



International Assistance to Afghanistan, 2008–2018

Part 5: Summary Report of Selected Evaluation Reports by Multilateral Organizations and NGO, 2008–2018

March 2020 | Reem Saraya and Christoph Zürcher



International Assistance to Afghanistan, 2008 – 2018

Part 5: Summary Report of Selected Evaluation Reports by Multilateral Organizations and NGO, 2008 – 2018

March 2020 | Reem Saraya and Christoph Zürcher

On behalf of

Federal Ministry for Economic Cooperation and Development (BMZ), Germany
Division for Afghanistan and Pakistan



Federal Ministry
for Economic Cooperation
and Development

Contents

1	Executive Summary	6
2	About this Report	8
2.1	Criteria for Including Evaluations	8
2.2	Methodologies Used by Included Evaluations	8
2.3	Evidence Base	8
3	Governance Programs	9
3.1	ARTF	9
3.1.1	Included Reports	9
3.1.2	Findings	9
3.2	Non-ARTF Governance Programs	11
3.2.1	Included Reports	11
3.2.2	Findings	11
3.3	Sub-National Governance	12
3.3.1	Included Reports	12
3.3.2	Findings	12
3.4	Rule of Law	13
3.4.1	Included Reports	13
3.4.2	Findings	13
3.5	Democracy Promotion	14
3.5.1	Included Reports	14
3.5.2	Findings	14
3.6	Summary of Findings: Governance Programs	17
4	Gender Programs	17
4.1	Included Reports	17
4.2	Findings	22
4.3	Summary of Findings: Gender	22
5	Rural Development Programs	22
5.1	Included Reports	22
5.2	Findings	23
5.3	Summary of Findings: Rural Development	25
6	Education Programs	26
6.1	Included Reports	26
6.2	Findings	26
6.3	Summary of Findings: Education	27

7	Health Programs	28
7.1	Included Reports.....	28
7.2	Findings.....	28
7.3	Summary Findings: Health.....	29
8	Country Programs of Multilateral Donors	29
8.1	Included Reports.....	29
8.2	Findings.....	29
9	Conclusions	31
10	References	33
	Imprint	35

Abbreviations and Acronyms

ADB	Asian Development Bank
ARAP	Afghanistan Rural Access Program
ARTF	Afghanistan Reconstruction Trust Fund
ASGP	Afghanistan Sub-National Governance Project
BMZ	Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung/ The Federal Ministry for Economic Cooperation and Development, Germany
CERP	Commanders Emergency Response Program
COIN	Counterinsurgency
DDA	District Development Assemblies
EQUIP	Education Quality Improvement Program
EVAW	Eliminating violence against women
IDLG	Independent Directorate of Local Governance
IEC	Independent Electoral Commission of Afghanistan
IRDP	Irrigation Restoration and Development Project
LOTFA	Law and Order Trust Fund
MBAW	Making Budgets and Aid Work
NDPG	National State Governance Project
NGO	Non-Governmental Organization
NIBP	National Institution Building Project
NSP	National Solidarity Program
SIGAR	Special Inspector General for Afghanistan Reconstruction
UNDP	United Nations Development Program
TA	Technical Assistance

1 Executive Summary

This report summarizes the findings of 36 evaluation reports (excluding impact evaluations, which we discuss in the Impact Evaluations Report¹) of development projects implemented by multilateral donors and NGOs in Afghanistan. The reports were published in English between 2008 and 2018. The main findings for each sector follow.

Governance

The Afghan Reconstruction Trust Fund (ARTF) provided donors with an efficient vehicle for allocating aid and served as a forum for donor coordination. ARTF was also an important vehicle for financing the Afghan administration. However, ARTF has had little impact on better governance or economic growth. Most importantly, it has not contributed to improving capacity in the Afghan Government.

Capacity development programs for the Afghan administration largely failed. This is especially the case for programs that aimed to build capacity for managing relations between the center and the provinces on the basis of meaningful decentralization. In addition, none of the reviewed projects led to increased capacity for subnational administration.

The two projects reviewed in the area of the rule of law failed because they were overly ambitious, were not based on the political-economic realities on the ground and were based on an ideological and unrealistic theory of change.

Development aid was able to provide the capacity needed for presidential elections, but projects aimed at democratic awareness-raising failed.

Gender

The effectiveness of gender programming was very low. Modest projects embedded in traditional structures helped to increase access to health and education; others marginally improved livelihoods for women. By contrast, larger, more ambitious projects aimed directly at changing gender norms and relations failed. To some extent, that failure is attributable to weak leadership and weak capacity in the lead agency. To a larger extent, the projects failed because they were overly ambitious and ignorant of the context, overestimated the capacity and political will of the partner, and underestimated the traditional cultural fabric of Afghanistan.

¹ Christoph Zürcher, Andrew Coon, Marissa de la Torre Ugarte, Patrick Labelle, Binxi Li, Razan Masad, Hassina Popal, Reem Saraya Maryam Shah, Michael Swenson, Ella Sylvester, Anna Vanderkooy, and Mengrou Wang (2019), Systematic review of impact evaluations of development aid in Afghanistan, 2008-2018. BMZ/University of Ottawa Graduate School of Public and International Affairs.

Rural Development

Many rural development projects, often implemented through newly created community-level organizations, were successful at providing small infrastructure, such as roads, irrigation, and access to energy. Better access to services and basic infrastructure contributed to improved livelihoods. However, there was no reduction in overall poverty rates, and much of the infrastructure is not sustainable. There is some evidence that smaller projects implemented by NGOs with more community buy-in worked better.

Education

Well-targeted and well-managed interventions were able to improve the quality of primary teaching, provide primary education to children in remote regions, increase access to education for girls, and increase support for girls' education. The need for infrastructure and qualified teachers remains very high. Despite massive investment by donors, the state of the physical infrastructure and the quality of education still need improvement, and many gains are not sustainable.

Health

The three evaluations on midwifery training suggest that investing in midwifery is an efficient, effective, and sustainable way to improve health outcomes in child and maternal health.

2 About this Report

This report summarizes the findings of 36 evaluation reports of development projects in Afghanistan published in English between 2008 and 2018. The evaluation reports summarized here are formative evaluations, mid-term evaluations, performance audits, forward-looking evaluations and final evaluations. They look at projects, programs or multiple programs. In short, they cover the whole bandwidth of evaluation reports with the exception of strict impact evaluations. Note that we summarize impact evaluations in another report in this series, *Systematic review of impact evaluations of development aid in Afghanistan, 2008-2018* (Zürcher et al., 2019). We suggest reading these two reports in conjunction.

The reports included here focus on important aid sectors: governance, stabilization, rule of law, democracy promotion, rural development, gender, education and health. We also include a section on the multi-sector country-level evaluations of multilateral donors.

2.1 Criteria for Including Evaluations

As mentioned, we discuss impact evaluations in another report. The present report summarizes findings from all other evaluations that did not meet the very demanding inclusion criteria for impact evaluations but nevertheless, in our view, contain important lessons. While these evaluations cannot assess impacts, they nevertheless shed light on relevance, effectiveness, efficiency, outputs and, occasionally, outcomes.

2.2 Methodologies Used by Included Evaluations

The included evaluations are typically based on desk review, including the analysis of project documentation and other relevant secondary literature. Often, stakeholder interviews or focus groups are also used. Sometimes, direct observations during field visits to selected sites are employed. Baseline and/or endline surveys, or other quantifiable data, are not usually used.

2.3 Evidence Base

In total, we included 37 studies. Of these, eleven were on governance (including six on ARTF and three on sub-national governance), three on promotion of democracy, two on the rule of law, eight on gender, five on rural development, three on education and three on health; there were also two evaluations of World Bank and UNDP country programs. To some extent, this is a convenience sample: we included any studies that we thought were of acceptable quality and contained valuable lessons. However, the findings are very much in line with those of our systematic review of impact evaluations and of our bilateral donor evaluation review. This increases our confidence in the validity of the findings emerging from this report.

3 Governance Programs

3.1 ARTF

3.1.1 Included Reports

Management Systems International (MSI). (2017a). Year 1 annual report (November 2015–October 2016). Afghanistan Reconstruction Trust Fund.

Management Systems International (MSI). (2017b). Year II annual report (November 2016–October 2017). Afghanistan Reconstruction Trust Fund.

Scanteam. (2008). Afghanistan Reconstruction Trust Fund: External evaluation. Final report. World Bank.

Scanteam. (2012). ARTF at a crossroads: history and the future. Final report. World Bank.

Scanteam. (2017). Taking charge: Government ownership in a complex context. External review. Afghanistan Reconstruction Trust Fund. World Bank

Sida Decentralized Evaluation. (2015). Review of Afghanistan Reconstruction Trust Fund, ARTF, internal and external studies and evaluations in Afghanistan - final report.

3.1.2 Findings

Management Systems International (MSI) was contracted as an external supervisory agent to conduct 1,900 monitoring/verification missions annually for ARTF sub-project sites of infrastructure programs across all 34 provinces in Afghanistan. Approaches used included participatory monitoring by citizen monitors and remote monitoring. The reports are mainly concerned with verification of outputs, not necessarily effectiveness, efficiency and impact. The reports look at small infrastructure built by various programs. For example, the Afghanistan Rural Access Project (ARAP) builds small infrastructure, mainly roads; the Irrigation Restoration and Development project (IRD) is active in the irrigation sector; and the National Solidarity Programme (NSP) provides small infrastructure at the community level. Most of these outputs were rated as average or good (MSI, 2017a; 2017b).

Projects in EQUIP II (infrastructure relating to education, mainly school buildings) received an overall average grade of 3.0 (satisfactory). The report notes a visible lack of operational and management plans and funding mechanisms at many school sites.

These findings suggest that less complex infrastructure fares better than more complex projects and that the Afghan Government does not yet have enough capacity to plan, build, and maintain infrastructure such as school buildings.

Three more ARTF evaluations come from Scanteam (2008; 2012; 2017). The 2008 Scanteam report looked at the outputs of the NSP. At that time, the NSP had completed around 18,000 small infrastructure projects in water/sanitation (24.3%), transport (22.3%), irrigation (16%), power (15.2%), education (13.4%), and livelihoods (6.6%). Scanteam (2008) noted the lack of either a solid baseline or an agreed set of indicators to monitor, particularly regarding governance. As a result, the extent and sustainability of results (whether poverty reduction, gender distribution impact, strengthened local governance, etc.) could not be documented.

The 2012 Scanteam report (Scanteam 2012) noted that ARTF was an efficient and effective channel for on-budget donor funding. Funding was heavily weighted towards NSP/rural development. Good outputs were achieved in the NSP, health and education, but there were still no outcomes being tracked. Results in capacity development also remained weak.

In their 2017 report (Scanteam 2017), Scanteam reiterated that in Afghanistan's challenging and rapidly changing context, ARTF remained a cost-efficient tool for channeling financial and technical support to Government priorities. Some poverty indicators had worsened, but ARTF had provided critical funding and acted as a stabilizing resource when other external funding decreased. Without ARTF funding for broad-based social services and community development, the situation would have been even more precarious. However, investment in business and infrastructure did not translate into economic growth. The capacity development of ARTF was disappointing, and monitoring and outcome tracking reporting was still much weaker than one would expect for a program that disburses USD 800-900 million a year.

It is interesting to note that the same themes are repeated in these Scanteam reports, which span a decade. For donors, ARTF is an efficient way to allocate their aid, while for Afghanistan, the Fund is important, since it finances the administration. The Fund had created some remarkable outputs, especially in small infrastructure, but this had not led to better overall economic performance nor had an impact on Government capacity or governance in general. Monitoring, reporting and evaluation remained weak, with no real outcome and impact assessment possible.

Finally, a report by Sida (2015) notes that ARTF is not sustainable by definition, as it depends on external aid flows and has a finite life. According to the report, ARTF contributed to widening access to public goods, notably education and health, with better health outcomes being achieved. ARTF also contributed to building a social contract between the state and its citizens through its support for service delivery. However, the efforts of the government's flagship NSP to foster that contract have had limited results. Sida also found room for considerable improvement in reporting on outcomes of ARTF investment projects.

Summing up, all the reports reviewed paint a similar picture: donors like ARTF because it is an efficient vehicle for allocating aid and serves as a forum for donor coordination and dialogue with the Afghan Government. ARTF is an important vehicle for financing the Afghan administration and, through its investment window, has financed an impressive number of outputs over the years. However, given ARTF's weak and unimproved monitoring and evaluation system, we do not know much about the outcomes and impacts of these outputs. There is some evidence that implementing and operating complex infrastructure is still beyond the capacities of the Afghan Government. In addition, ARTF has not had an impact on better governance or economic growth and, most importantly, has not contributed to better capacity in the Afghan Government.

3.2 Non-ARTF Governance Programs

3.2.1 Included Reports

Collin, C., Jawhary, A. M., & Robertson, L. (2014). UNDP-Afghanistan CPD outcome 3 evaluation: Final evaluation report. United Nations Development Program (UNDP).

Rao, M. P., & Alam, T. (2014). Final external evaluation report: national institution building project. United Nations Development Program (UNDP) Sub-National Governance.

3.2.2 Findings

Two reports on governance programs other than ARTF were produced by Collin et al. (2014) and Rao and Alam (2014).

Collin et al. (2014) evaluate one outcome of the UNDP country program: 'The state has improved ability to deliver services to foster human development, and elected bodies have greater oversight capacity' (p. 13). Four projects were intended to contribute to this outcome, all focused on building executive branch capacity. These were: the National Institution Building Project (NIBP), targeting the Independent Administrative Reform for Civil Service Commission to provide capacity development support to the Afghan Government at the national and sub-national levels; the National State Governance Project (NSGP), working directly with the Office of the President (OoP) to enable more effective policy and decision making at the center of government; the Afghanistan Sub-National Governance Project (ASGP), supporting the Government in developing capacity and systems to ensure effective implementation of sub-national governance strategies; and the Making Budgets and Aid Work (MBAW) project, supporting the Ministry of Finance (MoF) to develop capacity, processes and systems for public financial management and aid coordination.

Collin et al. (2014) find that the overall results of these projects were very limited. Structural problems impeded capacity building. The tensions between a formal, highly centralized state and de facto strong periphery - and differences between the merit-based model and patronage-based practices - hampered capacity building. In addition, management and partners focused on outputs more than outcomes. As a result, the effectiveness of the overall program was rated as 'medium to low' and its sustainability as 'low' (Collin et al. 2014, p. 33).

Rao and Alam (2014) evaluated the UNDP's flagship National Institution Building Project (NIBP), which sought to change the way the Government operates at both the national and sub-national levels. It placed national and international capacity development advisors (CDAs) in 17 line ministries and departments in Kabul and in 13 regional training centers (RTCs) and line departments in the provinces.

The report notes that the project lacked user buy-in and was felt to be forced on the ministries with little regard for their core institutional requirements. The project could have been more effective if, prior to its launch, there had been a diagnostic study of the requirements of the specific ministries, which could have developed a demand-driven project instead of a supply-driven one. The project also suffered from clashes within the civil administration and power dynamics between ministries. Finally, one of the glaring defects of the NIBP was the non-establishment of strong monitoring and evaluation. Rao and Alam (2014) report that 'NIBP could have identified exactly what it wanted to achieve' and should have been

demand driven and that 'every line ministry should be approached before a project document is written' (Rao and Alam 2014, pp. 56, 63). Long-term capacity development advisors should have been replaced with short-term technical advisors procured for specific deliverables on a case-by-case basis. In the end, the report concludes, 'there is no sufficient evidence to show the achievement of visible results' and 'no efforts were made on sustainability' (Rao and Alam 2014, pp. 52-53).

Summing up, both reports clearly suggest that these capacity-building programs had little impact. The main reasons appear to be that power struggles between center and periphery, and between and within ministries, made it difficult to build capacity, that there was little demand for these programs, and that programs were poorly planned and monitored.

3.3 Sub-National Governance

3.3.1 Included Reports

Chiwara, R. M., Afridi, H. R., & Jawhary, A. M. (2014). Afghanistan Sub-National Governance Programme phase II, January 2010–December 2014: final evaluation report. United Nations Development Program (UNDP).

Saed, I. (2017). Local governance program (LoGo), October 2015–November 2017: Midterm evaluation report. United Nations Development Program (UNDP).

Transtec. (2013). Evaluation report: Enhancing the capacity for inclusive local governance through synergies and sustainable linkages between communities and government in North Afghanistan. UDF-AFG-08-249. UNDEF.

3.3.2 Findings

Chiwara et al. (2014) evaluated the Afghanistan Sub-National Governance Programme Phase II, run by the UNDP. The project aimed to build the capacity of the Independent Directorate of Local Governance (IDLG), the Government agency tasked with administering sub-national administration. IDLG's main task, in the eyes of the donors, was to build institutions for meaningful decentralization of administrative relations. The anticipated outcomes of the project were improvement in the state's ability to deliver services to foster human development and greater oversight capacity on the part of elected bodies.

The evaluation found that the project lacked a clearly defined pathway to change and that there was no progress towards the overall objective of improving local public service delivery and management of public affairs. In addition, the project had weak financial reporting capacity: budget expenditure consistently increased every year but did not seem to reflect increased output delivery. Chiwara et al. (2014) concluded that decentralization involved the redistribution of power and access to resources as well as shifting lines of accountability. Since such reforms tend to be highly politicized, the project failed.

Saed (2017) evaluated another UNDP program, the Local Governance Program (LoGo), covering 13 provinces and 22 municipalities and designed to increase the capacity of provincial governors' offices, provincial councils, and IDLGs (Independent Directorates of Local Governance).

The evaluation found the project very slow in delivery, lacking a proper analysis that identified capacity development needs for the provincial governors' offices, civil society organizations, municipalities, and

IDLGs and suffering from weak buy-in and lack of political will by the central Government for meaningful decentralization. As a result, the project did not achieve much.

Transtec (2013) evaluated the project called Enhancing the Capacity for Inclusive Local Governance through Synergies and Sustainable Linkages between Communities and Government in North Afghanistan, run by United Nations Democracy Fund (UNDEF). The project aimed to promote suitable linkages between sub-national government, community development councils (CDCs) and civil society groups by working with local community, district and provincial levels in the project's three target provinces: Baghlan, Badakhshan and Takhar. Interventions included training local stakeholders on advocacy and gender-sensitive development and capacity-building for community leaders and local organizations.

The evaluation found that the training provided was rated very highly by the project participants. Most of the trainees interviewed were able to give examples of the ways in which they had been able to put their newly acquired skills to work. A partial exception was the course on gender-sensitive development. However, the project did not lead to more and better linkages between sub-national government, CDCs, and civil society groups and did not make a significant impact at the institutional level.

In sum, none of the reviewed projects led to increased capacity for subnational administration. The projects ignored the fact that there was no political will for meaningful decentralization. In addition, reforms in the field of decentralization were heavily politicized, and key stakeholders were involved in political struggles. Finally, these overly ambitious projects lacked a sound analysis by recipients of their real needs.

3.4 Rule of Law

3.4.1 Included Reports

Abbott, K., & Naderi, A. (2017). Independent evaluation of the UNDP JHRA phase II project.

Puric, O., Brooks, J., & Massoudi, K. (2018). Mid-term review of the Afghanistan access to justice program. UNDP.

3.4.2 Findings

Two studies on the promotion of the rule of law were carried out by Abbott and Naderi (2017) and Puric et al. (2018).

Abbott and Naderi (2017) evaluated UNDP's Justice and Human Rights in Afghanistan Project (JHRA), Phase II. The project aimed to develop frameworks for delivering legal aid: building institutional capacity in the courts, Attorney General's Office and Ministry of Justice; expanding access to the formal justice system at the district level; promoting human rights and legal remedies for violence against women; and other initiatives. The evaluation finds that the project failed to deliver results and provides a number of reasons for this.

To begin with, the theory of change behind the project was ideological and not in line with Afghan realities: there was no reason for the UNDP to think it could increase public trust in government rule-of-law institutions to any measurable degree, as that would have meant addressing personal interests and

corruption in that sector. Furthermore, providing 'access to justice' in a system that does not reliably deliver justice serves no purpose. The entire conceptual framework for the UNDP intervention was flawed, according to the evaluators. Moreover, capacity building focused too much on central institutions and not enough on rural areas, where most Afghans reside, and the project was too ambitious and understaffed and often underfunded.

Puric et al. (2018) provided a mid-term review of the Afghanistan Access to Justice Program run by the UNDP, the successor of the above-discussed JHRA Phase II project. Reading this evaluation, it is clear that important lessons from the predecessor project were ignored. Puric et al. (2018) found that it was very difficult to gauge impact because of the limited data and information available and the lack of any impact analysis conducted at the project level. The project was found to have had very little impact on improving access to justice for the people of Afghanistan. The report also noted that without proper evidence-based programming and a political economy analysis, the project would not be implementable; the project's theory of change should also be realistic, attainable and informed by the political realities of the national context.

In sum, the attempted rule-of-law projects were overly ambitious, were not based on the political-economic realities on the ground and were based on an ideological, unrealistic theory of change as well as weak impact assessments. As a result, the programs did not achieve their objectives.

3.5 Democracy Promotion

3.5.1 Included Reports

Gomez, C., & Baker, T. (2015). Final evaluation of the enhancing legal and electoral capacity for tomorrow (ELECT) phase II: January 2012-July 2015. United Nations Development Program.

Transtec. (2013). Evaluation report: Citizens' platform for democratic debates and dialogues in Afghanistan. UDF-AFG-09-316. United Nations Democracy Fund.

Transtec. (2015). Evaluation report: Involving women and youth CSOs in strengthening democratic debate and public news media around elections in Afghanistan. UDF-AFG-12-508. United Nations Democracy Fund.

3.5.2 Findings

There are three reports on democracy promotion programs headed by the United Nations Development Program (UNDP) and the United Nations Democracy Fund (UNDEF).

Gomez and Baker (2015) evaluated UNDP's Enhancing Legal and Electoral Capacity for Tomorrow (ELECT) program, which provided technical assistance designed to enhance the institutional capacity of the Independent Election Commission of Afghanistan (IEC). The project also provided technical and operational assistance for the 2014 presidential and provincial council elections and 2015 parliamentary elections.

The evaluation found that ELECT II succeeded in building the capacity of individuals within the IEC and the Independent Electoral Complaints Commission (IECC) to conduct technically viable elections in 2014. It helped the IEC to independently manage a portion of project funds during the 2014 election cycle. The report also said that a lack of ownership by Afghanistan's Government, donor fatigue, and uncertainty and erosion of security would be a problem in the future. It should be noted that ELECT was widely criticized as being very inefficient.²

Transtec (2013) evaluated the Citizens' Platform for Democratic Debates and Dialogues in Afghanistan, an UNDEF project implemented by Saba Media Org (SMO). The overall objective was to deepen understanding of democratic values among the Afghan people through messages on radio and TV and to establish a media platform by organizing live debates.

This project's targets were ambitious for the timeframe and resources of the grant, and the project implementation lacked adequate strategic perspective. SMO needed greater experience and capacity in order to meet the challenge of achieving social change through media. Ultimately, Transtec found that the project was not effective.

Transtec (2015) also evaluated another UNDEF project called Involving Women and Youth CSOs in Strengthening Democratic Debate and Public News Media around Elections in Afghanistan, a 15-month, USD 250,000 project implemented by Development and Humanitarian Services for Afghanistan (DHSA), an Afghan NGO based in Kabul.

The main expected outcomes of the project were to increase awareness among women and young people about the importance of participation in the 2014 presidential election and its relevance to Afghanistan's development. Increasing the role of CSOs for women and young people in generating effective debate around the democratic elements of the electoral process and the accountability of elected officials was also a target. The report found that the project's objective was ambitious for the timeframe and resources of this grant, that the project lacked a strategy to achieve or assess progress and effectiveness, and that no assessment mechanism to measure progress or impact was implemented.

In sum, development aid could substitute for the capacity needed for presidential elections. However, projects aimed at democratic awareness-raising did not reach their objectives.

² Norway (2016), A good ally: Norway in Afghanistan 2001-2014, Norwegian Ministry of Foreign Affairs and Ministry of Defence, Official Norwegian Reports NPOU 2016:8 (English translation from Norwegian): 105. <https://www.regjeringen.no/contentassets/09faceca099c4b8bac85ca8495e12d2d/en-gb/pdfs/nou201620160008000engpdfs.pdf>

3.6 Summary of Findings: Governance Programs

Governance and Capacity

The ARTF provided donors with an efficient vehicle for allocating their aid, serving as a forum for donor coordination and a vehicle for financing the Afghan administration. Through its investment window, ARTF financed an impressive number of outputs. However, implementing and operating complex infrastructure appears to be beyond the capacities of the Afghan Government. Nor has ARTF had an impact on better governance, economic growth or improved capacity in the Afghan Government.

Programs aimed at improving administrative capacity largely failed, especially in decentralizing from the center to the provinces. Power struggles between center and periphery, and between and within ministries, made it difficult to build capacity. Ultimately, there was little demand for these programs, which built no increased capacity for the sub-national administration.

Rule of Law

The two projects reviewed were overly ambitious, were implemented without considering the political-economic realities on the ground, were based on an ideological, unrealistic theory of change, and had weak impact assessment mechanisms. The projects failed.

Democracy Promotion

Development aid could substitute for the capacity needed for presidential elections. Projects aimed at democratic awareness-raising failed.

4 Gender Programs

4.1 Included Reports

Bernard, A. (2014). Forward-looking strategic evaluation of the UNICEF-supported female literacy program (2010-2013) in 34 provinces of Afghanistan. UNICEF.

FAO UN (Food and Agriculture Organization of the United Nations, Office of Evaluation). (2016). Cluster evaluation of: Strengthening the Role of Women in Agriculture Development for Improved Household Food; Strengthening Policy Development and Coordination for Food and Nutrition Security in Afghanistan; and Support of Extension Systems, Household Food and Livelihood Security. Project Evaluation Series.

Haarr, R. N. (2015). External evaluation of the elimination of violence against women (EVAW) special fund 2008-2014. UN Women Afghanistan Country Office.

Mushinga, M., & Fattahi, Z. (2016). UNDP gender equality project II: Final evaluation report. United Nations Development Program (UNDP).

Ojha, G. P., & Fattahi, A. Z. (2015). Women's empowerment and gender equality project: Mid-term evaluation project. United Nations Development Program (UNDP).

Sayara Research. (2015). Impact evaluation report: EU empowerment project. Oxfam. Intervention.

Transtec. (2014). Evaluation report: Raising awareness about women's social, political, and economic rights in Afghanistan. UDF-AFG-10-379. UNDEF.

Watkins, F., & Azarbaijani-Moghaddam, S. (2019). Evaluation of UN women country program in Afghanistan: Final report. Sida Decentralized Evaluation.

4.2 Findings

We have eight evaluation reports on gender.

The evaluated programs have a wide number of objectives, among them increasing literacy for women and supporting women in agriculture, but also working towards gender equality, eliminating violence against women, and advocacy for women's rights. As we will see, modest projects with more tangible and immediate benefits for women tended to be much more effective than projects with more ambitious, transformative goals.

Bernard (2014) provided a forward-looking strategic evaluation of the UNICEF-Supported Female Literacy Program (2010-2013) in 34 provinces of Afghanistan. The project aimed to contribute to national literacy targets through supporting a 9-month Literacy Center training course for rural women. The training was delivered in daily two-hour classes in private homes and mosques by purpose-trained female teachers and using a gender-tailored curriculum. The project was generally effective insofar as it enrolled the 'right' people (adolescent girls and women denied the opportunity of a formal education). It worked as intended in exposing learners to reading. Learner and teacher satisfaction levels were good, and dropout rates and absenteeism were minimal.

FAO UN (2016) reports results from an evaluation of its project Strengthening the Role of Women in Agriculture Development for Improved Household Food. The project supported, among other things, income-generating activities, including kitchen gardens in Parwan, greenhouses and mushroom farms in Badakhshan, and orchards in Istalif, Kabul. Mushroom production, vegetable production, and fruit and vegetable processing are some of the activities specifically targeted at women.

The evaluation found that the program contributed to increased knowledge and skills development for women in mushroom farming and technology transfer in mushroom cultivation. These skills were also transferred from the women trainees to community farmers. This means that women were acting as informal extension workers in the community, a role previously filled only by men.

Although women had increased access to extension services, their access to markets to sell their products was not significantly increased. Nevertheless, mobility of trained women was increased, and they could visit the extension office/training center to request information. Their prestige at home and status in the family increased as a result of their ability to discuss nutrition and food issues and as a result of earned income, which is largely used for family necessities such as food, clothes, and children's school supplies.

In sum, the program had some positive effects on women's agricultural knowledge and improved both their role in the household and their mobility. It seems that these achievements were possible because the programs targeted 'women-specific' activities within agriculture.

Sayara Research (2015) evaluated the Women Empowerment Project (WEP) implemented by Oxfam, which provided capacity training to civil society organizations to increase their capacity to help empower the women they assist. Women's Rights Committee (WRC) meetings of advocates from Government and NGOs were held regularly in each province to discuss women's rights and challenges and to build linkages between groups. Community seminars and dialogue sessions were held by civil society organizations with local men and women to raise awareness of women's rights and to understand the views of the general population on the topic.

Through the WEP, there was a significant increase in women's access to healthcare, derived primarily from two sources: 1) building and staffing clinics in areas that needed them, and 2) increasing awareness via health-related talks during WRC and community meetings. Increased access to education and knowledge for women of all ages has also been attributed to the WEP. Advocacy of a girl's right to attend school and the use of religious scholars to support this claim resulted in a perception among targeted communities that there is now a higher rate of female school attendance. Advocating for women's rights and empowerment through an Islamic lens increased WEP's efficacy. The topics discussed during WEP programming were much better received as a result of the active recruitment of religious figures to promote women's rights awareness.

The evaluation concluded that Oxfam's Women Empowerment Project (WEP) resulted in greater access to healthcare and education for beneficiaries and helped to develop linkages between Government and NGOs advocating for women's empowerment.

The three evaluation reports discussed so far referred to gender projects with relatively modest, yet tangible, outcomes, such as literacy, improved nutrition, and better access to health. The remaining evaluations refer to much more ambitious objectives. First, Haarr (2015) evaluated the Elimination of Violence Against Women (EVAW) Special Fund 2008-2014, run by the UN Women Afghanistan Country Office. The EVAW Special Fund provided small grants to civil society organizations (CSOs) to implement

community-based initiatives. The evaluation noted that many of the civil society organizations (CSO) had little capacity and that focusing on overcoming these weaknesses would improve effectiveness. The Fund itself faced challenges regarding funding, staffing and management, which impacted efficiency and effectiveness. The evaluation notes that it was too early to assess sustainability but that the Fund was 'well-focused on one of its long-term goals of receiving more substantial proposals from strengthened CSOs for projects that will produce measurable positive change' (Haarr 2015, p.29). In other words, after three years, the Fund had not yet received many substantial proposals that led to positive change.

Ojha and Fattahi 2015 and Mushinga and Fattahi (2016) both evaluated UNDP's Gender Equality Project II, the former being a mid-term evaluation and the latter a final evaluation. The Gender Equality Project II supported gender mainstreaming in various sectors of government at national and sub-national levels with three main pillars: 1) capacity development, 2) economic empowerment of women, and 3) access to justice and human rights.

The midterm report (Ojha and Fattahi 2015) observed that the overall effectiveness of the project was low and that achievements of the project were unlikely to be sustained mainly because of the low technical and financial capacity of service providers. The report noted that 'some participating women entrepreneurs have expanded their business and have created additional employment; some vulnerable women have earned income engaging themselves in microenterprises, but their number is too small' (Ojha and Fattahi 2015, p. x).

In the final evaluation, Mushinga and Fattahi (2016) report that 'commendable results across all three pillars have formed strong foundations for women[s] empowerment and gender equality in Afghanistan which, however, still need some additional support to continue producing tangible outputs as well as strengthening prospects for the activities' sustainability and potential impact' (Mushinga and Fattah 2016, p. 7). This is a convoluted way of saying that the evaluation had not identified impacts by the project and sustainability was questionable. The report also noted a 'weak monitoring and evaluation system and high operation costs combined with few activities in some areas' (Mushinga and Fattah 2016, p. 7). The report recommends that a proper management record should be in place and that activities should be planned with a committed budget and hints that this may have been 'a huge project with limited resources.' Future projects should take 'a "deeper" versus a "broader" approach', which would help to 'produce tangible/measurable results as well as developing latitude for measuring impact as compared to going for broader activities whereby results may remain superficial, unsustainable and with minimal or zero impact' (Mushinga and Fattah 2016, p. 38).

Transtec (2014) evaluated the Raising Awareness about Women's Social, Political, and Economic Rights in Afghanistan project run by UNDEF. The project sought 'to promote the democratic rights and participation of women' (Transtec 2014, p.1). Key outcomes were to increase awareness and support of women's rights in 15 provinces in Afghanistan and to increase participation in advocacy for women's rights through regional women's rights networks in the target provinces. This was a very small project with a grant of USD 300,000. The report found it 'impossible to confirm that the project was implemented in full based on the information available to them,' and 'there is nothing to suggest that the project had any impact' (Transtec 2014, p. 3).

Finally, Watkins and Azarbaijani-Moghaddam (2019) provided an evaluation of the UN Women Country Program in Afghanistan. The evaluation was commissioned by Sweden's Sida, which, together with Norway and Finland, contributed 40% of the total USD 39.37 million budget, which makes this by far the largest gender project in our sample. The evaluation covers the period 2014-2017 and, while looking

primarily at support from Sida, Norway, and Finland, considers the implementation of the program as a whole. The interventions of the program consisted of coordination and advocacy, political and economic empowerment, developing partnerships with institutions and individuals who can serve as national/ sub-national voices on the issues of eliminating violence against women (EVAW), extensive support to ministries - mostly the Ministry of Women's Affairs (MoWA) - and building institutional capacity both for Government and non-governmental organizations.

Watkins and Azarbaijani-Moghaddam (2019) find that while the program was broadly relevant, its priorities were driven more by funding than by a strategic or long-term vision of what was needed. There was some progress made in the achievement of outputs, but much less progress in the achievement of outcomes or their sustainability. The lack of results was partially due to the difficult context, but many of the issues affecting the implementation also came from the management of the program. The program was 'overambitious in terms of what could be achieved in a difficult context working with partners with limited capacity,' and 'resources available were spread thinly across a complex program' (Watkins and Azarbaijani-Moghaddam 2019, p. 21). The evaluation found considerable concern about UN Women's lack of overall leadership, slow decision making and poor communication. In sum, it 'was difficult to find evidence of a direct contribution that UN Women has made to poverty reduction and to gender equality results' (Watkins and Azarbaijani-Moghaddam 2019, p. 5).

The report also mentions a range of cultural factors which made legislation ineffective, as well as unintended consequences of such a program being implemented in the Afghan context, especially with regard to one aspect of the program, namely women's shelters: it is worth quoting the relevant section:

.....

'There is strong evidence of unintended effects of efforts to bring in legislation on EAW and to provide protection for women who are affected by violence. There is evidence from a number of interviews of continued controversy and argument about the efforts to introduce the EAW Law: including significant disagreements between women's organizations over the approaches taken to introducing the law and concerns over the introduction of the Anti-Harassment Law in 2016. There is also evidence from interviews of very negative attitudes to women's shelters. Examples were given where women who have used shelters have then been unable to maintain links with family or to move on, although some efforts are being made to address this in retrospect. However, the strongest evidence comes from the 2017 Samuel Hill research on the EAW Commissions. With regard to legislation the research found that while most women interviewed stated knowing that the EAW law exists (that is to say, had a general understanding that this law existed to protect them from violence and abuse), there was very limited impact because of social stigma, norms and restrictions. Additionally, women reported severe flaws in the implementation of the law. Flaws related to implementation included uncertainty over the sensitive handling of their cases, corruption (fear/perception that the perpetrator's family could pay officials to decide in their favour) and as a result of corruption, the fear of backlash for bringing their domestic case in the public.

With regard to protection the research found that the most unanimous finding across every single person interviewed for this study is the negative perception associated with women's shelters. Shelters are considered by many as places where prostitution occurs, where women may be raped by the police, and where women indulge in illicit activities. Many women interviewed strongly believe that once a woman goes to a shelter, even if she is guaranteed safety, she can never come back into society or join her family due to the stigma attached to these centers and misconceptions of shelters'.

.....

(Watkins & Azarbaijani-Moghaddam, 2019, p. 30)

4.3 Summary of Findings: Gender

Overall, the effectiveness of gender programming appears to be extremely low, but with pockets of modest success.

For example, a small NGO-run literacy program for rural women met its objectives. Another NGO project increased access to health and education for women by setting up and staffing health clinics and by advocating, with the help of religious authorities, for girls' rights to attend school. Another project trained women in women-specific activities within agriculture, such as mushroom farming and kitchen gardening, and improved women's mobility, income and standing in their communities. In sum, small, modest projects embedded in traditional structures helped to increase access to health and education and modestly improved livelihoods for women.

By contrast, larger, more ambitious projects aimed directly at changing gender norms and relations failed. To some extent, the failure is attributable to the weak leadership and capacity of the lead agency (in this case UN Women and UNDP). To a much larger extent, the projects failed because they were overly ambitious and ignorant of the context. They also overestimated the capacity and political will of the partner and underestimated the traditional cultural fabric of Afghanistan. Thus, there was hardly any impact from the large UN Women country program 2014-2017 on capacity for and awareness about eliminating violence against women (EVAW), of the UN Women Special Fund for Elimination of Violence Against Women (EVAW), or of the UNDP gender mainstreaming program.

5 Rural Development Programs

5.1 Included Reports

Altai Consulting. (2017). AFD agriculture programs in Afghanistan (2005-2014): Final report. AFD Evaluation 65.

Bhattacharjee, A., Postgate, D., & Andersen, H. (2013). Evaluation of CPD outcome 6: Diversified livelihoods, private sector development and public-private partnership. UNDP Afghanistan.

Emmott, S., & Jawhary, A. M. (2014). Evaluation of the national area-based development program (NABDP) in Afghanistan. United Nations Development Program (UNDP).

Hussain, T., & Wasim, M. (2017). Mid-term review of UNDP GEF-LDCF2 project Afghanistan (27 April 2014-30 June 2017): Strengthening the resilience of rural livelihood options for Afghan communities in Panjshir, Balkh, Uruzgan and Herat Provinces to manage climate change-induced disaster risks. UNDP.

5.2 Findings

We have four evaluations on rural development program in our sample.

Bhattacharjee et al. (2013) evaluated one outcome of the UNDP county program (2010–2014), 'diversified livelihoods, private sector development and public-private partnerships' (Bhattacharjee et al. 2013, p. 6). There were five projects under outcome 6, contributing to such diverse objectives as strengthening district level institutions and governance structure, improving access to services and infrastructure for the rural poor, stabilization in less secure regions and districts by addressing economic and livelihoods issues; support agricultural enterprises, and supporting gender equality and empowerment for women by enhancing the Ministry of Women's Affairs' capacity for policy making, by supporting women's entrepreneurship's skills and by increasing access to justice for women.

Given the wide array of objectives, it is not surprising that the evaluation is often rather vague in its findings and does not address all objectives in the same amount of detail. Some of the key findings worth highlighting are the following:

The various projects contributed to creating participatory structures including district development assemblies, provincial Women's Development Councils, and provincial Peace Councils. However, it was unclear whether these structures contributed to real outcomes, and the report rated effectiveness as modest, efficiency as low, and sustainability as questionable. One of the weaknesses of the district level structures was that their *raison d'être* was perceived to be linked to project implementation rather than 'governance' as such (Bhattacharjee et al. 2013: 26).

Moreover, it remained unclear whether women's participation in these district forums had increased, since no impact assessment was conducted (Bhattacharjee et al. 2013, p. 26). However, projects were modestly successful in providing livelihoods for rural women, engaging them in mostly farm-based and traditional skill-based activities. (Bhattacharjee et al. 2013, p. 28). Supporting agricultural small businesses proved to be difficult and did not contribute to poverty reduction (Bhattacharjee et al. 2013, p. 28).

Emmott and Jawhary (2014) evaluated another UNDP program, the National Area-Based Development Program (NABDP), a joint intervention of UNDP and the Ministry of Rehabilitation and Rural Development (MRRD). The program sought to reduce poverty and vulnerability through a focus on productive rural infrastructure and institutionalization of district development assemblies (DDAs). The project helped to establish DDAs in 388 districts. Through these DDAs, small infrastructure was implemented in communities, ranging from improved access to energy, irrigation to transport infrastructure. Thousands of short-term jobs were also created, mainly through 'cash for work'.

Emmott and Jawhary (2014) conclude that the program was 'effective in delivering a vast number of requested and appreciated infrastructure (...) but that its ambition has been greater than its capacity' (Emmott and Jawhary 2014, p.26). They also state that effectiveness on gender equality has been limited (Emmott and Jawhary 2014, p.26). The report suggests that the program did have impacts at the 'very local household or small community level, and that may have been important, even if only in the short term' (Emmott and Jawhary 2014, p.27), but there was no systemic impact on reduction in poverty and food security. Finally, the authors also caution that the program was unlikely to be sustainable.

Hussain and Wasim (2017) provided a mid-term review of UNDP's project, Strengthening the Resilience of Rural Livelihood Options for Afghan Communities in Panjshir, Balkh, Uruzgan and Herat Provinces to Manage Climate Change-Induced Disaster Risks. One component of the project refers to rural development and intends to support communities by providing them with livelihood options, improved water management and irrigation infrastructure. Expected outcomes included enhanced and diversified rural income and livelihood opportunities for vulnerable communities, and rehabilitated and improved productive irrigation infrastructure.

This is a rather small project with modest yet tangible outcomes. The evaluation found that the project implemented effective livelihood training and supported local small infrastructure, especially in the irrigation sector. As a result, livelihoods at the local level were moderately improved.

Finally, Altai Consulting (2017) provided a cluster evaluation of Agence Française de Développement (AFD) agriculture programs in Afghanistan (2005-2014), which funded 15 agriculture-related projects worth a total of EUR 42 million. All projects were managed by MAIL (Ministry of Agriculture, Irrigation and Livestock) and included the following: the North-East Agriculture Support Project (NEASP); the New Afghanistan Project for Cotton and Oil Development (NAPCOD); the Horticulture and Cooperative Development Project (HCDP); the Beekeeping Development Project (BDP); and the Animal Husbandry Development Project (fish farming/artificial insemination). A number of other NGO projects were also supported, among them the Yakawlang Rural Development Project, the Passive Solar Houses Project, the Wardak Rural Development Project, and the Central Highlands project.

Results were very mixed. The cotton cluster project ultimately failed as an economically profitable activity because it was based on unrealistic assumptions stemming from a lack of understanding of the local context. Developing the horticulture sector through a dedicated cooperative faced numerous challenges related to poor project management and an overall lack of focus. The fish farming project failed when, sadly and unexpectedly, all fingerlings died.

On the other hand, beekeeping projects have had significant achievements in contributing to more secure livelihoods. Projects for supporting the dairy sector (increased milk and dairy production through genetic improvement of livestock, strengthening the dairy value chain benefits to producers, and promoting professional organizations) has succeeded in creating a viable economic activity with a large potential impact through its artificial insemination (AI) project. It has had less success so far on activities targeting other aspects of the dairy value chain.

NGO projects fared a little better than those managed by MAIL, as they were well defined from inception, with scoped work, milestones, precise timeframes, and lists of expected deliverables. By contrast, MAIL projects did not perform at the anticipated level because of a lack of capacity and ownership on the part of MAIL.

5.3 Summary of Findings: Rural Development

Many rural development projects were successful at providing small infrastructure, such as roads, irrigation, access to energy, greenhouses, and so on. Access to services and basic infrastructure has improved, and this contributed to improved livelihoods at the very local level.

There is some evidence that smaller projects implemented by NGOs with more community buy-in worked better. However, there was no reduction in overall poverty rates, and much of the infrastructure is not sustainable. There has also not been much improvement in employment creation or in developing agriculture businesses with value chains.

Many of these projects were implemented through newly created community or district level organizations. These organizations were often useful as project implementers, but they did not develop into organizations that could exercise governance in rural areas. It is also not clear whether these organizations increased participation, including participation by women.

The capacity of the Government partner institutions in rural areas was weak, and it proved difficult to build more capacity.

6 Education Programs

6.1 Included Reports

Integrity Watch Afghanistan. (2018). Education compromised? A survey of schools in 10 provinces of Afghanistan.

Samuel Hall Consulting. (2015). School-in-a-box 2015 evaluation. The Womanity Foundation.

UNICEF. (2015). Let us learn (LUL): Formative evaluation. UNICEF Afghanistan Country Office.

6.2 Findings

There are three evaluations on education in our sample.

While one looks primarily at the state of infrastructure, two are concerned with primary education for girls. Given the importance of the sector for Afghanistan, the number of evaluations is very small. Moreover, the flagship education program EQUIP is not among the programs evaluated here. Nevertheless, the three evaluation reports provide interesting vignettes.

Integrity Watch Afghanistan (2018) surveyed schools in 10 provinces, visiting 276 schools over three years to assess their operation and physical condition. Their findings suggest that despite billions of dollars invested by donors, the state of physical infrastructure and quality of education remained poor. Access to quality education faced challenges including security concerns, insufficient and low-quality school buildings, limited access to water and sanitation, limited availability of qualified teachers (especially female teachers), insufficient learning materials, and corruption and mismanagement of resources.

The report suggests that local communities should be more engaged in the planning, construction, and maintenance of infrastructure by developing an institutional framework for effectively channeling community involvement. It also recommends that low cost options for school buildings be explored because Afghanistan will not have the resources to meet the growing demand. In order to bolster female participation, all schools should have boundary walls, well-maintained and gender segregated toilets, and a crèche facility on school premises for female teachers.

Samuel Hall Consulting (2015) evaluated a small NGO program run by the Womanity Foundation called School-in-a-box, which was launched in 2007 in Kabul and replicated in 11 other state schools. The program aimed to improve the quality of girls' primary and secondary education in Afghanistan through teacher training, student counselling, improved infrastructure, and community outreach. The report found that the program had a positive effect on community attitudes toward girls' education and created a snowball effect: as more girls were seen attending school, more were likely to follow. Womanity's contributions to improved sanitary conditions was very important. Challenges remained, including a lack of qualified teachers and insecurity, which remains the main challenge to girls' education.

UNICEF (2015) evaluated their own Let Us Learn (LUL) program, aimed at providing primary school education (grades 1 to 6) for children aged 9 to 15, mainly girls who otherwise would not have a chance to attend school and obtain an education. LUL works via community participation to create Accelerated Learning Centers (ALCs) in remote communities without access to a school. Communities provide the learning spaces and identify educated members who can be trained as ALC teachers. The Ministry of Education is responsible for providing textbooks, while UNICEF trains teachers and provides their salaries, teaching and learning materials, stationery, carpets, and firewood.

The evaluation found that LUL met its objectives, reaching its intended participants. The program produced opportunities for female teachers to teach in their local community, making it possible for girls to obtain an education as well as contributing to improved equality for girls in Afghanistan. An increase in community support for ALCs, and in particular for girls' education, was found across all five provinces. Increased learning outcomes were also observed, primarily in the areas of literacy and numeracy, and students seemed to be ahead of their formal school counterparts.

6.3 Summary of Findings: Education

Despite massive investment by donors, the state of physical infrastructure and the quality of education still need improvement, and many gains are not sustainable. The need for infrastructure and qualified teachers remains very high. However, well-targeted and well-managed interventions have the potential to improve the quality of primary teaching, to provide primary education for children in remote regions, to increase access to education for girls, and to increase support for girls' education.

7 Health Programs

7.1 Included Reports

Speakman, E. M., Shafi, A., Sondorp, E., Atta, N., & Howard, N. (2014). Development of the community midwifery education initiative and its influence on women's health and empowerment in Afghanistan: A case study. *BioMed Central* 14, 111.

Turkmani, S., Currie, S., Mungia, J., Assefi, N., Javed Rahmanzai, A., Azfar, P., & Bartlett, L. (2013). Midwives are the backbone of our health system: Lessons from Afghanistan to guide expansion of midwifery in challenging settings. *Midwifery* 29, 1166–1172.

Zainullah, P., et al. (2014). Establishing midwifery in low-resource settings: Guidance from a mixed-methods evaluation of the Afghanistan midwifery education program. *Midwifery* 30, 1056–1062.

7.2 Findings

All three reports on health projects focus on midwifery in Afghanistan, and all three report positive results.

Speakman et al. (2014) evaluated the Community Midwifery Education (CME) program, which began training rural midwives in 2002 and was scaled up nationally in 2005. The report found that the program consistently contributed to positive indicators, including a six-fold increase in qualified midwives since 2002, an increase of up to 28% in skilled deliveries, and a reduction of up to 1,273 per 100,000 in maternal mortality rates. Many factors contributed to these achievements, including readily accessible funding, international technical expertise, implementing NGOs on the ground, establishment of regulatory and accreditation bodies and, crucially, midwives ready to become community role models. The program is considered by stakeholders to be a positive model for promoting women's education, employment and health and highlights that maternal care provides an example of female empowerment.

Turkmani et al. (2013) evaluated the pre-service midwifery education program at accredited midwifery schools, finding that midwives described overall satisfaction with the quality of their education. Midwives and stakeholders perceived that women were more likely to use maternal and child health services in communities where midwives had been deployed. The report concluded that the contributions of midwives to Afghanistan - from increased maternal health care service utilization to changing community perceptions of women's education and professional independence - is overwhelmingly positive.

Zainullah et al. (2014) examined the costs and graduate performance outcomes of the two types of pre-service midwifery education programs, Institutes for Health Science (IHS) and Community Midwifery Education (CME). They found that the pre-service midwifery education experience of Afghanistan can serve as a model for rapidly increasing the number of skilled birth attendants. Improving midwifery education has supplied competent skilled birth attendants in Afghanistan and has expanded the human resources needed to promote reproductive health, with high acceptability among local officials, community members, and midwives themselves.

7.3 Summary Findings: Health

The three evaluations on midwifery training suggest that investing in midwifery is an efficient, effective, sustainable way to improve health outcomes in child and maternal health.

8 Country Programs of Multilateral Donors

8.1 Included Reports

Independent Evaluation Group. (2013). Evaluation of World Bank programs in Afghanistan, 2002-11. Washington, DC: World Bank Group. <https://openknowledge.worldbank.org/handle/10986/15768>

UNDP. (2014, July). Assessment of development results: Evaluation of UNDP contribution in Afghanistan. Independent Evaluation Office, United Nations Development Program.

8.2 Findings

Lastly, we also very briefly discuss two country program evaluations: an evaluation of World Bank programming from 2002 to 2011 by the Independent Evaluation Group, and an evaluation of UNDP programming from 2009 to 2014.

Both evaluations cover long time spans and a very large number of programs and projects in almost all development sectors. Given this scope and breadth, it is clear that evaluation findings are at a very general level.

We briefly report the main findings of both evaluations for the sectors in which we are primarily interested. It is noteworthy that, by and large, this bird's-eye view supports the findings of the 37 evaluation reports that we have summarized above.

The World Bank's strategy for Afghanistan was focused on what it calls three pillars: building the capacity of the state and its accountability to its citizens; promoting growth of the rural economy and improving rural livelihoods; and supporting growth of the formal private sector, including through infrastructure development.

With regard to the first pillar, the report notes progress in public management and especially in public financial management. Nevertheless, the report warns that public administration remains vulnerable, as there is little evidence that the new laws, procedures, and regulations are translating into improved civil service performance and that many achievements have 'relied mainly on a "second civil service" of contracted staff who are paid relatively high salaries' (Independent Evaluation Group 2013, p. xii). Other sectors where progress was made are public health and telecommunications. The report also notes that good results were achieved in primary education, especially in regard to enrolment.

Success in the second pillar was rare. While the bank helped to provide a large amount of small infrastructure, above all rural roads, 'the overall impact on the core objective of promoting growth in the rural economy and improving rural livelihoods has been modest' (Independent Evaluation Group 2013, p. xiv).

Results in the third pillar were again mixed: results have been noteworthy in microfinance and communications but much weaker in strengthening an investment-friendly climate and promoting the private sector. Both lack of security and a difficult policy and governance environment explain the modest results.

The report also warns that sustainability of development gains remained a major risk because of capacity constraints.

The UNDP (2014) report is mainly focused on listing outputs and is not set up to assess effectiveness or impacts. However, in key aspects, the report closely echoes the findings of the World Bank. The UNDP report notes that poor security had strong negative effects on timelines and costs, which were exceptionally high. Sustainability was assessed as very low, with capacity relying on the donor-funded 'second civil service'. Capacity building measures were often 'stop-gap', and did not lead to sustainable increase in capacity. Furthermore, few initiatives were based on a thorough capacity assessment. In the area of gender, cultural sensitivities proved a significant constraint, and 'several generations will likely pass before deeper changes can begin to take root' (UNDP 2014, p. xiv). Few to no results were achieved in sectors such as anti-corruption, justice sector, and anti-narcotics. Finally, the report found that UNDP achieved little in the area of poverty reduction and income generation.

9 Conclusions

1. There are **clear differences between sectors with regard to effectiveness**. Among the **most effective interventions** were interventions in the sectors of **primary education and health**. **Much less effective** were interventions in **gender, rule of law, capacity building for sub-national administration and decentralization, and promotion of democratic values**. There are mixed results in rural development, with some strengths in providing small infrastructure and weaknesses in providing governance institutions and supporting growth.
2. With regard to education, **well-targeted and well-managed interventions were able to improve the quality of primary teaching**, provide primary education to children in remote regions, increase access to education for girls, and increase support for girls' education. Yet the need for infrastructure and qualified teachers remains very high. Despite massive investment by donors, the state of physical infrastructure and the quality of education still need improvement, and many gains are not sustainable.
3. With regard to the health sector, three evaluations on midwifery training suggest that **investing in midwifery is an efficient, effective, and sustainable way to improve health outcomes** in child and maternal health.
4. The **Afghan Reconstruction Trust Fund (ARTF)** provided donors with an efficient **vehicle for allocating aid** and served as a forum for donor coordination. ARTF was also an important vehicle for financing the Afghan administration.
5. **ARTF has financed a large number of small infrastructure projects**. The evaluation reports suggest that less complex infrastructure projects usually fared better than more complex projects and that the Afghan Government does not yet have enough capacity for planning, building, and maintaining more complex infrastructure, such as school buildings
6. **ARTF has not had an impact on better governance or on economic growth. Most importantly, it has not contributed to increased capacity in the Afghan Government.**
7. Monitoring, reporting, and evaluation of ARTF remained weak with no real outcome and impact assessment possible, despite recurrent criticism of donors.
8. **Capacity-building programs for the Afghan administration, both at national and at sub-national level, largely failed**. This is especially the case for programs that aimed to build capacity for managing relations between the center and the provinces on the basis of meaningful decentralization. Decentralization involves the redistribution of power and access to resources. This is highly politicized, and reforms in politicized fields are rarely effective.
9. Other factors that hampered capacity building were entrenched patronage-based practices within the Government, a lack of buy-in from the Government, donor-driven project design with little regard for the core institutional requirements of the partner institutions, and lack of political will on the part of the Government, especially for decentralization.
10. We only have two evaluations of **rule of law programs**. These programs aimed at developing frameworks for delivering legal aid, building institutional capacity in the courts, expanding access to the formal justice system, and promoting human rights. **Evaluations suggest that both programs failed** because they were overly ambitious, were not based on the political-economic realities on the ground, and were based on an ideological and unrealistic theory of change.

11. **The effectiveness of gender programming was in general low.** Modest projects embedded in traditional structures helped to increase access to health, education and economic opportunities at the household level. Literacy programs for rural women were also effective. By contrast, larger, more ambitious projects aimed directly at changing gender norms and relations failed. To some extent, that failure is attributable to weak leadership and weak capacity in the lead agency. To a larger extent, the projects failed because they were overly ambitious and ignorant of the context, overestimated the capacity and political will of the partner, and underestimated the traditional cultural fabric of Afghanistan. This is why there was hardly any impact from a large UN Women country program on capacity and awareness about eliminating violence against women, the UN Women special fund for elimination of violence against women, or UNDP gender mainstreaming program.
12. Many **rural development** projects, often implemented through newly created community-level organizations, **were successful at providing small infrastructure**, such as roads, irrigation, and access to energy. Better access to services and basic infrastructure contributed to improved livelihoods. However, there was **no reduction in overall poverty rates**, and much of the infrastructure is not sustainable. There is some evidence that smaller projects implemented by NGOs with more community buy-in worked better.
13. Many of these projects were implemented through newly created community- or district-level organizations. These organizations were often useful as project implementers, but they did not develop into organizations which could exercise governance in rural areas.
14. The capacity of the Government partner institutions in rural areas was weak, and it proved difficult to build more capacity.
15. Many of the above-mentioned findings are echoed in two **country program evaluations of World Bank and UNDP programs**. The reports point to some **progress in public management** and especially in public financial management. Nevertheless, the report warns many achievements were due to the 'second civil service' of contracted international staff. The reports also note that **decent results were achieved in primary education and health**. With regard to **rural development**, the report notes that while a large number of small infrastructure projects have been implemented, the overall **impact on growth in the rural economy and on improving rural livelihoods has been modest**. Interventions aimed at promoting the private sector were also not effective. In the area of gender, cultural sensitivities proved a significant constraint. **The reports also warn that most achievements are not sustainable.**

10 References

- Altai Consulting.** (2017). AFD agriculture programs in Afghanistan (2005–2014): Final report. AFD Evaluation 65.
- Bernard, A.** (2014). Forward-looking strategic evaluation of the UNICEF-supported female literacy program (2010–2013) in 34 provinces of Afghanistan. UNICEF.
- Bhattacharjee, A., Postgate, D., & Andersen, H.** (2013). Evaluation of CPD outcome 6: Diversified livelihoods, private sector development and public–private partnership. UNDP Afghanistan.
- Chiwara, R. M., Afridi, H. R., & Jawhary, A. M.** (2014). Afghanistan subnational governance program phase II, January 2010–December 2014: Final evaluation report. United Nations Development Program (UNDP).
- Collin, C., Jawhary, A. M., & Robertson, L.** (2014). UNDP–Afghanistan CPD outcome 3 evaluation: Final evaluation report. United Nations Development Program (UNDP).
- Emmott, S., & Jawhary, A. M.** (2014). Evaluation of the national area-based development program (NABDP) in Afghanistan. United Nations Development Program (UNDP).
- FAO UN (Food and Agriculture Organization of the United Nations, Office of Evaluation).** (2016). Cluster evaluation of: Strengthening the role of women in agriculture development for improved household food; Strengthening policy development and coordination for food and nutrition security in Afghanistan; and Support of extension systems, household food and livelihood security. Project Evaluation Series.
- Gomez, C., & Baker, T.** (2015). Final evaluation of the enhancing legal and electoral capacity for tomorrow (ELECT) phase II: January 2012–July 2015. United Nations Development Program.
- Haarr, R. N.** (2015). External evaluation of the elimination of violence against women (EVAW) special fund 2008–2014. UN Women Afghanistan Country Office.
- Hussain, T., & Wasim, M.** (2017). Mid-term review of UNDP GEF-LDCF2 project Afghanistan (27 April 2014–30 June 2017): Strengthening the resilience of rural livelihood options for Afghan communities in Panjshir, Balkh, Uruzgan and Herat Provinces to manage climate change-induced disaster risks. UNDP.
- Independent Evaluation Group.** (2013). Evaluation of World Bank programs in Afghanistan, 2002–11. Washington, DC: World Bank Group. <https://openknowledge.worldbank.org/handle/10986/15768>
- Integrity Watch Afghanistan.** (2018). Education compromised? A survey of schools in 10 provinces of Afghanistan.
- Management Systems International (MSI).** (2017a). Year 1 annual report (November 2015–October 2016). Afghanistan Reconstruction Trust Fund.
- Management Systems International (MSI).** (2017b). Year II annual report (November 2016–October 2017). Afghanistan Reconstruction Trust Fund.
- Mushinga, M., & Fattahi, Z.** (2016). UNDP gender equality project II: Final evaluation report. United Nations Development Program (UNDP).
- Ojha, G. P., & Fattahi, A. Z.** (2015). Women’s empowerment and gender equality project: Mid-term evaluation project. United Nations Development Program (UNDP).
- Rao, M. P., & Alam, T.** (2014). Final external evaluation report: National institution building project. United Nations Development Program (UNDP) Sub-National Governance.

- Saed, I.** (2017). Local governance program (LoGo), October 2015–November 2017: Midterm evaluation report. United Nations Development Program (UNDP).
- Samuel Hall Consulting.** (2015). School-in-a-box 2015 evaluation. The Womanity Foundation.
- Sayara Research.** (2015). Impact evaluation report: EU empowerment project. Oxfam. Intervention.
- Scanteam.** (2008). Afghanistan Reconstruction Trust Fund: External evaluation. Final report. World Bank.
- Scanteam.** (2012). ARTF at a crossroads: History and the future. Final report. World Bank.
- Scanteam.** (2017). Taking charge: Government ownership in a complex context. External review. Afghanistan Reconstruction Trust Fund. World Bank
- Sida Decentralized Evaluation.** (2015). Review of Afghanistan Reconstruction Trust Fund, ARTF, internal and external studies and evaluations in Afghanistan — final report.
- Speakman, E. M., Shafi, A., Sondorp, E., Atta, N., & Howard, N.** (2014). Development of the community midwifery education initiative and its influence on women's health and empowerment in Afghanistan: A case study. *BioMed Central* 14, 111.
- Transtec.** (2013). Evaluation report: Citizens' platform for democratic debates and dialogues in Afghanistan. UDF-AFG-09-316. United Nations Democracy Fund.
- Transtec.** (2013). Evaluation report: Enhancing the capacity for inclusive local governance through synergies and sustainable linkages between communities and Government in North Afghanistan. UDF-AFG-08-249. UNDEF.
- Transtec.** (2014). Evaluation report: Raising awareness about women's social, political, and economic rights in Afghanistan. UDF-AFG-10-379. UNDEF.
- Transtec.** (2015). Evaluation report: Involving women and youth CSOs in strengthening democratic debate and public news media around elections in Afghanistan. UDF-AFG-12-508. United Nations Democracy Fund.
- Turkmani, S., Currie, S., Mungia, J., Assefi, N., Javed Rahmanzai, A., Azfar, P., & Bartlett, L.** (2013). Midwives are the backbone of our health system: Lessons from Afghanistan to guide expansion of midwifery in challenging settings. *Midwifery* 29, 1166–1172.
- UNDP.** (2014, July). Assessment of development results: Evaluation of UNDP contribution in Afghanistan. Independent Evaluation Office, United Nations Development Program.
- UNICEF.** (2015). Let us learn (LUL): Formative evaluation. UNICEF Afghanistan Country Office.
- Watkins, F., & Azarbaijani-Moghaddam, S.** (2019). Evaluation of UN women country program in Afghanistan: Final report. Sida Decentralized Evaluation.
- Zainullah, P., et al..** (2014). Establishing midwifery in low-resource settings: Guidance from a mixed-methods evaluation of the Afghanistan midwifery education program. *Midwifery* 30, 1056–1062.

Published by

Christoph M. Zürcher

Address

University of Ottawa
Graduate School of Public and International Affairs
120 University, Ottawa, ON K1N 6N5, Canada
T +1 613-562-5800-8997
F +1 613-562-5241

E christoph.zuercher@uottawa.ca
I www.sciencesociales.uottawa.ca

Authors

Reem Saraya and Christoph Zürcher, Ottawa, Canada

Design/Layout

Barbara Reuter, Oberursel, Germany
E barbarareuter-grafik@web.de

URL links

This publication contains links to external websites. Responsibility for the content of the listed external sites always lies with their respective publishers.

The present report was commissioned by the BMZ. The responsibility for the content lies solely with the publisher and does not reflect the views of the BMZ.

On behalf of

Federal Ministry for Economic Cooperation and Development (BMZ), Germany
Division for Afghanistan and Pakistan
Thomas Feidieker
Berlin, Germany
I www.bmz.de/en
I www.ez-afghanistan.de/en



Federal Ministry
for Economic Cooperation
and Development

March 2020

