200.4) Girls and boys that report positive attitudes towards ending GBV			
Gaza	All	55.95	No discernible difference
West Bank	All	61.74	
1200.5) Girls and boys that exhibit context-appropriate social/emotional skills			
Gaza	All	31.87	Levels significantly higher in West Bank at p<0.00 at p<0.00
West Bank	All	70.00	
1300.1) UNRWA staff who efficiently use referral mechanisms to address gender equality and GBV			
Gaza	All	86.70	No discernible difference
West Bank	All	80.00	
1300.2) Caregivers and community members that respond to gender-based education needs of girls and boys			
Gaza	All	20.00	No discernible difference
West Bank	All	10.00	
1300.4) Teachers who receive support from education supervisors to deliver gender-responsive learning			
Gaza	All	71.67	No discernible difference
West Bank	All	61.36	
1310.1) Education supervisors/school principals with positive attitudes towards gender-responsive learning			
Gaza	All	36.17	No discernible difference
West Bank	All	30.43	

Time	Sub-group	Proportion (in %)	Pearson's Chi-squared test with Yates' continuity correction
1320.1) Caregivers with knowledge of where psychosocial support services and where and how to formally report GBV			
Gaza	All	52.54	No discernible difference
West Bank	All	66.67	
1320.2) Mothers, fathers, and caregivers with positive perception of gender equality			
Gaza	All	37.63	Levels significantly higher in West Bank at p<0.00 at p<0.00
West Bank	All	63.33	

Note: Annex 6 presents a comparison between both territories that is also disaggregated by gender.

TEACHERS

For teachers, MA'AN was perceived to be effective in building their capacity (skills and knowledge) and confidence as teachers as well as practically helping them in their teaching practices. Teachers reported that the MA'AN trainers were skilled, which made the training more engaging. The training tackled aspects of physical, emotional, social, and cognitive development. The training manual was useful in linking the theoretical and practical aspects together. Overall, they felt that the approach promoted positive thinking and connected play with day-to-day life situations. Receiving feedback in the training and engaging in group work gave teachers concrete ideas of how to use games in their everyday teaching practice. This happened through coaching, which was facilitated during the classes attended by the education supervisors. At least four teachers (out of 10) reported that the education supervisors played a vital role in this process. However, one teacher reported that the role of the education supervisors in following the MA'AN project was not significant and stated that this should not be the role of the education supervisors—rather that there should be a person designated to follow up progress in achieving the outcomes of the MA'AN project. Interestingly, the training gave teachers a space to escape their intense workload and have some time for fun. Further, the MA'AN training fits well with other training organised by UNRWA on inclusive education,

EXAMPLE: Some teachers previously used to divide the students into groups based on their school/academic achievement. After the training, teachers became aware how this may exclude students and deepen differences. Teachers therefore now use random numbers in assigning students into groups, which causes boys and girls to work together more.

managing illiteracy, and learning difficulties. It encouraged teachers to develop individualised approaches to students with disabilities, when needed.

Teachers reported that the training improved the class atmosphere, making it more cooperative and inclusive. This was reported equally in both West Bank and Gaza (at least three out of five teachers reported that in each of the territories). One teacher gave an example of how this cooperative and inclusive atmosphere changed as a result of MA'AN intervention. He said: "There was a student [female] who suffered from a hand injury, the students were competing in providing support and help to that student especially in writing and catching up with the lessons." Existing learning pedagogies and teaching methods are based on narration and repetition. Teachers and education supervisors reported that the training equipped teachers with skills and knowledge to adopt more

EXAMPLE: One teacher (33-year-old, West Bank) gave an example of explaining a lesson on children with disabilities by asking students to play handball while seated. At the beginning, students kept forgetting the rules of the game and jumped to pick up the ball. However, by the end of the class, they had learned more about the challenges children with disabilities face in playing and learning, which was a starting point to speak about the need to address bullying of these children.

inclusive teaching and learning practices. MA'AN's GR-PBL approach was novel and practical in teaching activities that included all students, regardless of their gender, disability status, or level of school achievement. Although this finding may seem contradictory to the reported results under indicator 1110.2, it is worth mentioning that there could be differences between what teachers perceived as a good impact of the project, and what students, on the other hand, experienced in reality. It is equally important to highlight that qualitative tools allowed for more exploration and investigation of themes beyond what was reported in the quantitative evaluation.

Beyond academic work, teachers manage and respond to various other issues and behaviours on a daily basis. Teachers reported that MA'AN helped to instil values and skills like respect, accepting others, forgiveness, and imagination, which helped in addressing behavioural issues. For example, it was reported that bullying and violence are common in schools, but MA'AN helped address this. One teacher (Gaza), for example, reported how teachers used different tools to address bullying in his school, including the use of educational pamphlets, theatre, and role-play, that managed to address 80 percent of cases of bullying.

Many examples were provided of how teachers used the MA'AN training to work with students. They reported that it was easier to guide younger students through dialogue and instruction. For older students, teachers used other techniques to change behaviour and attitudes, including

those related to gender stereotypes and inclusion—for example, using life examples and games. Theatre and other media were also used to increase respect students had for each other. Overall, the teachers reported that MA'AN was empowering in responding to these issues in an inclusive manner.

Although teachers were using games before receiving the MA'AN training, their use of such activities was not structured or organised. Providing GR-PBL training that included both theoretical values and practical applications made teachers more confident and active in using these approaches. The resource toolkits helped in smoothly implementing the activities. This impact was felt by many teachers interviewed. It is also interesting that some reported that this impact was transferred to other colleagues who did not participate in the MA'AN training. The transfer of knowledge happened mainly through the established Communities of Practice (CoPs), as explained below, but also informally (as reported by at least two teachers).

COUNSELLORS

Likewise, counsellors appreciated the activities and resources made available as part of the MA'AN project, reporting that activities-based learning approaches and counselling were more effective than traditional counselling. While counsellors did not seem to have the full picture about the MA'AN project at the outset, gradually through the intervention, they became more aware of its objectives and the potential benefits of including play activities as part of their work. Although counsellors regularly receive other training, the topics of the MA'AN training were particularly relevant in their day-to-day work with students. The structure and content of the training was specifically useful as it includes dealing with parents, referrals, a games toolkit, and practical detailed manuals, all of which were helpful in applying the training in practice. MA'AN also includes various meetings for monitoring and coaching/mentoring visits, a space for feedback, and sessions to raise and discuss challenges. All these have helped in implementing activities and achieving the goals related to the project and have enabled seeing the impact of the activities on students.

It was reported that the topics discussed in the training helped in developing the technical aspects of the work and their overall performance of the counsellors. A specific example given of this is how the MA'AN training helped counsellors work more effectively with teachers: Counsellors closely coordinate with teachers in managing and addressing the student concerns. They meet to discuss individual student cases and plan and implement school-wide campaigns (for example, the 16-day campaign to eliminate GBV). It was reported that MA'AN helped to increase synergy and collaboration in performing these tasks. This collaboration between school staff in general, and between teachers and counsellors in particular, was reported by the majority of staff interviewed (both in the West Bank and Gaza).

EDUCATION SUPERVISORS

Relevant to the role played by the counsellors and the teachers in MA'AN is the work of the education supervisors. Education supervisors are UNRWA staff who visit schools regularly and supervise and monitor teachers' performance. The involvement of education supervisors was reported as positive and effective in achieving the objectives of MA'AN. Education supervisors were central to promoting the core principles of the training, including among those teachers who had not directly participated. Working closely with the education supervisors proved to be an effective way of bringing all the moving parts of the MA'AN project together and ensuring that work implemented throughout the years of the project is sustained (see the discussion on Sustainability). Teachers reported that education supervisors play a supportive role in implementing MA'AN through their class visits, providing support and guidance to apply best practices, and working closely with teachers in organising and changing the planning notebook and the way activities are implemented to ensure active learning. The education supervisors also discuss and provide feedback after the teaching session and advise on other games that could be used for specific topics and subjects as well as sending guiding principles and manuals to support teachers. The importance of their role is noted by views that the frequency of their visits, on average twice a year, are not sufficient and do not allow time for teachers to properly learn and benefit from the extensive experience of the education supervisors.

While findings from the quantitative data showed a decrease in education supervisors' attitudes towards GR-PBL (1310.1), this could partially be attributed to a difference in methods used. One way of reading this contradiction is that other school staff, as reported in FGDs and SSIs, appraise and appreciate the role of the education supervisors, possibly because they are placed higher in the professional hierarchy than teachers; while principals—who reported on the gender-responsive attitudes of themselves and education supervisors—are more critical. Another explanation could result from there being fewer principals and education supervisors, thus subjecting the quantitative finding to more volatility.

VARIATIONS IN STAFF AND PARENTS' EXPERIENCES

Although the selection of participants in the qualitative evaluation was intended to provide a balanced account drawing on a sample that was composed of a 50/50 male/female ratio, the actual fieldwork resulted in more women (school staff and parents) participating in the FGDs and interviews. This happened because the sample included another layer of selection that was based on school type (male-only, female-only, co-ed). The likelihood of having female staff (and not male staff) working in co-ed schools is higher, and this resulted in shifting the gender composition of the sample in favour of female participants. In addition, women (mothers) were more

responsive to the invitation to take part in FGDs than men (fathers). This could be explained by the dominant norm of men being the sole breadwinners in the communities involved. As a result, it was not possible to draw conclusions on the variations in the experience of women and men taking part in the MA'AN project. In any case, unexpectedly, there was not a significant difference between the attitudes of women and men around the questions of gender equality or the use of play in both the parents' groups and the school staff interviews.

CHALLENGES

While overall MA'AN was perceived as effective, especially in the context of the challenges faced during the intervention period, there were challenges that limited its effectiveness.

For counsellors, there are many obstacles to complete work in a meaningful and effective manner. Some of these challenges are structural and beyond the scope of the project, but nonetheless impact its effectiveness. For example, many work on a part-time basis. This results in counsellors often having too many cases that need attention, which means continually postponing work on critical but slightly less urgent cases. It also does not allow for long-term monitoring and supervision of the issues emerging among the students. A counsellor (female, North of Gaza) reported that sometimes these cases are behavioural and related to violence while others are psychological and emotional (anxiety, lack of concentration) or related to school performance.

For teachers and counsellors, the limited infrastructure and physical spaces for play and reading restrict staff's ability to put their skills in practice. In mixed-gender schools, girls are unable to use the available spaces because they are overcrowded and because of the aggressive behaviours that are predominantly undertaken by male students. Many of these issues cannot be addressed by the counsellors, especially due to their limited time and lack of resources. Furthermore, some teachers and counsellors reported that the training added an additional burden on them, especially as it took place during working hours. One recommendation was therefore that the training could be more effective if it takes place during summer breaks.

Many participants in all groups of stakeholders reported that MA'AN fell short of responding to the increasing demands and needs in the schools and the targeted communities. This finding should be read in the context of two other different findings: 1) the persistence of the relevance of the MA'AN intervention to the issues experienced in the schools and communities (see the section on Relevance), and 2) the high expectation of school staff from Right to Play in terms of supporting schools with more material resources, including renovating and equipping play spaces that are not part of the current intervention. In fact, the staff requests and recommendations

included supplying schools with interactive screens or virtual reality devices, which can be seen as relevant to the overall play-based learning experience, but which is not within the scope of the MA'AN project.

Relevant to the first point is the fact that society is still resistant to many of the ideas of the MA'AN project (reported equally in the West Bank and Gaza). For example, while educational supervisors observed that teachers undertook considerable effort to promote the core values and learnings from the training, including with the students' families, MA'AN is still being implemented in a context where gender equality as a topic is controversial and difficult to tackle. This was acknowledged, both in the West Bank and Gaza, by school staff, education supervisors, and parents. Generally, school staff reported that they actively engaged with the issue of gender equality, especially in mixed-gender schools, and the project provided a framework for their work on this area. At the same time, both school staff and parents admitted that more work on gender equality would be needed at different levels, especially in a context where the society is generally resistant to the very idea of equality. In a focus group discussion in Hebron in the West Bank, for example, most of the participants (all women) were against women working as car mechanics, listing several reasons for their views, including that this work is not dignifying for women, the incompatibility of this job with marriage and *femininity*, and that it would allow women to deal with men, which is not generally acceptable. This demonstrates the good work being achieved as part of this intervention, but also the necessity of examining other approaches to effectively change dominant attitudes around gender and inclusion.

To what extent has the project contributed to the main outcomes?

Note: This section focuses on understanding the **project's impact**. To do that:

- It uses **quantitative data** from surveys to **compare** the performance of **treatment schools** to **control schools**.
- This data is based on the "treatment and control sample" described earlier, which only focused on primary schools in Gaza.
- It also uses **qualitative data** from FGS, KIIs, and SSIs to **describe** how **participants view the impact of the project**.

OVERVIEW

In addition to changes between baseline and midline, comparing treatment schools with control schools demonstrates the ways that MA'AN is uniquely contributing to project outcomes. Both primary boys and girls in treatment schools (Gaza only) were more likely to report having **positive perceptions of their teachers** and exhibit context-appropriate **social-emotional skills**. In addition, **teachers** were more likely to self-report the **use of the referral mechanisms** while **parents** and caregivers were more likely to have **positive perceptions of gender equality**. Interviewed research participants emphasised the impact that MA'AN has had, detailing its impact on **school staff** and **students** and **the learning environment**.

IMPACT ON TREATMENT SCHOOLS VERSUS CONTROL SCHOOLS

Table 10 below presents the results of the statistical tests performed to compare treatment and control. Annex 5 presents the respective results disaggregated by gender.

TABLE 10: DIFFERENCES AT MIDLINE BETWEEN TREATMENT AND CONTROL

Time	Subset	Proportion (in %)	Pearson's Chi-squared test with Yates' continuity correction
1000.1) Girls and boys with feelings of social connectedness			
Control	All	39.29	No discernible difference
Treatment	All	46.96	
1000.2) Girls and boys that report an overall improvement in their psychosocial well-being			
Control	All	14.29	Treatment significantly higher at p<0.00
Treatment	All	31.49	
1000.3) Classrooms that meet criteria for emotional safety/emotional well-being			
Control	All	53.17	No discernible difference
Treatment	All	59.12	
1100.1) Teachers that demonstrate GR-PBL methodologies			
Control	All	45.83	No discernible difference
Treatment	All	58.46	
1100.2) Girls and boys who feel empowered to participate in classroom activities			
Control	All	86.11	No discernible difference
Treatment	All	92.27	

Time	Subset	Proportion (in %)	Pearson's Chi-squared test with Yates' continuity correction
1100.3) Girls and boys who have a “positive” perception of their teacher			
Control	All	43.65	Treatment significantly higher at p<0.00
Treatment	All	66.85	
1100.4) Girls and boys that exhibit context-appropriate social/emotional skills			
Control	All	15.48	Treatment significantly higher at p<0.00
Treatment	All	32.60	
1110.2) Teachers that know the strategies to include girls in educational activities			
Control	All	48.61	No discernible difference
Treatment	All	58.46	
1200.5) Girls and boys that exhibit context-appropriate social/emotional skills			
Control	All	27.38	Treatment significantly higher at p<0.00
Treatment	All	47.51	
1300.1) UNRWA staff who efficiently use referral mechanisms to address gender equality and GBV			
Control	All	80.56	Treatment significantly higher at p<0.05
Treatment	All	95.38	
1300.2) Caregivers and community members that respond to gender- based education needs of girls and boys			
Control	All	22.62	No discernible difference
Treatment	All	13.11	
1300.4) Teachers who receive support from education supervisors to deliver gender-responsive learning			
Control	All	56.94	Treatment significantly higher at p<0.00
Treatment	All	83.08	
1310.1) Education supervisors/school principals with positive attitudes towards gender-responsive learning			
Control	All	16.67	No discernible difference
Treatment	All	28.57	
1320.1) Caregivers with knowledge of where psychosocial support services are and where and how to formally report GBV			
Control	All	39.29	No discernible difference

Time	Subset	Proportion (in %)	Pearson's Chi-squared test with Yates' continuity correction
Treatment	All	49.18	
1320.2) Mothers, fathers, and caregivers with positive perception of gender equality			
Control	All	22.62	Treatment significantly higher at $p < 0.00$
Treatment	All	62.30	

Treatment effects materialised across all intermediate outcome areas, particularly around student well-being. Students of teachers within the treatment group were more likely to report social-emotional skills (1100.4/1200.5), to have positive perceptions of their teachers (1100.3), and to report improvements in their overall well-being (1000.2).

In the immediate outcome areas, the treatment/control comparison only found discernible effects in terms of parents' attitudes towards gender equality (1320.2). As already explained above, general agreement levels in terms of gender equality may have gone down between baseline and midline. However, they may have more rapidly declined among school communities not served by the project (e.g., the control schools) than within school communities supported by the project. This in turn may explain the positive treatment effects explained.

Apart from indicator 1320.2, treatment effects were not observed within the immediate outcome areas. This is somewhat surprising because immediate outcomes may be seen as the avenue through which treatment effects in intermediate outcome areas materialise. One explanation for the lack of treatment effects on the immediate level as well as on school staff related indicators could be related to methodology. Indicators for which the evaluation did not find any effects among adult participants are largely reliant on self-reporting. For example, across both treatment and control schools, primary teachers are asked to rate themselves in terms of their ability to include girls in classroom activities (i.e., indicator 1110.2) and parents are asked to indicate whether or not they know how to obtain psychosocial support to address their child's needs without actually having to show how to make use of them (i.e., indicator 1320.1) or whether or not they respond to gender-based education needs (1300.2). By contrast, treatment effects were observed especially on those indicators that relied on (self-)reports by students. If the previous explanation holds true, then this may also imply that students feel more free to report what they truly experience and see. In technical terms, students may be less likely to fall prey to social desirability effects than parents or teachers. Positive effects were also observed in relation to indicator 1300.4, but it required teachers to report on support they received from education supervisors. This in turn may have been a question that teachers did feel able to talk about.

What students report may also say something about their school environment. This may be particularly obvious in the case of indicator 1100.3, which is about students' perceptions of their teachers. At treatment schools, students hold more positive perceptions of their teachers than at control schools. For indicator 1100.3, both female and male students are more likely to report positive perceptions of their teachers than their counterparts in control schools. However, male students seem to have developed more positive perceptions of their teachers than female students. Whereas differences between treatment and control samples were about 14 percentage points in the case of female students, it was about 32 percentage points in the case of male students. Annex 5 presents the gender breakdown of the treatment/control results.

IMPACT ON SCHOOL STAFF

Principals reported that MA'AN has contributed to building the capacity of staff, especially of older teachers and staff who tend to use traditional teaching and counselling methods. As a result of the project, teachers and counsellors are working together more to tackle issues and behaviours observed in the classrooms and reported by parents. The approach of MA'AN seems to be resulting in a more holistic intervention where school resources, including staff, are mobilised to serve children's best interests. Principals stated that there is an improvement in the

EXAMPLE: One participant said that the project contributed to enhancing the existing referral mechanisms and recounted the story of a specific child who had suicidal thoughts and how activating the referral mechanisms has helped in addressing the situation (principal, female, Nablus in the West Bank).

relationship between female and male teachers due to the establishment of ways of communicating based on trust. Teachers have incorporated what they have learned through MA'AN in their activities, and this has generated a significant impact on children and the overall learning environment. Education supervisors reported positive changes in teachers' approaches and attitudes and the use of play-based learning activities in their classes, although not all teachers benefited in the same way or had the same impact, with individual personalities and commitment reflecting the extent of benefit gained from the project. Education supervisors also reported that the response of the teachers improved significantly and, as a result, the incorporation of more marginalised students into games and activities was achieved. Teachers and counsellors themselves also reported increased levels of confidence in using GR-PBL activities in the classroom, especially after witnessing the impact of this approach on their students (*discussed later in this section*). The MA'AN training has equipped them with new tools and skills and has refreshed their knowledge on the stages of child development as well as child rights, pedagogy, and participatory active learning.

The fact that MA'AN did not target all teachers in all the targeted schools was reported as one of the limitations of the intervention. Nonetheless, the impact could be observed beyond those targeted. For example, teachers reported that their colleagues, who did not take part in the training, are also using the training resources and manual to implement GR-PBL activities. This may have been a result of their engagement in the CoPs. Importantly, MA'AN has encouraged schools to use the existing referral mechanisms and to inform parents about them.

THE IMPACT ON CHILDREN AND THE LEARNING AND HOME ENVIRONMENT

EDUCATION SUPERVISORS

Education supervisors agree that MA'AN has resulted in a decrease in violence and an increase in the acceptance of underachieving and marginalised students by teachers. A more gender-responsive and inclusive form of learning has been adopted with one of the education supervisors stating that a change in the schools' internal cultures towards gender equality is a strategic goal to be achieved in the coming two years.

PARENTS

Most of the parents reported that they saw a direct impact of the parents' meetings on the way they deal with their children. Some parents became more aware of positive methods to deal with children. According to parents, the children have become calmer, their behaviour has changed, and for some, children have become more disciplined. Parents have also become more accepting of their children and their school achievement, which can improve behaviour and school achievement. The meetings also changed their

EXAMPLE: One of the parents shared that he used to use punishment to change his child's behaviour before realising through the parents' meeting that this was unfair and unhelpful in changing his child's behaviour.

attitudes toward their children from being characterised by stress and tension to being more patient and adopting a constructive dialogue approach. For example, the meetings provided a starting point for a dialogue and discussion between the two generations as parents engaged in discussing the topics learned with their children. While the parent meetings were viewed as impactful, some parents reported that meetings were not conducted, and they were not invited to take part. This could signal that project implementation was not uniform across all schools and could be the result of limited staff capacity to run meetings. However, given the reported impact of these meetings, this could be an area that warrants more attention. It was not however

possible to disaggregate these findings by territory, especially since not all planned FGDs in the West Bank were conducted due to the resumption of the local workers strikes.

TEACHERS

Overall, it was reported that because of MA'AN, students gained new skills in leadership, cooperation, empathy, and self-esteem. The more inclusive and activities-based approaches used by teachers as a result of the training helped to build trust among girls and boys and better include children with physical and learning disabilities who became more active in class discussions and learning activities.

Teachers reported that GR-PBL activities helped students respond to and resolve daily issues and challenges they face. The activities helped to increase a sense of belonging, respect for school rules, and attention and concentration in class. This increased the efficiency of the learning process. With time saved on the actual lessons, teachers had more time to introduce other games and activities. Furthermore, school staff reported a decrease in school dropouts,¹⁶ an enhancement of the overall educational level, and an increase in students' motivation and discipline, with a positive impact on their health and physical safety. Another demonstrable impact of MA'AN was the sense of students' ownership of the classroom. Rearranging and using the space for activities motivated students to take more care in keeping their class tidy and clean. The impacts were also seen beyond the classroom. Many students took these activities home and used them with their siblings. This increased their enthusiasm and positive attitude.

To what extent is the project gender responsive?

Note: This section focuses on understanding the **project's effectiveness** as it relates to gender outcomes. To do that:

- It provides a **deeper analysis** on **indicators** related to gender from both the **logframe sample** and the **treatment/control**.
- It also uses **qualitative data** from FGS, KIIs, and SSIs to describe how **participants view** the **effectiveness** of the project on **indicators related to gender**.

¹⁶ Principals, teachers, and counsellors reported on the issue of dropout in both of Gaza and the West Bank, and they attributed it to early marriage in the case of girls or work in the case of boys. This issue was explored briefly, but it is expected that the quality of education, and addressing psychological challenges and gender equality issues, are likely to increase students' commitment to schooling.

OVERVIEW

Our findings show a more complicated picture about gender dynamics in the targeted communities. With the quantitative analysis, indicators related to gender may be observed at two different levels. First, one may observe positive effects on indicators that have a gender focus. For example, indicator 1320.1 is about the extent to which caregivers hold positive perceptions around gender equality. Respective results have already been reported within the section on the baseline/midline comparison as well as the treatment/control comparison. The current section will just provide a summary in this respect. Second, one may observe differential project effects that vary between boys and girls. For example, one may observe project outcomes for one gender but not for the other. The second part of this section aims at discussing to what extent there are indeed gender-differential effects.

INDICATORS RELATED TO GENDER-RESPONSIVE ATTITUDES AND ENVIRONMENT

Project effects have been analysed both within the baseline/midline comparison as well as within the treatment/control comparison.

Indicators that related to gender attitudes are 1100.1, 1110.2, 1200.3, 1200.4, 1200.5, 1300.2, 1300.4, 1310.1, 1320.1, and 1320.2. They also relate to important aspects of the school environment (e.g., attitudes of school staff and caregivers) and thus provide an indication to what extent the school environment is gender responsive. Several indicators point to MA'AN contributing to school environments that are more gender responsive. A joint analysis of both baseline/midline data as well as treatment/control data suggests that primary teachers are now more likely to report the adequate use of referral systems to address gender equality and GBV (1300.1) as well as to report feeling supported by education supervisors in delivering gender-responsive teaching and learning (1300.4). The baseline/midline comparison also revealed that teachers are more familiar with strategies to include girls in educational activities (1110.2), a finding not supported by the treatment/control analysis.

The school environment becoming more gender-responsive is also echoed in interviews with evaluation participants. It was reported that MA'AN has been effective in activating the referral mechanisms, especially through parents' meetings organised as part of the project. These referral mechanisms are used not only to tackle GBV but are part of the UNRWA child protection work. This includes referring students for medical attention, specialised psychological services, or in cases of learning difficulties (glasses prescription, speech therapy, etc.). Counsellors reported that they are now using these referral mechanisms, connecting with parents, and conducting consultations on how best to support students, especially children with disabilities. In addition, counsellors use existing protection networks with NGOs and others to refer cases when needed

and undertake regular meetings with parents to inform them about the referral mechanisms in place and encourage them to use them when needed.

Finally, as mentioned above, attitudes towards ending GBV (1200.4), gender-responsive learning (1310.1), and views toward gender equality (1320.2) may have deteriorated between baseline and midline. However, counterfactual-based evidence suggests that respective attitudes may have deteriorated even worse within control school communities between both time points. This in turn may suggest that the project might have had a buffering effect on some of the contextual factors affecting school communities. Nonetheless, awareness of issues affecting gender equality may still not be high. In the questionnaires for both secondary students and parents, research participants were asked what they think GBV means. Only about 58 percent of parents and 61 percent of students indicated knowing what GBV means, although they were not asked to define it. Therefore, there was no way to test if those that reported knowing what GBV means do know or vice versa. There is, therefore, a possibility that the actual levels of students and parents who know the meaning of GBV is actually lower. Irrespective of this, at a minimum, around **40 percent of students and their parents do not know about GBV**; this suggests that there is a need for additional awareness-raising campaigns around GBV. A recommendation thus could be to provide additional information material on what GBV constitutes and on symptoms on how to detect victims of GBV. The project's media campaign to be delivered within the final year of the project implementation may be a step into the right direction.

The qualitative data has provided some additional insights into project effects on gender dynamics. The MA'AN training created space for counsellors to discuss areas of professional conduct and to plan and implement gender equality approaches with students. For example, some counsellors used the training and meetings as a starting point for adopting creative methods and integrating theatre activities and drama in their work, which were used to tackle gender equality. Counsellors reported a decline in gender stereotyping, which was achieved through diversifying role plays and ensuring that girls are not represented in a weaker manner, that boys are encouraged to take part in cleaning activities, that girls are reminded of their rights and that they are equal to their boy peers, and so on. Other activities that were useful in challenging gender stereotypes included school parliaments, awareness-raising workshops organised at the schools, and other supporting audio-visual materials.

Another aspect of counsellors' work in MA'AN involves organising regular meetings with parents to raise their awareness on gender equality and to introduce referral mechanisms enacted at the school level. Some parents reported that they had prior knowledge of the topics discussed, while for others the topics were new. Parents reported an overall change in their own attitudes from participating in the meetings and felt that their change in attitudes had a positive and noticeable impact on their children's behaviour. Most of the parents reported that they would continue

attending future meetings if invited, stating that they don't want to miss any opportunities that can potentially benefit their children. Some reported that children themselves are keen about their parents' participation in these meetings. This points to the existing relevance of the activities to the lived reality and the effectiveness of these meetings in handling existing concerns, but it suggests that more work is needed to ensure the sustainability of impact.

OUTCOMES BETWEEN BOYS AND GIRLS

Another lens through which to study the gender responsiveness of the project is to look at differential effects on both girl and boy students. Annexes 4, 5, and 6 provide a gender disaggregation of the statistics associated with the baseline/midline comparison, the treatment/control comparison, and the comparison of both Gaza and West Bank. For Gaza/West Bank, there are differences between both genders across the different indicators. However, these differences do not seem to follow a specific pattern.

Within the baseline/midline comparison, data was disaggregated by territory. At this level, it often suggests that females have attained more positive outcomes compared to males. The only levels where males seem to have surpassed females are in indicators 1100.1 and 1110.2. Male teachers seem to be more likely to demonstrate GR-PBL methodologies (1100.1) and to know strategies to include girls in educational activities (1110.2). This is notable because at baseline, awareness levels between females and males were at par in terms of both indicators. This would suggest that male teachers in particular have improved in these areas.

In most outcome areas, however, higher levels were observed among female participants than among male participants. Compared to boys, girls were more likely to express feelings of social connectedness (1000.1), report an overall improvement in their psychosocial well-being (1000.2), use counselling mechanisms and GBV services (1200.3), and exhibit context-appropriate social/emotional skills (1200.5). In addition, female counsellors were more likely to report the use of play-based learning methodologies in their psychosocial support activities for students (1200.1). At baseline, it seems that respective outcomes were already slightly more favourable for female than for male participants. These trends seem to have continued.

Lastly, the midline data also allows a comparison between co-ed and single-sex schools. While differences in gender outcomes were not commonly found in the evaluation, on certain indicators where there were differences, **both boys and girls at single-sex schools demonstrated higher levels of attainment than boys and girls at co-ed schools**. Boys and girls at single-sex schools were more likely to see improvement in their psychosocial well-being (1000.2), have positive perceptions of their teachers (1100.3), and exhibit higher social/emotional skills

(1100.4/1200.5). In short, it appears that single-sex schools may be associated with higher student well-being than co-ed schools.

In summary, the counterfactual-based impact-assessment element of the evaluation suggests that the project has indeed demonstrated some gender-responsive effects that have particularly emerged in school environments. This also resonates with some gender-differential effects that have been observed among students, especially effects of the project on student well-being that appear to be more pronounced in the case of girls.

To what extent are any outcome changes likely to continue beyond the end of the project?

Note: This section focuses on understanding the **project's sustainability**. To do this, it uses **qualitative data** from FGDs, KIIs, and SSIs to describe how participants view the future sustainability of the project.

The future of the MA'AN project was of high interest to interviewed stakeholders. They highlighted aspects of the project that were most important in the long term, offered suggestions of things that could be improved, and shared ideas for how to sustain and expand the project.

TRAINING

When asked about the training component and the need for future follow-up refresher training, teachers reported that the theoretical part of the training was sufficient in that it allowed for further development of the games that are used in the classroom. However, further training has the potential to improve the way the games are implemented by benefiting from feedback that is based on actual practice. The teachers generally agreed that the training helped them to focus on life skills in the classroom and prepared them to implement the play-based learning activities in that both were easier after building a solid understanding of the objectives of the games. It is believed that these skills will continue to be utilised beyond the period of the MA'AN project, but this sustainability could be enhanced by refresher training.

COMMUNITIES OF PRACTICE (CoPs)

The establishment of CoPs were an integral part of the project. CoP gatherings take place regularly and were reported to most immediately and practically help teachers implement what they have learned in the training. Some created WhatsApp groups to facilitate communication, and these are used to share lessons learned, videos, images, and activities. Further, the CoPs also

work in providing knowledge and building new skills for teachers as well as transferring skills to teachers who have not participated in the MA'AN training. This therefore extends the reach of the project by providing spaces where teachers establish forms of communication with each other and share their teaching experiences. Education supervisors also supported CoPs established as part of this project, as they are seen as important spaces for exchanging ideas and benefiting from feedback. Through CoPs, education supervisors helped in ensuring that plans are being effectively implemented and that counsellors and teachers are supported. The use of these communities even during the pandemic, via online tools that were created midway through the project to respond to the impact of COVID-19, points to the impact they have had on teachers' practice and the effectiveness of this practice in developing staff knowledge and skills around GR-PBL. The CoP is potentially one of the key components that would ensure sustainability of MA'AN as they seem to be useful spaces for extending the learning of MA'AN beyond those who were directly involved.

Toolkits and Curriculum

Despite viewing the training as effective and other resources provided as part of the gender-responsive continuum of teacher training (GR-CoTT) as helpful, teachers still reported facing many challenges in applying what they learned in the training in the classrooms. One of the challenges was utilising the limited space in the class in organising activities that are inclusive for everyone. In addition to that, the seating arrangements limit mobility and raise concerns about the potential injury to students while moving in limited space in activities that need movement. They overcame these challenges by implementing smaller activities or dividing students into groups while ensuring that everyone continued to take part.

Another challenge that was reported was that the toolkit did not include all the resources needed for implementing the games that were part of the training. The pandemic also affected the way activities were implemented. For many teachers, there was not sufficient time to teach students the basics of the curriculum, especially as significant time was lost during lockdowns and school closures. Some teachers therefore continued to prioritise the use of traditional methods as they usually take less time than using GR-PBL activities. This concern is key in planning future training packages in discussion with UNRWA, which should allow for more flexibility and take the existing workload of teachers and the curriculum more into consideration.

Further, while many of the MA'AN activities were implemented as planned and deemed to be culturally- and age-appropriate, some teachers saw the need to introduce adaptations. Examples of adaptations included implementing some classroom-based activities in larger public spaces of the school as these games needed more space than the classrooms allowed for. Some worked to adapt the games when the classrooms were composed of a single-sex group (boys or girls); for

example, creating role plays and theatre activities in which boys take girls' roles (or vice versa). The classroom environment was also physically adjusted into a U shape setting, which was beneficial for communication and increased attention from the students.

These adaptations are important as they demonstrate that teachers had ownership of activities and made them their own and that the games were flexible and allowed for creativity and customisation. However, the extent of the adaptations also signal that the games were not fully appropriate and responsive to time, space, and resource constraints experienced by teachers and the composition of their classes. This will impact the project's future use and relevance. The evaluation team concluded that the games allow for high adaptability, depending on the school, its geographical location, cultural sensitivities, school type (single-sex or co-ed schools) etc. However, some variables (related to UNRWA) are similar across all schools and in both territories (i.e., class is 45 minutes, classroom space is 5X5 metres, mandatory curriculum and minimum learning content is quite standardised). Therefore, it would be helpful if a comprehensive consultation with school staff is initiated by MA'AN on the adaptation they made in response to these limitations, especially as these are constants. For example, some teachers used the public communal space at schools to implement their activities, as a form of adaptation to the limited classroom space, but this raises questions around building the cohort identity, privacy, and other concerns about the learning environment. In this example, students could easily be distracted by what is happening around them, especially those who have attention and concentration impairments. Therefore, it is vital to engage in a discussion on how to creatively adapt the games while ensuring that no harm or learning compromises result. Such a process could ideally be conducted by MA'AN trainers and pedagogy specialists, and the CoPs seem to be an ideal platform to facilitate these discussions.

OTHER RESOURCES

All stakeholders highlighted human and material resource constraints¹⁷ that limited current implementation and identified resource expansion that would help its impact. This was most acutely felt in staff workload. Teachers and counsellors reported that the magnitude of the issues any of the schools deal with exceeds their current capacity, resulting in adding additional burden on teachers and counsellors and affecting the quality of their teaching and counselling practices. They also pointed to the negative impact of only being able to train some teachers, highlighting that because not all teachers were targeted, it creates a difference between students whose teachers took part in the MA'AN project, and those whose teachers did not. Both principals and teachers argued that more field visits by the education supervisors would be beneficial, especially soon after receiving the training to ensure that teachers receive feedback on their

¹⁷ Possibly due to 2018 downsizing that UNRWA had to do from aid cut-offs.

implementation and make any necessary adjustments. However, as reported above, education supervisors generally are only able to undertake one to two visits on average a year to the targeted schools. Parents reported that there is a need to decrease the number of students in the classrooms and provide sufficient spaces for their children to play.

Other highlighted resource constraints were around physical spaces and supplies. Recommendations for addressing these included that some of the tools/sport equipment that were lost or damaged during the previous year's escalation in violence should be replaced. Stakeholders highlighted an overall lack of games (tools and sports equipment) in the schools. One parent recommended supplying schools with rooms equipped with devices, although it was not clear what type of devices, that encourage learning "without using pencils" (i.e., more engaging and using active learning approaches), which can be used by all students at school. Classes can take turns in using these rooms, and participation could be part of the curriculum and therefore mandatory (Rafah, Gaza, mixed FGD). One school staff recommended changing the games because he felt they were repetitive. Education supervisors recommended investing in resources such as smart boards or screens to help implement the activities inside the classroom and better communicate ideas.

PARENT ENGAGEMENT

Parents who participated in the parent meetings run by counsellors largely viewed these as helpful and something they would continue to attend. Some parents recommended increasing the number and frequency of the meetings with them and targeting more parents. Further, in one of the FGDs, parents requested open days for parents and children to be organised, where families can all spend time together away from the school environment.

PROJECT EXPANSION

MA'AN's relevance throughout the project period and future sustainability is perhaps best demonstrated by the calls from all stakeholders for its expansion. These include increasing the number of counsellors in each school, increasing the number of staff trained and involved in future phases of the project, expanding MA'AN to other schools including non-UNRWA schools, and supporting schools with more toolkits and resources. The call to increase the number of counsellors and decrease the students per counsellor ratio is reflected in UNRWA Strategic Plan (see the section on Relevance to population needs). Outcome 2 of Objective 3 (section 172) refers to UNRWA's efforts to move from the existing approach of one counsellor for every 3,000 students to one counsellor for every 800 students. This would represent a significant change that would ensure that students receive quality counselling services and would ensure that

counsellors have more capacity to engage with students and communities and build stronger synergies with teachers.

Both supervisors and teachers recommended expanding the scope to include more teachers—increasing the number of classrooms involved in MA’AN and increasing the number of teachers nominated for training. Some teachers recommended increasing the support in the toolkits and resources available to each teacher in their classroom. Educational supervisors suggested integrating preparatory schoolteachers who were not targeted for training, and expanding the work done with parents to transfer changes made in the classrooms/learning to households. One of the education supervisors suggested holding workshops to highlight the success stories of creative teachers. Another recommended holding meetings between Gaza and the West Bank to transfer experiences—suggesting these could focus on reviewing activities that were included in the training material but not covered in the training itself and tackling challenging trends such as children who are likely to drop out of school or experiencing other forms of marginalisation.

Overall, discussions with the different stakeholders point to many ways that the project’s impact will be long-lasting. To support the **sustainability of the project**, there are several considerations for MA’AN staff.

Recommendations

How could the project be adapted or strengthened to improve performance against the intended outcomes of the project in the final year and in future iterations?

Note: This section details recommendations that could be implemented in the final year of the project and in future iterations to help it improve its relevance, effectiveness, impact, and sustainability.

Based on these findings, the evaluation has identified recommendations to help MA’AN make key improvements in its final year of implementation towards achieving the project’s intended outcomes. The evaluation has also identified a set of recommendations for possible future iterations of the project in Palestine or elsewhere.

Recommendations for Final Year of Project

The consistent finding throughout the evaluation is the high trust and appreciation of MA'AN. This was expressed across all stakeholders. Therefore, as MA'AN concludes its final year of activities, it does so with strong collateral and can leverage this high buy-in to implement a strategic final year focus as well as sustainability measures.

SUSTAINABILITY

BUILD SYSTEMS TO SUPPORT THE LONGEVITY OF THE PROJECT.

To ensure the key achievements of MA'AN last, efforts in the final year should be strongly focused on institutionalising elements of the project and building systems to support them. Focus should shift from what MA'AN staff are doing, to how MA'AN staff are equipping other key stakeholders to carry the work forward.

INITIATE CONVERSATIONS AND PLANNING WITH UNRWA FOR LONG-TERM SUSTAINABILITY MEASURES.

As MA'AN works with UNRWA to integrate GR-PBL guidelines into UNRWA's development department, it should also begin conversations with UNRWA about longer-term sustainability initiatives. The **following three recommendations outline practical applications** of what system-building and sustainability measures **could look like**. While these are measures that would run beyond the end point of the project, initiating conversations, and planning now—in the final year—will be crucial to help put the structures in place to make them possible moving forward.

FORM MULTI-STAKEHOLDER LEADERSHIP TEAMS IN COMMUNITIES AND TERRITORIES.

The strong appreciation reported by all stakeholders, and a desire for the project to expand, indicates a likely willingness of stakeholders to be further involved in it. A leadership team or task force could be formed to take responsibility for the future implementation of MA'AN. It could involve representatives from different groups of stakeholders (e.g., teachers, counsellors, PE teachers, parents, principals, and education supervisors) as well as UNRWA staff and Right to Play in an advisory role and/or could be set up locally to increase community ownership.

INITIATE A TRAIN THE TRAINER (TOT) PROJECT TO EQUIP STAFF WITH THE KNOWLEDGE AND SKILLS TO CARRY MA'AN FORWARD.

It seems that a ToT is an essential approach to be used in the long term as part of a sustainability approach. While its creation or use may not be applicable in the final year of the MA'AN project,

initiating work on implementing a ToT model that involves school staff who have already completed the training and have interest in training other teachers in their schools/community as well as UNRWA staff from various departments (teachers, principals, supervisors, gender officers, and child protection officers) could have several benefits. It can help to eventually expand ownership of the MA'AN philosophy and approach, embed expertise in local schools, expand the frequency and reach of trainings, and ensure a holistic approach to training and mentoring implementation beyond the end term of the project. It can further be adapted by UNRWA's development and training department so that the GR-PBL approach becomes one of the required courses for UNRWA teachers.

WORK WITH UNRWA TO SUPPORT AND DEFINE THE ROLE OF EDUCATION SUPERVISORS IN THE FUTURE OF THE PROJECT.

The role of educational supervisors was demonstrated to have considerable importance in supporting teachers' use of MA'AN and could be critical in its future viability. MA'AN could work with UNRWA to help advise on what their roles could be. They could include scheduling regular times for the training and determining when the best time is for training (while participants recommended summer, this was previously and unsuccessfully attempted), booking CoP times for teachers, ensuring the GR-PBL annex that was developed in year two of the project is integrated into the UNWRA school visit report, and supporting GR-PBL school-wide initiatives.

TOOLS AND MATERIALS

USE THE GR-PBL APPROACH TO CONTINUE TO EQUIP EDUCATIONAL SUPERVISORS AND PRINCIPALS WITH TOOLS TO RESPOND TO THEIR NEEDS AND HELP THEM ASSESS AND PROMOTE THE ADOPTION OF THE PROJECT.

In addition to the importance of education supervisors outlined above, principals spoke positively about MA'AN and were interested in how it could continue. Principals and education supervisors are strategically positioned to help advance the longevity of the project and it would be important to work with these staff—possibly through a few consultation sessions—to identify the tools that would continue to be the most helpful to them and respond to their needs. Such tools should focus on improving what is already available within the UNRWA system or use and improve what has been provided by the MA'AN project. This might include adapting checklists to be used for staff supervision, adding the already-developed GR-PBL annex to the school visit report, using templates for meeting minutes and tasks and ways to document and track progress and challenges related to GR-PBL, discussion guides for staff meetings to discuss key principles of the approach, and using the existing tools in the supervisors' reports, such as monitoring tools to understand progress and implementation. A section on gender-responsive progress could be

added to templates being developed to document challenges and progress being made in this aspect of the project. To stimulate ownership, this could be developed in collaboration with UNRWA's development and gender departments and project teams.

REINFORCE THE DEFINITION OF GR-PBL.

It appears that staff are adopting MA'AN's GR-PBL methodology. However, these are largely based on self-reports and both gender-responsiveness and play-based learning activities could be areas prone to over reporting. When reporting, school staff may be recognising that both are good things and therefore believe they are doing them because they value them conceptually. MA'AN should work to continue to develop very clear examples of what is meant by GR-PLB to ensure that certain practices are not conflated as part of the methodology when they are not. A simple example of this could involve a laminated work tool for teachers to keep at their desk that specifies in different columns what GR-PBL learning is and what it is not.

DEVELOP SIGNAGE THAT COMMUNICATES KEY GR-PBL MESSAGES AND GBV SERVICES.

Another way to help reinforce the definition of GR-PBL would be through signage. While often overlooked, signage and visual cues in a physical space can help reinforce service approaches and key messages. MA'AN staff could work to design a package of materials that could be installed throughout schools—posters that teachers could hang up in classes, flyers or infographics that could be posted in places frequented by students to promote the use of the GBV referral mechanism, or larger signage throughout the schools that emphasise key GR-PBL values and messages. Further, it could be to provide additional information material on what constitutes GBV and what the possible signs are of someone experiencing GBV. The project's media campaign to be delivered within the final year of the project implementation may thus be a step into the right direction.

SUPPORT TEACHERS TO FEEL CONFIDENT IN MAKING KEY ADAPTATIONS TO THE TRAINING MATERIALS.

While the activities provided are extensive and meant to be adapted as most curricula are, many staff in FGDs reported needing to make frequent adaptations to activities from the training. To some it appeared that they did not feel confident in doing this and for others it appeared that the need to make adaptations lessened their commitment to the curriculum as it required more time and effort on their part. It would be important in initial trainings, and then reinforced through CoPs, that teachers should feel empowered to make adaptations. It would be important to identify if there are common themes in the adaptations that staff are making as they might be areas where the toolkit needs updating. Since teachers frequently reported feeling overcapacity, if implementing the curriculum requires significant time and effort from staff to adapt it, this

could lessen their buy-in overtime. Organising workshops with school staff on the challenges they are facing in adopting GR-PBL activities could help identify games and activities that need to be further developed based on the actual experience of GR-PBL over the project intervention period thus far.

GBV

AMPLIFY AND INCREASE AWARENESS OF THE REFERRAL MECHANISM.

The evaluation found signs that teachers and staff are increasing their use of the referral mechanism and view it as a helpful resource. As explained in the findings, the project seems to have a buffering effect around gender attitudes. Within control school communities, gender attitudes seem to have deteriorated faster than in treatment school communities. This means gender attitudes seem to have deteriorated faster in school communities not served by the project compared to school communities served by the project. Despite this buffering effect of MA'AN, gender attitudes seem to have decreased within project school communities, nonetheless. It is hypothesised that some of the decrease in gender attitudes may result from school closures due to COVID-19, escalating violence in both territories, the local workers' strikes, and the fact that the physical school context is a strategic place where norms and attitudes of gender are reinforced—therefore, this is impacted by any closures. The referral mechanism is potentially a particularly strategic way to promote outcomes related to gender in the final year, as it provides a very tangible way to help connect students with needed services in response to GBV or other issues of child protection, which if unaddressed, could adversely impact their quality of education. Promoting the use of the referral mechanism specifically in response to GBV can also serve as a platform to initiate broader conversations about GBV and gender equality, which could help inform positive attitudes on these issues. Thus, amplifying the referral mechanism in the final year could be very important.

INCREASE AND/OR ENSURE ALL SCHOOLS ARE HOLDING PARENTS' MEETINGS.

The parents' meetings run by school counsellors were received positively on several fronts. However, there was also some indication that they weren't happening uniformly across all schools. MA'AN should work to ensure in the final year that all schools are holding regular parent meetings and/or identify if there are ways to expand their frequency. These could be a key forum where the GBV referral mechanism is highlighted, the importance of which is outlined above.

INCREASE ACTIVITIES AND EXAMPLES DESIGNED FOR TEACHERS IN SINGLE-SEX SCHOOLS.

It appeared that the teachers who had to make the most adaptations to the GR-PBL activities were teachers at single-sex schools. Further, in FGDs, teachers and counsellors seemed to have

a general conception that the project was less relevant in single-sex schools. Regardless of the school where staff worked, they saw the project as an important intervention in co-ed schools, but its relevance was less apparent and important to them in single-sex schools. It could be important to focus on the experiences of teachers at single-sex schools and make adaptations to the training and activities to support these schools specifically and to help staff understand its relevance even in single-sex contexts.

Future Iterations of the Project

PROJECT DESIGN

INCREASE ALIGNMENT OF INTERVENTION AND OUTCOMES.

There seemed to be some incongruencies between the indicators the project was attempting to achieve and specific day-to-day intervention. For example, there are many outcomes and indicators related to gender, such as improving resilience of girls to respond to and overcome conflict and violence and GBV (outcome 1200), or improving the response of mothers, fathers, and community members to the gender-based education needs of girls and boys and GBV (outcome 1300). While there have been improvements against these indicators, it would be important in future iterations to ensure better communication about the approach to gender responsiveness at all levels of the programme to increase buy-in and understanding.

FOCUS AND DECIDE ON THE MOST STRATEGIC TARGETING OF INTERVENTIONS.

The MA'AN project engages with several different education staff in the UNRWA system—teachers, counsellors, PE teachers, principals, and education supervisors. Beyond teachers, there was strong evidence of the role of counsellors and education supervisors. However, there were indications that the importance of their role far exceeded their capacity; some counsellors were part-time; education supervisors might only see a teacher one to two times a year. Long-term plans need to account for this, as having staff with such crucial roles to the project, but also such stretched capacity, will impact effectiveness. In relation to this, the role of PE teachers and principals was less clear. It would be important for future iterations to better articulate the unique role they play or make strategic decisions to focus attention elsewhere. For example, a scaled down version of the project could exclude the training of PE teachers.

SUPPORT UNRWA AND IDENTIFY KEY RESOURCE NEEDS.

A remarkable positive aspect of the project was the partnership between UNRWA and MA'AN. However, as UNRWA will be central in maintaining future efforts, Right to Play could focus its

future work on supporting and collaborating with UNRWA, both in an advisory capacity on the GR-PBL methodology and also in helping to identify and secure critical resource needs. For example, the roles of education supervisors and counsellors were viewed as very important by different stakeholders but also that these staff were the most limited in time and numbers. Stakeholders in interviews and FGDs expressed the need for more individuals in these roles. UNWRA has also established a goal of increasing “counselling support, from current norms of one counsellor for every 3,000 children to one counsellor for every 800 children.”¹⁸ This could be an area that Right to Play helps to support UNRWA in addressing. Further, many participants in all groups of stakeholders reported that MA’AN fell short of responding to the increasing demands and needs in the schools and the targeted communities. This reinforces the overall need of the project but that more resourcing and funding is needed for it to be effective.

STUDENT OUTCOMES

MAKE A LONGER INVESTMENT TO UNDERSTAND STUDENT-LEVEL OUTCOMES.

Given that the unit of intervention was teachers, the spill-over impacts to students’ psychosocial outcomes and parents in a short period of time is notable. Any continued support to teachers is likely to contribute to students’ well-being. However, because the project intervention was training and supporting school staff, more work is needed to understand its impact on students: This could include identifying funding to increase the frequency or intensity of services in a small cluster of schools to identify if more GR-PBL activities lead to better student outcomes. This could also include identifying funding for a longitudinal study of students to assess the same outcomes of current students at key intervals down the line.

BUDGET FOR QUALITATIVE PARTICIPATORY RESEARCH IN THE FINAL EVALUATION TO CAPTURE THE PROJECT’S IMPACT MORE HOLISTICALLY IN STUDENTS’ OWN WORDS.

Although the Proteknôn-Birzeit consortium proposed qualitative research with students, the time and budget constraints on the evaluation limited their participation to the survey only. However, since students are the group that the intervention is designed to ultimately impact, it is crucial to incorporate their voices, and the final evaluation should budget sufficient time and resources to do this in ethical and meaningful ways. Using a case study approach would be beneficial in further exploring the qualitative part of the project. This could take the school as a unit of intervention with different stakeholders, including students and parents, participating in qualitative methods.

¹⁸ UNRWA (2023). UNRWA Strategic Plan 2023-2028. p.36.

A major limitation of the quantitative midline surveys was that they relied solely on self-reporting. A crucial part of the project is to promote the use of GR-PBL. Self-reports by teachers and counsellors as to the extent to which they apply these methodologies in class may be a relatively crude measure and most likely subject to overreporting. Standardised classroom observations that assess classroom interactions and climate may be better positioned to gauge the applicability of GR-PBL to project classrooms. Classroom observations can be cost-intensive and the ability to compare observations from one point in time to another can be inadequate if careful observation protocols are not kept. However, they could be a great source of reflection and learning in a final evaluation, especially when combined with quantitative self-reports and qualitative data collection.

CREATE SPECIFIC STRATEGIES FOR SCHOOL TYPE.

The experiences and impact of the project in single-sex and co-ed schools appeared to differ. The evaluation identified instances where students at single-sex schools were performing better than students in co-ed schools. Single-sex schools were also preferred among parents. However, staff at single-sex schools experienced more difficulty and had to make more adaptations to the GR-PBL activities. School type should be taken into consideration during the training sessions, CoP discussions, and monitoring visits to make sure that strategies are appropriate. For example, co-ed schools might need more intensive services with more teachers receiving training and more counsellor support.

STRENGTHEN GENDER COMPONENTS OF THE PROJECT BY LEVERAGING PARENTS' AND STUDENTS' IDEAS AND LEADERSHIP.

From baseline to midline, the survey data found decreasing levels of attitudes towards gender equality. However, in treatment and control comparisons, treatment schools had statistically significant higher levels of parents and caregivers with perceptions of gender equality and of staff who used the referral mechanisms to address gender equality. Interviews demonstrated that stakeholders themselves supported gender equality, but that cultural stereotypes and biases still dominated. Parents specifically demonstrated mixed support for gender equality—both affirming its value while also demonstrating differences in how they see their children's prospective education and employment opportunities. To increase the impact on gender equality, the project could strengthen its gender components by focusing on engaging parents and students as agents of change. This could entail holding focus groups with parents and students to identify their needs and ideas and equipping them either through existing parent committees and school leadership groups or starting new ones to execute the ideas that emerge. This could also include developing a toolkit focused on school-wide campaigns and initiatives that happen throughout the year across the whole school, not just in individual classes. This could

help with changing culture, reinforcing of positive norms, and being a way to celebrate and highlight good practices.

CREATE A CLASS- AND SCHOOL-LEVEL GENDER MONITORING FRAMEWORK.

To localise gender mainstreaming, it could be important to work with UNRWA to identify a few key indicators at both the classroom- and school-level that teachers and principals could use to regularly monitor gender equality. This would help stakeholders to more easily identify what changes look like and keep attention focused on this. Further, as more data is collected on gender equality, it could be important to identify what gender equity measures might be promoted within schools to improve gender equality.

PLAN FOR INCREASED INTEGRATION WITH COMMUNITY-BASED INITIATIVES TO CHANGE ATTITUDES TOWARDS GENDER EQUALITY.

Although significant work has been done with parents, this needs to be complemented by organising workshops in the local communities and not only in schools. The impact of MA'AN could potentially be higher if dominant societal attitudes are challenged. This can best be done in collaboration with civil society community-based organisations. Parents who were engaged in the parents' meetings tended to have an interest in their children's education and well-being, views which are, however, not representative of the whole society. Therefore, it seems key to go beyond this group in future interventions.

Annex 1

OVERVIEW OF EVALUATION QUESTIONS AND LINKED EVALUATION CRITERIA[1]

#	Evaluation Question / Key Sub-questions	Evaluation Criteria	Qualitative Data Sources	Quantitative data Sources
To what extent does the project respond to the needs and priorities of the affected population?				
	Is the project still relevant in the context of The Gaza Strip and West Bank?	Relevance / Appropriateness	Parents: FGD Counsellors/Teachers : SSI Principals: KII Education Supervisors: SSI	N/A
	To what extent is the project aligned to the UNRWA education strategy in Palestine?		Principals: KII Education Supervisors: SSI	N/A
To what extent is the project on track to achieve its intended outcomes?				
	Is the project improving:	Effectiveness	All qualitative tools	Additionality gauged through treatment/control comparison
	i. Children’s psychosocial well-being and social connectedness		Parents: FGD Counsellors/Teachers : SSI	Well-being survey (1–4/5–9)

#	Evaluation Question / Key Sub-questions	Evaluation Criteria	Qualitative Data Sources	Quantitative data Sources
	ii. The emotional and physical learning environment in schools (where applicable)		Counsellors/Teachers : SSI	School assessment form/Child survey (1–4/5–9) ***
	iii. Children's positive perception of their teachers		Parents: FGD Education Supervisors: SSI Counsellors/Teachers : SSI	Child survey (1–4/5–9)
	iv. Children's Social and Emotional Learning (SEL) competencies		Parents: FGD Counsellors/Teachers : SSI	SEL questionnaire
	v. The participation of children in class		Counsellors/Teachers : SSI Education Supervisors: SSI	Child survey (1–4/5–9) ***
	vi. Teachers, PE Teachers, Counsellors' application of skills in GR-PBL methodologies	Effectiveness	Counsellors/Teachers : SSI Principals: KII Education Supervisors: SSI	Counsellor's survey/ Child survey 5- 9 *** (only partly)
	vii. Children's use of school counselling services and GBV services		Counsellors/Teachers : SSI	Child Survey 5–9

#	Evaluation Question / Key Sub-questions	Evaluation Criteria	Qualitative Data Sources	Quantitative data Sources
	viii. Children’s attitude towards ending GBV		Parents: FGD Counsellors/Teachers : SSI Education Supervisors: SSI	Child Survey 5–9
	ix. UNRWA staff and parents use of referral mechanisms, contribution towards enhanced gender equality, GBV elimination and gender-responsive learning		Parents: FGD Counsellors/Teachers : SSI Principals: KII	Teacher survey/ Counsellor’s survey/ Parents’ survey/School assessment form
	x. Capacity of UNRWA staff to ensure gender-responsive learning environment in schools		Principals: KII	N/A
	xi. Parents’ awareness about gender equality, GBV elimination, psychosocial well-being, and gender-responsive learning		Parents: FGD	Parents’ survey
To what extent has the project contributed to the main outcomes?				
	How has the project affected different outcomes for treatment versus comparison groups? Include gender & geographic comparison analysis.	Impact	Parents: FGD Principals: KII Counsellors/Teachers : SSI	Additionality of project will be gauged through treatment/control comparison. N/A
	Have there been any unintended outcomes, either positive or negative?		Parents: FGD Principals: KII	

#	Evaluation Question / Key Sub-questions	Evaluation Criteria	Qualitative Data Sources	Quantitative data Sources
			Counsellors/Teachers : SSI	
To what extent is the project gender-responsive?				
	To what extent did the project manage to reduce gender inequalities in schools/community?	Effectiveness	Parents: FGD Education Supervisors: SSI Counsellors/Teachers : SSI	
	To what extent did the project change attitudes towards GBV and the use of referral mechanisms?		Parents: FGD Counsellors/Teachers : SSI	
To what extent are any outcome changes likely to continue beyond the end of the project?				
	To what extent has the project contributed to sustainable access to psychosocial well-being and GBV services?	Sustainability	Parents: FGD Principals: KII Counsellors/Teachers : SSI	N/A
	What school, community, or institutional level structures and processes will need to be in place for key activities to continue once the project ends?		Principals: KII Counsellors/Teachers : SSI	N/A
	How has the project increased integration of the GR-PBL into UNRWA schools’ plans and budgetary processes?		Principals: KII	N/A

#	Evaluation Question / Key Sub-questions	Evaluation Criteria	Qualitative Data Sources	Quantitative data Sources
	To what extent will the behavioural change, attained due to programme, be sustainable in the future?			
How could the project be adapted or strengthened to improve performance against the intended outcomes?				
	What are the strengths of the GR-PBL approach in contributing to the intended outcomes?	Effectiveness	Principals: KII Counsellors/Teachers : SSI	
	What would strengthen performance in the final year of the project?		Principals: KII Counsellors/Teachers : SSI Education Supervisors: SSI Parents: FGDs	
	In future iterations of the project, how could it be improved to achieve its intended outcomes?			

Annex 2

SUMMARY TABLES OF ALL INDICATORS

Indicator	Tool	Comparison of % attainment (SE)					
		Baseline	Midline	Control	Treatment	Male	Female
1000: improved quality of education children, especially girls, enrolled in UNRWA schools in the West Bank and Gaza							
1000.1: %/total girls and boys with feelings of social connectedness (m/f)	Well-being survey	31%	50% (0.01)	39% (0.03)	47% (0.04)	44% (0.02)	55% (0.02)
1000.2: %/total girls and boys that report an overall improvement in their psychosocial well-being (m/f)	Well-being survey	14%	28% (0.01)	14% (0.02)	31% (0.03)	21% (0.02)	33% (0.02)
1000.3 %/total of classrooms that meet criteria for emotional safety/emotional well-being	School assessment form	32%	58% (0.02)	53% (0.03)	59% (0.04)	60% (0.02)	57% (0.02)
1100 increased integration of GR-PBL by primary teachers into teaching practice							
1100.1 %/total of teachers demonstrating GR-PBL methodologies (m/f)	Teachers' survey	37%	49% (0.03)	38% (0.06)	52% (0.04)	N/A	N/A
1100.2 %/total girls and boys who feel empowered to participate in classroom activities (m/f)	Child survey (1–4/5–9)	81%	87% (0.01)	86% (0.02)	92% (0.02)	86% (0.01)	88% (0.01)

Indicator	Tool	Comparison of % attainment (SE)					
		Baseline	Midline	Control	Treatment	Male	Female
1100.3 %/total of girls and boys who have a “positive” perception of their teacher (m/f)	Child survey (1-4 /5–9)	47%	61% (0.02)	44% (0.03)	67% (0.03)	60% (0.01)	61% (0.01)
1100.4 %/total of girls and boys that exhibit context-appropriate social/emotional skills (m/f)	Well-being survey	29%	26% (0.01)	15% (0.02)	33% (0.03)	25% (0.02)	28% (0.02)
1200 Improved resilience of girls to respond to and overcome conflict and violence, including GBV (intermediate outcome)							
1200.1 %/total of counsellors demonstrating play-based learning methodologies in their psychosocial support activities for students (m/f)	Counsellor’s survey	22%	94% (0.03)			86% (0.07)	100% (0.00)
1200.2 %/total of PE teachers demonstrating play-based learning methodologies in their psychosocial support activities for students	PE teacher survey	16%	55% (0.05)			60% (0.07)	50% (0.08)
1200.3 % total of girls and boys who use school counselling mechanisms and GBV services (m/f)	Child Survey 5–9	57%	25% (0.01)			9% (0.03)	32% (0.03)
1200.4 % total of girls and boys reporting positive attitudes towards ending GBV (m/f)	Child Survey 5– 9	77%	58% (0.03)			55% (0.05)	60% (0.04)
1200.5 %/total of girls and boys that exhibit context-appropriate social/emotional skills (m/f)	Well-being survey	20%	38% (0.02)	27% (0.03)	48% (0.04)	32% (0.02)	43% (0.02)

Indicator	Tool	Comparison of % attainment (SE)					
		Baseline	Midline	Control	Treatment	Male	Female
1300 Improved response of school officials, UNRWA staff, mothers, fathers, and community members to the gender-based education needs of girls and boys, and GBV							
1300.1 %/total of UNRWA staff who efficiently use existing referral mechanisms to address gender equality and monitor the safety of girls and report on GBV (m/f) Quantitative, and triangulated through KIIs and FGDs	Teacher survey Counsellor’s survey Parent survey School assessment form	20%	86% (0.02)	81% (0.05)	95% (0.02)	94% (0.04)	85% (0.02)
1300.2 %/total mothers, fathers, caregivers, and community members responding to gender-based education needs of girls and boys (m/f)	Parent survey	25%	19% (0.02)	23% (0.05)	13% (0.02)	20% (0.04)	18% (0.03)
1300.3 %/total of girls and boys who receive focused psychosocial and GBV services (m/f)	School assessment form	N/A	average of 389 per school	average of 315 per school	average of 429 per school	N/A	N/A
1300.4 %/total of teachers who receive support from education supervisors to deliver gender-responsive learning	Teacher survey	31%	70% (0.03)	57% (0.06)	83% (0.03)	71% (0.08)	70% (0.03)

Indicator	Tool	Comparison of % attainment (SE)					
		Baseline	Midline	Control	Treatment	Male	Female
1110 Improved skills of teachers to implement GR-PBL into teaching practice (m/f)							
1110.1 %/total of trained teachers scoring satisfactorily on post-GR-COTT training tests (m/f)	Pre/post data	N/A	N/A	N/A	N/A	N/A	N/A
1110.2 %/total of teachers that know the strategies to include girls in educational activities (m/f)	Teacher survey	N/A	N/A	N/A	N/A	N/A	N/A
1210 Improved skills of school counsellors to provide gender-responsive, play-based psychosocial support for girls and boys (M/F)							
1210.1 %/total of trained counsellors scoring satisfactorily on post-GR-COTT training tests (m/f)	Pre/post data	N/A	N/A	N/A	N/A	N/A	N/A
1210.2 %/total of counsellors that feel prepared to implement psychosocial support activities to girls and boys (disaggregated by sex)	Pre/post data	N/A	N/A	N/A	N/A	N/A	N/A
1220 Improved skills of PE teachers to lead gender-responsive sport and play activities for boys and girls (M/F)							
1220.1 %/total of PE teachers who feel prepared to deliver interventions to change attitudes on GBV for boys and girls (m/f)	Pre/post data	N/A	N/A	N/A	N/A	N/A	N/A
1220.2 %/ Total of PE teachers scoring satisfactorily on post-GR-COTT training tests (m/f)	Pre/post data	N/A	N/A	N/A	N/A	N/A	N/A

Indicator	Tool	Comparison of % attainment (SE)					
		Baseline	Midline	Control	Treatment	Male	Female
1310 Improved capacity of school principals, education supervisors, and UNRWA staff to ensure gender-responsive learning environment in schools (immediate outcomes)							
1310.1 %/total of education supervisors and school principals with positive attitudes towards ensuring gender-responsive learning environments in schools (m/f)	School Assessment Form	61%	34% (0.06)	17% (0.15)	29% (0.12)	N/A	N/A
1310.2 % total of education supervisors with knowledge of gender-responsive learning environments in schools (m/f)	Pre/post data	N/A	N/A	N/A	N/A	N/A	N/A
1320 Increased awareness of mothers, fathers, and caregivers on how to support gender equality and psychosocial well-being of girls and boys, and respond to GBV (immediate outcomes)							
1320.1 % total of mothers, fathers, and caregivers with knowledge of where psychosocial support services are provided and where and how to formally report GBV (m/f)	Parent survey	42%	54% (0.03)	39% (0.05)	49% (0.04)	55% (0.05)	53% (0.03)
1320.2 % total of mothers, fathers, and caregivers with positive perception of gender equality (m/f)	Parent survey	61%	40% (0.03)	23% (0.05)	62% (0.04)	42% (0.05)	39% (0.03)

Annex 3

Indicator	Tool	Questions (midline)
1000 Improved quality of education children, especially girls, enrolled in UNRWA schools in the West Bank and Gaza Strip		
1000.1: %/total girls and boys with feelings of social connectedness (m/f)	Student survey (1–4)	SEF4/SEF5/SEF6/SAF1/SAF2/SAF3/RFS1/RFS2/RFS3/WELL1/WELL2/WELL3/WELL4
	Student survey (5–9)	SLEC2/SLEC3/SLE8/SLEC9/SLEC10/SLC11/SLEC14/SLEC15/SLEC16/SLEC19/SLEC21
1000.2: %/total girls and boys that report an overall improvement in their psychosocial well-being (m/f)	Student survey (1–4)	SEF3/SEF4/SEF5/SEF6/SAF1/SAF2/SAF3/RFS1/RFS2/RFS3/WELL1/WELL2/WELL3/WELL4
	Student survey (5–9)	SLEC1 through SLEC21
1000.3 %/total of classrooms that meet criteria for emotional safety/emotional well-being	Student survey (1–4)	PS5/PS7/PS9/PS10/PS11/SEF3/SAF1/SUP2
1100 Increased integration of GR-PBL by primary teachers into teaching practice		
1100.1 %/total of teachers demonstrating GR-PBL methodologies (m/f)	Teacher survey	TA7/TA8/TA12/TA13/TA14/TA15/TA16/TA17/TA18
1100.2 %/total girls and boys who feel empowered to participate in classroom activities (m/f)	Student survey (1–4)	PS4/PS5/PS6
1100.3 %/total of girls and boys who have a “positive” perception of their teacher (m/f)	Student surveys 1–4/5–9	PS7/PS8/PS9/PS10/PS12
1100.4 %/total of girls and boys that exhibit context-appropriate social/emotional skills (m/f)	Student surveys 1–4	COP1/GOAL1/GOAL2/GOAL3/PERS1/PERS2/RFS1/RFS2/RFS3/P12
1200 Improved resilience of girls to respond to and overcome conflict and violence, including GBV (intermediate outcome)		

Indicator	Tool	Questions (midline)
1200.1 %/total of counsellors demonstrating play-based learning methodologies in their psychosocial support activities for students (m/f)	Counsellor's survey	CO10/C011
1200.2 %/total of PE teachers demonstrating play-based learning methodologies in their psychosocial support activities for students	PE Teacher survey	PE7/PE8/PE9/PE10/PE11
1200.3 % total of girls and boys who use school counselling mechanisms and GBV services (m/f)	Student surveys 5–9	SS7/SS9/SS10
1200.4 total of girls and boys reporting positive attitudes towards ending GBV (m/f)		SS13/SS14/SS15/SS16/SS17/SS18
1200.5 %/total of girls and boys that exhibit context-appropriate social/emotional skills (m/f)	Student surveys 1–4	COP1/GOAL1/GOAL2/GOAL3/ PERS1/PERS2/RFS1/RFS2/RFS3
1300 Improved response of school officials, UNRWA staff, mothers, fathers, and community members to the gender-based education needs of girls and boys, and GBV		
1300.1 %/total of UNRWA staff who efficiently use existing referral mechanisms to address gender equality and monitor the safety of girls and report on GBV (m/f)	Teacher survey	TA19
1300.2 %/total mothers, fathers, caregivers, and community members responding to gender-based education needs of girls and boys (m/f)	Parent survey	PR11/PR12/PR13/PR15
1300.3 %/total of girls and boys who receive focused psychosocial and GBV services (m/f)	School assessment form	SA18/SA19/SA24

Indicator	Tool	Questions (midline)
1300.4 % total of teachers who receive support from education supervisors to deliver gender-responsive learning	Teacher survey	TS25.1/TA26/TA27/TA28/TA29/TA30
1110 Improved skills of teachers to implement GR-PBL into teaching practice (m/f)		
1110.2 %/total of teachers that know the strategies to include girls in educational activities (m/f)	Teacher survey	TA12/TA13/TA14/TA15 /TA16/TA17/TA18
1310 Improved capacity of school principals, education supervisors, and UNRWA staff to ensure gender-responsive learning environment in schools (immediate outcomes)		
1310.1 %/total of education supervisors and school principals with positive attitudes towards ensuring gender-responsive learning environments in schools (m/f)	School Assessment Form	SA28/SA29/SA30/SA31/SA32/SA33/SA35/SA36/SA37
1320 Increased awareness of mothers, fathers, and caregivers on how to support gender equality and psychosocial well-being of girls and boys, and respond to GBV (immediate outcomes)		
1320.1 % total of mothers, fathers, and caregivers with knowledge of where psychosocial support services are provided and where and how to formally report GBV (m/f)	Parent survey	PR3/PR7
1320.2 % total of mothers, fathers, and caregivers with positive perception of gender equality (m/f)	Parent survey	PR23/PR25/PR26/PR27/PR28/PR29/PR30/PR31/PR32/PR33/PR34

Annex 4

GENDER-DISAGGREGATED COMPARISON BETWEEN BASELINE AND MIDLINE

Time	Sub-group	Proportion (in %)
1000.1) Girls and boys with feelings of social connectedness		
Baseline	All	31.00
Baseline	Female	36.00
Baseline	Male	26.00
Midline	All	50.41
Midline	Female	55.09
Midline	Male	44.03
1000.2) Girls and boys that report an overall improvement in their psychosocial well-being		
Baseline	All	14.00
Baseline	Female	17.00
Baseline	Male	9.00
Midline	All	28.06
Midline	Female	33.14
Midline	Male	21.14
1000.3) Classrooms that meet criteria for emotional safety/emotional well-being		
Baseline	All	N/A
Baseline	Female	N/A
Baseline	Male	58.27
Midline	All	56.83
Midline	Female	60.00
Midline	Male	37.00
1100.1) Teachers that demonstrate GR-PBL methodologies		
Baseline	All	37.00
Baseline	Female	39.00
Baseline	Male	32.00
Midline	All	61.01

Time	Sub-group	Proportion (in %)
Midline	Female	58.68
Midline	Male	77.14
1100.2) Girls and boys who feel empowered to participate in classroom activities		
Baseline	All	81.00
Baseline	Female	82.00
Baseline	Male	81.00
Midline	All	87.14
Midline	Female	87.92
Midline	Male	86.19
1100.3) Girls and boys who have a “positive” perception of their teacher		
Baseline	All	47.00
Baseline	Female	52.00
Baseline	Male	41.00
Midline	All	60.86
Midline	Female	61.19
Midline	Male	60.48
1100.4) Girls and boys that exhibit context-appropriate social/emotional skills		
Baseline	All	29.00
Baseline	Female	32.00
Baseline	Male	26.00
Midline	All	26.49
Midline	Female	27.52
Midline	Male	25.24
1110.2) Teachers that know the strategies to include girls in educational activities		
Baseline	All	20.00
Baseline	Female	19.00
Baseline	Male	21.00
Midline	All	62.45
Midline	Female	59.92
Midline	Male	80.00

Time	Sub-group	Proportion (in %)
1200.1) Counsellors that demonstrate play-based learning methodologies in their psychosocial support activities for students		
Baseline	All	22.00
Baseline	Female	26.00
Baseline	Male	18.00
Midline	All	94.03
Midline	Female	100.00
Midline	Male	85.71
1200.2) PE teachers that demonstrate play-based learning methodologies in their psychosocial support activities for students		
Baseline	All	16.00
Baseline	Female	21.00
Baseline	Male	8.00
Midline	All	55.42
Midline	Female	50.00
Midline	Male	60.47
1200.3) Girls and boys that use school counselling mechanisms and GBV services		
Baseline	All	57.00
Baseline	Female	54.00
Baseline	Male	61.00
Midline	All	24.73
Midline	Female	32.29
Midline	Male	8.79
1200.4) Girls and boys that report positive attitudes towards ending GBV		
Baseline	All	77.00
Baseline	Female	77.00
Baseline	Male	77.00
Midline	All	58.30
Midline	Female	59.90
Midline	Male	54.95

Time	Sub-group	Proportion (in %)
1200.5) Girls and boys that exhibit context-appropriate social/emotional skills		
Baseline	All	20.00
Baseline	Female	22.00
Baseline	Male	17.00
Midline	All	38.05
Midline	Female	42.77
Midline	Male	32.38
1300.1) UNRWA staff who efficiently use referral mechanisms to address gender equality and GBV		
Baseline	All	20.00
Baseline	Female	19.00
Baseline	Male	21.00
Midline	All	85.61
Midline	Female	84.71
Midline	Male	94.12
1300.2) Caregivers and community members that respond to gender-based education needs of girls and boys		
Baseline	All	25.00
Baseline	Female	25.00
Baseline	Male	26.00
Midline	All	19.08
Midline	Female	18.47
Midline	Male	19.80
1300.4) Teachers who receive support from education supervisors to deliver gender-responsive learning		
Baseline	All	31.00
Baseline	Female	30.00
Baseline	Male	32.00
Midline	All	70.04
Midline	Female	69.83
Midline	Male	71.43

Time	Sub-group	Proportion (in %)
1310.1) Education supervisors/school principals with positive attitudes towards gender-responsive learning		
Baseline	All	61.00
Midline	All	34.29
1320.1) Caregivers with knowledge of where psychosocial support services are and where and how to formally report GBV		
Baseline	All	42.00
Baseline	Female	65.00
Baseline	Male	52.00
Midline	All	53.85
Midline	Female	53.15
Midline	Male	55.45
1320.2) Mothers, fathers, and caregivers with positive perception of gender equality		
Baseline	All	61.00
Baseline	Female	64.00
Baseline	Male	56.00
Midline	All	40.00
Midline	Female	39.19
Midline	Male	41.58

Annex 5:

GENDER-DISAGGREGATED COMPARISON BETWEEN TREATMENT AND CONTROL

Group	Subset	Proportion (in %)
1000.2) Girls and boys that report an overall improvement in their psychosocial well-being		
Control	All	14.29
Control	Female	20.98
Control	Male	5.50
Treatment	All	31.49
Treatment	Female	40.23
Treatment	Male	23.40
1100.3) Girls and boys who have a “positive” perception of their teacher		
Control	All	43.65
Control	Female	46.15
Control	Male	40.37
Treatment	All	66.85
Treatment	Female	60.92
Treatment	Male	72.34
1100.4) Girls and boys that exhibit context-appropriate social/emotional skills		
Control	All	15.48
Control	Female	19.58
Control	Male	10.09
Treatment	All	32.60
Treatment	Female	29.89
Treatment	Male	35.11
1200.5) Girls and boys that exhibit context-appropriate social/emotional skills		
Control	All	27.38
Control	Female	34.97
Control	Male	17.43

Group	Subset	Proportion (in %)
Treatment	All	47.51
Treatment	Female	56.32
Treatment	Male	39.36
1300.1) UNRWA staff who efficiently use referral mechanisms to address gender equality and GBV		
Control	All	80.56
Control	Female	80.56
Control	Male	N/A
Treatment	All	95.38
Treatment	Female	96.08
Treatment	Male	100.00
1300.4) Teachers who receive support from education supervisors to deliver gender-responsive learning		
Control	All	56.94
Control	Female	56.94
Control	Male	N/A
Treatment	All	83.08
Treatment	Female	86.27
Treatment	Male	71.43
1320.2) Mothers, fathers, and caregivers with positive perception of gender equality		
Control	All	22.62
Control	Female	23.21
Control	Male	21.43
Treatment	All	62.30
Treatment	Female	58.97
Treatment	Male	68.18

Annex 6

GENDER-DISAGGREGATED COMPARISON BETWEEN TERRITORIES

Territory	Subset	Proportion (in %)
1000.1) Girls and boys with feelings of social connectedness		
Gaza	All	48.25
Gaza	Female	52.67
Gaza	Male	43.54
West Bank	All	58.11
West Bank	Female	60.66
West Bank	Male	48.15
1000.2) Girls and boys that report an overall improvement in their psychosocial well-being		
Gaza	All	22.80
Gaza	Female	26.95
Gaza	Male	18.38
West Bank	All	46.79
West Bank	Female	47.39
West Bank	Male	44.44
1000.3) Classrooms that meet criteria for emotional safety/emotional well-being		
Gaza	All	59.23
Gaza	Female	58.71
Gaza	Male	59.79
West Bank	All	53.33
West Bank	Female	49.51
West Bank	Male	61.70
1100.1) Teachers that demonstrate GR-PBL methodologies		
Gaza	All	61.80
Gaza	Female	59.90
Gaza	Male	76.92

West Bank	All	56.82
West Bank	Female	51.43
West Bank	Male	77.78
1100.2) Girls and boys who feel empowered to participate in classroom activities		
Gaza	All	87.87
Gaza	Female	88.06
Gaza	Male	87.67
West Bank	All	83.33
West Bank	Female	87.38
West Bank	Male	74.47
1100.3) Girls and boys who have a “positive” perception of their teacher		
Gaza	All	60.26
Gaza	Female	60.20
Gaza	Male	60.32
West Bank	All	64.00
West Bank	Female	65.05
West Bank	Male	61.70
1100.4) Girls and boys that exhibit context-appropriate social/emotional skills		
Gaza	All	22.32
Gaza	Female	22.89
Gaza	Male	21.72
West Bank	All	48.00
West Bank	Female	45.63
West Bank	Male	53.19
1110.2) Teachers that know the strategies to include girls in educational activities		
Gaza	All	62.66
Gaza	Female	60.87
Gaza	Male	76.92
West Bank	All	61.36
West Bank	Female	54.29
West Bank	Male	88.89

1200.1) Counsellors that demonstrate play-based learning methodologies in their psychosocial support activities for students		
Gaza	All	97.50
Gaza	Female	100.00
Gaza	Male	94.44
West Bank	All	88.89
West Bank	Female	100.00
West Bank	Male	70.00
1200.2) PE teachers that demonstrate play-based learning methodologies in their psychosocial support activities for students		
Gaza	All	51.52
Gaza	Female	46.88
Gaza	Male	55.88
West Bank	All	70.59
West Bank	Female	62.50
West Bank	Male	77.78
1200.3) Girls and boys that use school counselling mechanisms and GBV services		
Gaza	All	10.71
Gaza	Female	15.48
Gaza	Male	5.95
West Bank	All	45.22
West Bank	Female	45.37
West Bank	Male	42.86
1200.4) Girls and boys that report positive attitudes towards ending GBV		
Gaza	All	55.95
Gaza	Female	53.57
Gaza	Male	58.33
West Bank	All	61.74
West Bank	Female	64.81
West Bank	Male	14.29

1200.5) Girls and boys that exhibit context-appropriate social/emotional skills		
Gaza	All	31.87
Gaza	Female	37.06
Gaza	Male	26.27
West Bank	All	70.00
West Bank	Female	65.05
West Bank	Male	80.85
1300.1) UNRWA staff who efficiently use referral mechanisms to address gender equality and GBV		
Gaza	All	86.70
Gaza	Female	85.99
Gaza	Male	92.31
West Bank	All	80.00
West Bank	Female	77.14
West Bank	Male	100.00
1300.2) Caregivers and community members that respond to gender-based education needs of girls and boys		
Gaza	All	20.00
Gaza	Female	19.19
Gaza	Male	21.05
West Bank	All	10.00
West Bank	Female	12.50
West Bank	Male	0.00
1300.4) Teachers who receive support from education supervisors to deliver gender-responsive learning		
Gaza	All	71.67
Gaza	Female	72.46
Gaza	Male	65.38
West Bank	All	61.36
West Bank	Female	54.29
West Bank	Male	88.89
1310.1) Education supervisors/school principals with positive attitudes towards gender-responsive learning		

Gaza	All	36.17
West Bank	All	30.43
1320.1) Caregivers with knowledge of where psychosocial support services are and where and how to formally report GBV		
Gaza	All	52.54
Gaza	Female	51.52
Gaza	Male	54.74
West Bank	All	66.67
West Bank	Female	66.67
West Bank	Male	66.67
1320.2) Mothers, fathers, and caregivers with positive perception of gender equality		
Gaza	All	37.63
Gaza	Female	36.36
Gaza	Male	40.00
West Bank	All	63.33
West Bank	Female	62.50
West Bank	Male	66.67