



16th Evaluation Report 2019–2020

Evaluate. Measure. Learn.

KFW

Independent evaluation by the Financial Cooperation Evaluation Unit

Vision

We improve the effectiveness of Financial Cooperation (FC) in line with the 2030 Agenda and its Sustainable Development Goals (SDGs) by supporting FC projects with our evaluation expertise, by critically assessing them, and by identifying lessons learned for future projects.

Mission statement

We combine internal and external evaluation expertise to systematically assess the effects of completed FC projects and to support ongoing FC projects by means of impact evaluations. We enhance institutional learning by facilitating evaluation results in products that are tailored to the target group and can be used efficiently. We are committed to professional excellence, continuous skills enhancement and innovation.

About us

As an agile team of sector experts with scientific and/or operational experience, we see ourselves as a “knowledge hub” for evaluation and impact measurement within KfW Development Bank. We cooperate with the Federal Ministry for Economic Cooperation and Development (BMZ) and network with other national and international institutions dedicated to evaluating development cooperation. The involvement of local partners and their interests and learning objectives are important to us. Our work incorporates relevant standards of the evaluation community (DeGEval, OECD-DAC) and of empirical research, in particular with respect to the protection of personal rights and the anonymity of personal data as well as an ethical code of conduct.

»»» The Evaluation Report goes digital



This 16th Evaluation Report 2019–2020 is the first KfW FC Evaluation Report published as a digital version
www.kfw-entwicklungsbank.de/evbe2021_start

Foreword

Dear Readers

We can all feel how our world is changing: the rapid pace of digitalisation and changes in our working environments create tremendous opportunities, while climate change and the consequences of the coronavirus pandemic are central challenges for society as a whole. I am personally very pleased that we at KfW, together with the German Federal Government, can contribute to seizing these opportunities and tackling the challenges in Germany, but also around the world.

“Evaluate – Measure – Learn” is how Professor Jochen Kluge, who has been in charge of the Evaluation Unit at KfW Development Bank since October 2019, defines his mission. He and his team take stock of the independent Evaluation Unit’s work over the past two years in this 16th Evaluation Report. In the spirit of change and ongoing development, however, the report also focuses on the future and presents a vision for the team’s long-term evaluation mandate (page 10).

What is the idea behind “EVALUATE – MEASURE – LEARN”?

It denotes the core products of the Evaluation Unit: impact evaluations for ongoing projects, measurement of programme success for completed projects and institutional learning. Impact evaluations support selected projects from start to finish and develop tailored evaluation designs for measuring impacts. This provides us with scientifically based and deeper insights into the effectiveness of Financial Cooperation (FC).

Christiane Laibach
Member of KfW Group’s Executive Board



For example, if we identify the effects of a new water pipeline on the lives and health of people in rural Tanzania while it is still under construction (page 19), I think this creates an exciting new perspective.

Completed FC projects are assessed by means of ex post evaluations, which show us, based on the OECD-DAC criteria, how successful our work has been – or how much room there is for improvement – in terms of relevance, coherence, effectiveness, efficiency, impact and sustainability. And whether our projects are evaluated on the ground or, as has recently become increasingly common, “remotely” (see page 28), all evaluations require the best and most innovative data possible, as an afforestation project in Vietnam shows: the geodata analysis used in the ex post evaluation tracks the increase in forest cover in detail. The challenges of using satellite and open data are explained on page 34, also showing that networking with other departments at KfW Development Bank is the key to success, which I am particularly pleased about as a member of the Executive Board.

From the same point of view, I also ask myself, what is the value of great data and studies if they do not contribute to improving new projects and institutional learning? In this context, I very much welcome the fact that the Evaluation Unit is systematically channeling its evaluation knowledge back into KfW Development Bank; one particularly innovative and unique feature is the new digital tool QUER (Quick Evaluation Results), which enables project managers to interactively access the Unit’s more than 1000 evaluation results from the last decades:

for instance, if you enter “biodiversity” as a search term, you will find all the relevant evaluations with their lessons learned, ratings and risks. Or you can filter by region, sector and rating, and see at the touch of a button, for example, all projects in the energy sector in West Africa that have been rated as “successful” (or not), and can thus improve new projects (page 36). In addition, KfW Development Bank’s operational staff also benefit from personally conducting ex post evaluations. They report on their exciting experiences as they “view the world through lenses of evaluation” and their travels to project countries on page 6.

The Evaluation Unit therefore continues to rely on a healthy mix of continuity and innovation, which is also reflected in the format of this publication. The 16th Evaluation Report continues a long and important tradition of knowledge transfer. At the same time, it is the first evaluation report that goes digital with exciting interactive and audiovisual content. I look forward to exploring the new and old world of our Evaluation Unit with you.

Christiane Laibach

Effectiveness of Financial Cooperation

We are an independent evaluation unit that uses renowned and state-of-the-art methods in three pillars of evaluation work:

Ex post evaluations

- Every year, we draw a representative random sample of completed FC projects, stratified by sector, for ex post evaluation (EPE).
- We conduct EPEs ourselves and work together with external experts and delegates from the operational units of FC.
- EPEs have been carried out systematically by the FC Evaluation Unit since 2000. They promote accountability and provide the key knowledge base for institutional learning.
- Ex post evaluations assess project success using the OECD-DAC criteria.
- EPEs are the main pillar of our evaluation work.

Impact evaluations

- We support ongoing FC projects by developing and implementing customised impact evaluation designs.
- We answer questions that are of particular relevance to FC.
- We use modern methods of rigorous impact evaluation (RIE) and customize the design from the RIE toolbox that is appropriate for the respective project (“form follows function”).
- We seek to collect and use data in innovative ways, especially satellite data.
- We cooperate with suitable partners on a case-by-case basis, especially research institutions and other development banks.

Institutional learning

- We are guided by the interests and learning objectives of our target groups and provide up-to-date information in modern formats.
- We offer topical learning formats, organise training on evaluation methods and contribute our knowledge to peer discussions and sector retreats.
- Our colleagues in operations actively carry out EPEs together with us.
- We report on new results via newsletters and publication series.
- We provide the content of all ex post evaluations since 2007 in digital and interactive form within an app. (**QUER** – **QU**ick **E**valuation **R**esults).



Contents



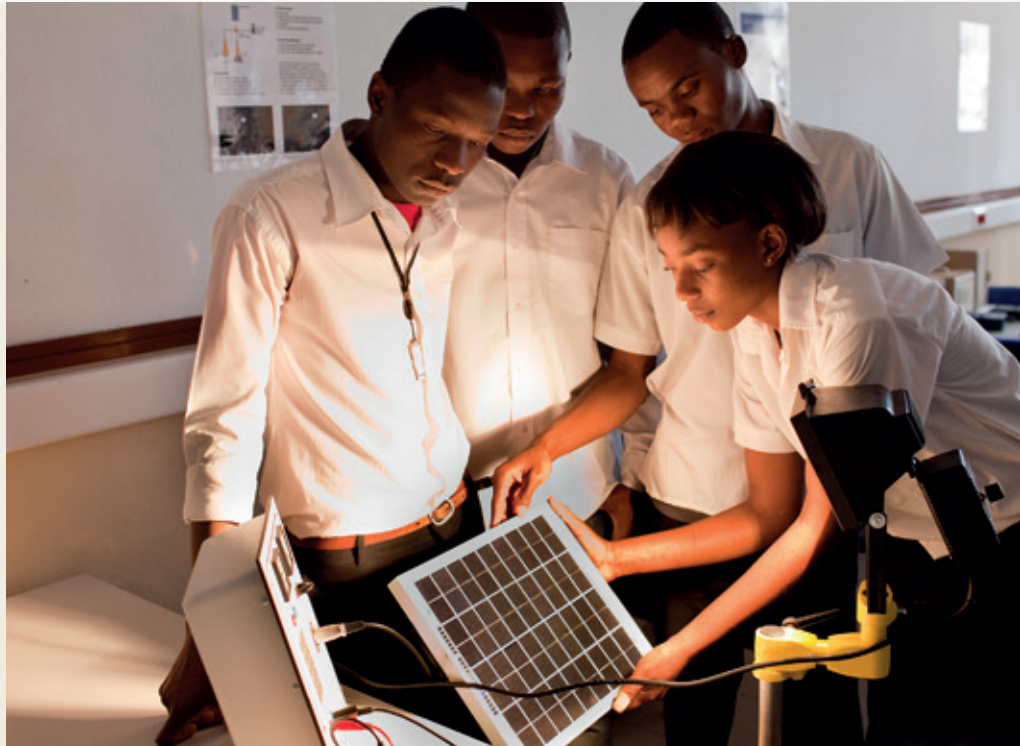
EVALUATE — evaluation of ongoing FC projects

- 06** Evaluations from around the world
- 10** Interview with Professor Jochen Kluge
- 12** Rigorous impact evaluation in development cooperation
- 16** Evaluation example: food security in Burkina Faso
- 18** Evaluation example: employment measures in Yemen
- 19** Evaluation example: water supply in Tanzania



MEASURE — evaluation of completed FC projects

- 22** Changing perspectives – a stroll through sectors and regions
- 28** Impacts 2.0 – improving measurement
- 29** Project example: forest development in Vietnam
- 30** Project example: electrification in Cambodia
- 31** Project example: irrigation measures in Kenya



LEARN — improving future projects

- 34** MAPME – planning, monitoring and evaluation
- 36** QUER – the interactive app for evaluation results
- 38** COVID-19 and employment



Methodology — Ex post evaluations 2019–2020

- 42** Working practices and methodology in the FC Evaluation
- 44** Sample size and distribution of ratings
- 46** Ex post evaluations – online
- 47** Imprint / photo details

Evaluations from around the world

From ecological forest conversion in China to small businesses in Tunisia and a microfinance line in Paraguay – the evaluations of Financial Cooperation (FC) projects are spread across the globe. To encourage learning from evaluations for current and future FC projects, employees at KfW Development Bank carry out evaluations on behalf of the independent Evaluation Unit – provided that they had no prior involvement with the project under evaluation to ensure an independent perspective. Five employees share their evaluation experience below:



More information is available online with a virtual journey through the projects!

Annika Schulte

Paraguay – Microfinance for microenterprises



The change in perspective was extremely interesting for me. Seeing where the development challenges are in the financial sector in Paraguay compared to the development challenges in the very same sector in the Middle East, which measures can solve them – that was a very important learning experience for me. Working with the Evaluation Unit and the regional department was a lot of fun. The responsible KfW office in Bolivia also provided me with excellent support.



Verena Quesnel

Tunisia – Promoting employment

I thought it was very exciting, not just to meet the partner institutions, but also Tunisian small business owners and to see what has changed for them – also thanks to the loans. The wide range of profiles of these companies was also interesting, as was the optimism expressed by the entrepreneurs. I would recommend carrying out an evaluation to everyone in FC. You learn so much about KfW's work and gain a lot personally.



Alexander Ehlert

Tajikistan – Expanding sustainable loan business

It was particularly exciting to broaden my own horizons and see how development is organised in other countries – it was challenging, but it taught me a lot for my own work.

Our partners welcomed us with open arms. Feedback was given mutually, which I found very helpful. One of the things I will remember was how appreciative and happy the people were who could afford their own house through the project.



Dirk Wenzel

China – Ecological landscape restoration

It has once again become clear that, especially in environmental projects, long-term support is essential, as is the case here with the ecological conversion of the forest.

Lea Stuff

China – Ecological landscape restoration



One key factor in the success of the project was a coherent political message behind it. It was interesting to see that environmental issues like biodiversity were linked to the government's priority issue of water supply for the Greater Beijing Region.



Involved from the beginning – the start of the project also marks the start of the impact evaluation. In close cooperation with the partners the project's impacts are analysed.

»» EVALUATE

Evaluation of ongoing FC projects



“An absolute highlight is our interactive QUER app which contains the evaluation knowledge of FC.”

Interview

Prof. Dr Jochen Kluge assumed the position as head of the FC Evaluation Unit at KfW Development Bank in October 2019. He has been a professor of economics at Humboldt University of Berlin since 2011. In this interview, he talks about his initial impressions, the challenges posed by the coronavirus and his vision for evaluation work in ten years' time.

Mr Kluge, you have now worked at KfW Development Bank for about a year and a half. How was your start?

Kluge: It has been a great start, in at least three respects. I inherited an established Evaluation Unit from my predecessor Eva Terberger, which has been systematically generating evaluation knowledge as an independent unit for more than 20 years now – we celebrated our anniversary last year. This is remarkable. I also joined a great, agile and committed team in the FC Evaluation Unit. And I met people who are just as dedicated throughout the entire development bank. I was and still am very enthusiastic. And of course there was also a steep learning curve in practical FC.

What prompted you to swap your position as professor at Humboldt University of Berlin for the head of the Evaluation Unit?

Kluge: I found the prospect of moving closer to practice – well, actually going straight into practice – extremely appealing. Researchers are committed to generating the best possible knowledge. This is essential. But the question is: what do we do with the knowledge? How can we use and apply it? How do we transfer evaluation findings into practice? I wanted to tackle these issues, and we have already made a lot of progress in the last year and a half. An absolute highlight is our interactive QUER app with the pooled evaluation knowledge of FC. I have always dreamed of having such a practical tool. And I haven't completely

given up academia, I will continue to do some research and teaching as part of my professorship.

What was the biggest surprise for you when you came to KfW?

Kluge: I am not sure if “surprise” is the right word. I found it remarkable, and still do, how committed all of my colleagues at the development bank are to the cause – to good development cooperation – at all levels. I am very happy to work in such a professional and open-minded environment.

The Federal Ministry for Economic Cooperation and Development (BMZ) adopted the 2030 Agenda just as you assumed your new position. How do you think the agenda will affect evaluation activities?

Kluge: Let me highlight one aspect. The 2030 Agenda emphasises the issue of impact, of measuring development effectiveness. This is important for BMZ, and this is exactly what we address as a modern Evaluation Unit in our trio of activities: firstly, we continue to carry out ex post evaluations. We have been systematically conducting ex post evaluations for more than 20 years. This is our core business and provides representative analyses of the impact and sustainability of FC. It is certainly the largest part of our work in terms of quantity. Secondly, we conduct impact evaluations that provide selected ongoing projects with thorough knowledge on their impacts –

including intermediary results that can feed directly into project implementation. Here, we also want to implement methodologically sophisticated evaluation designs – known as rigorous impact evaluations (RIE) – though always catering to the questions most relevant to the projects. Hence, following the principle “form follows function”, we ask: what do I want to know, and what method is best suited to answer that question? And our third activity is institutional learning. The aim here is to process all the findings from our first two fields of activity in a way that is useful and easily accessible for the users inside and outside of KfW Development Bank..

You mentioned that your team can only conduct a selected, i.e. limited, number of RIEs. How do you select projects for RIEs?

Kluge: We are particularly interested in questions with high operational relevance, i.e. where colleagues request support for impact measurement. This is particularly appropriate for projects in the design phase, as the methodology and data collection can be tailored to the impact questions of the project, and baseline data can also be collected. Some examples from Yemen, Tanzania and Burkina Faso are presented in this evaluation report. The examples also show that the FC Evaluation Unit supports projects over several years – and of course we can only do this for a limited number of projects;

in these evaluations we make a special effort to cooperate with other development banks – especially the World Bank – and universities in Germany and in partner countries.

And how exactly does an RIE work: what do project managers need if they want to conduct an RIE?

Kluge: Initially they just need interest. Then we sit down together and consider whether to set up an RIE and, if so, how. Of course, opportunities for cooperation and financial resources also play a role here, but there are many different case-specific solutions, for example, the integration of data collection into ongoing monitoring activities, the use of satellite data or cooperation with external partners.

The coronavirus pandemic is an exceptional crisis of global scale. You had to manage this crisis right at the beginning of your tenure at the development bank. How has the way you work changed?

Kluge: I wouldn't dwell too much on this issue here, other sectors of the economy and employees working in healthcare or nursing are feeling the effects of this extreme crisis much more intensely. However, in our actual evaluation work, of course, we have to work more “remotely”, which means performing desk reviews based on documents or using satellite

data. The latter definitely spurred innovation here, both in the evaluation and the operational departments we cooperate with in-house – there are some great initiatives. Many remote technologies are quite advanced and improving constantly, and it is now also possible to carry out technical assessments of infrastructure projects through video calls and similar means. Still, it remains the case that some development cooperation projects cannot really be adequately evaluated without an on-site visit. The personal impressions and ability to elicit on-site data are missing – but above all, remote assessments lack contact with the people who have benefited from the projects or implemented them.

How do you think evaluation work will look in the next ten years?

Kluge: Certainly similar to today in many ways. Innovative evaluation designs are still best created on the ground. But high-resolution satellite data will enable us to draw conclusions about many more projects and many more indicators than they already do today. Mobile technology is making on-site data collection increasingly easy. And databases and apps with evaluation knowledge like our QUER tool – and I think we are trailblazers with this tool – will be more robust and probably more widespread. This will certainly be the direction of the FC Evaluation Unit's work.

An on-site visit is essential – Jochen Kluge received important impressions during an evaluation field trip in El Salvador.



“The examples also show that the FC Evaluation Unit supports projects over several years.”



More personal insights as well as examples of impact evaluations can be found online.

Rigorous impact evaluations – more than just art for art’s sake

Researchers have used rigorous impact evaluations (RIE) as an important tool for impact assessment for nearly 20 years. Rigorous evaluation approaches are also gaining ground in international organisations and development banks. But what are they exactly? What are the advantages and disadvantages, and what role do RIEs play at KfW Development Bank?

Since the turn of the century, rigorous impact evaluations (RIEs) have become increasingly common in international development cooperation (DC). Inspired by the scientific community, they are now an integral part of many projects in international cooperation. The World Bank, for instance, has a Development Impact Evaluation (DIME) unit. The World Food Programme – winner of the 2020 Nobel Peace Prize – has been pursuing the WFP Impact Evaluation Strategy since 2019, while the International Initiative for Impact Evaluation (3ie) has been supporting and synthesising rigorous evidence from development projects since 2008. This led to a massive increase in the absolute number of RIEs carried out in development cooperation. While only about 50 RIEs of projects or policies in countries of the Global South were published worldwide by 2000, the next 15 years saw a boom with more than 4,000 RIEs.¹

The growth of RIEs was fuelled by the convergence of two trends. On the one hand, since the turn of the century political actors have actively pursued a stronger focus on impact in development cooperation. This was manifested in the Millennium Development Goals (MDGs) and in the Aid Effectiveness Agenda, in which the Federal Ministry for Economic Cooperation and Development (BMZ) played a significant role. On the other hand, researchers’ interest in analysing the causes of poverty and especially in possible mechanisms to alleviate it grew. They increasingly used improved statistical and econometric methods to evaluate projects. Researchers began to apply experimental methods – which were already common in natural science and medical research – to questions of development economics.

The Royal Swedish Academy of Sciences awarded these efforts in 2019 with the Prize in Economic Sciences in Memory of Alfred Nobel to development economists Abhijit Banerjee, Esther Duflo and Michael Kremer. The prize committee commented about the award: “Millions of people today benefit from effective interventions developed and tested with the new experimental approach for which they [the laureates] have laid the foundation.”²

¹ Source: Sabet, S.M. and Brown, A.N. (2018). Is impact evaluation still on the rise? The new trends in 2010–2015. Journal of Development Effectiveness, 10(3): 291-304.
² Source: Royal Swedish Academy of Sciences (2019). <https://www.nobelprize.org/prizes/economic-sciences/2019/press-release/>.



Disbursement of a cash transfer in Malawi – FC has been supporting the poorest families in Malawi since 2006 within the framework of Social Cash Transfer Programmes.

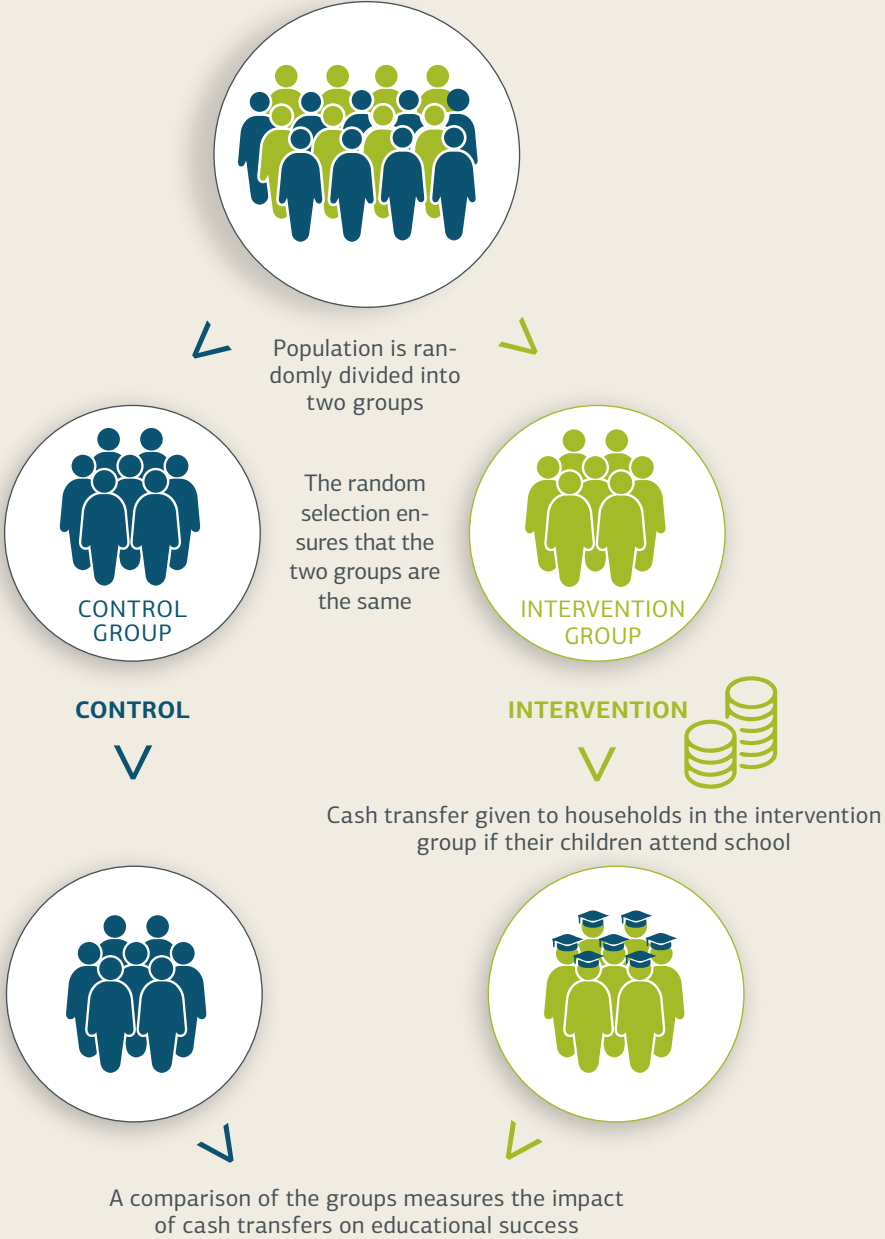
Why evaluate?

There are many good reasons for performing robust evaluations. Among the most important is accountability to the public and civil society in our partner countries as well as in Germany. Evaluations make it possible to identify particularly effective approaches, to modify them early on if necessary, and to quantify the cost-effectiveness of a project. Moreover, reliable impact measurements enable institutional learning. Finally, the findings contribute to external learning and the global evidence base.

Evaluation is particularly important in the context of global DC: on the one hand, DC projects do not have to compete the way private companies do. Traditional market mechanisms, such as bankruptcy when companies are poorly managed or crowded out by better products, do not exist. On the other hand, limited financial resources stand in contrast to a large number of urgently needed investments. A solid understanding of effectiveness is therefore extremely important.

RIEs are only one of several methods of evaluation or monitoring. For example, KfW Development Bank has successfully employed ex post evaluations since 1990 to systematically observe and assess projects as a whole and over time. However, if one is particularly interested in effects at the impact level, the most rigorous way to measure them is – as the name suggests – by means of an RIE.

Design of a randomised controlled trial (RCT) to test whether a cash transfer increases educational success



What exactly are rigorous impact evaluations?

RIEs describe a toolbox of experimental and semi-experimental methods that measure the causal effects of a project. The emphasis is on causality. In other words, on identifying those effects that can be attributed exclusively to the project and isolating them from concurrent developments or other connections between projects and target indicators. In addition to measuring specific impacts on the projects’ target groups, RIEs also analyse impacts on subgroups or mechanisms underlying the impacts. For example, a healthcare project may have significantly greater effects for women than for men, or a new connection to the electrical grid may only lead to productive electricity uses in areas that have access to markets.

The most rigorous methods in the IE toolbox are fully experimental methods, such as randomised controlled trials (RCTs) – which are also known as the “gold standard”. In RCTs, a project – or even parts of the project – is randomly assigned to a group of individuals, schools, communities or other (“intervention group”). The second group receives access to the project later, as described in the example of Yemen on page 18, or – as is the case with a placebo – not at all (control group). The principle of (controlled) random assignment, similar to medical research, ensures the comparability of the two groups: for example depending on the measure, they are on average the same age, similarly healthy, ambitious, vulnerable or wealthy. This means that all post-intervention differences between the groups can be attributed to the project itself. A well-known example is cash transfers, which are disbursed to households in the target group if their children attend school.

If a purely experimental (random) assignment is not reasonable or feasible, semi-experimental methods are often a useful alternative. For example, comparison groups can be defined along threshold values of certain selection criteria (Regression Discontinuity Design, RDD). If a project targets children under two years of age, participants who are almost two years old can be compared with participants who are just over two years old, as the example from Burkina Faso shows on the following pages.

RCTs and RDDs are only two examples from the IE toolbox. Depending on the type of project, the level of implementation and the criteria for selecting beneficiaries, the toolbox provides a range of methodological options. One thing is certain, however: the earlier an impact evaluation is integrated into a project’s implementation, the more likely it is that reliable conclusions can be drawn about its impacts. Collecting data before the start of the project (baseline), for example, can greatly improve evaluations. These lessons learned can then be transferred to similar or follow-up projects to increase their effectiveness.





See what the local water supplier has learned and implemented following the impact analysis of Sebastian Tonke.

A water project in Pristina, Kosovo, shows that, in addition to the questions commonly addressed in an impact evaluation, it can also be worthwhile to evaluate behaviour. The project aimed to build effective structures for water supply and sanitation. The goal was to improve the drinking water supply and, with it, living conditions. To examine the payment behaviour of customers, so-called “nudges”, i.e. incentives to change behaviour, were applied and tested for their impact. These incentives included, for example, attaching the bill to the front door (instead of leaving it in the letterbox) or sending letters appealing to the customers’ sense of responsibility. The different incentives were randomly assigned. Depending on the type of incentive and the wording of the message, on-time payments increased by up to 62 %. According to the responsible water supplier, the approaches were adopted beyond the originally planned two-month period. This evaluation example shows how relevant results can be achieved without spending a lot of time and money, and how valuable it can be to test new and creative approaches using randomised methods.



RIEs also have critics, and debates about the pros and cons have been passionately waged for years. The criticism includes:

- Ethical reservations: participation in projects is not assigned based on needs but based on randomisation. This criticism is valid and important. The right to participate in a project must always follow fair and reasonable criteria. However, RCTs can (and must) adhere to high ethical standards, for example, if their design exploits regional, budgetary or time limits.
- Results of RIEs do not necessarily generalise across contexts, populations or timeframes: this criticism applies, as it does to any other evaluation method of individual projects. Existing possibilities to increase generalisability must therefore be fully exploited in the implementation of RIEs, and the transferability to other projects must be scrutinised on a case-by-case basis. By the way, an increasing number of meta-analyses and systematic reviews of RIEs are seeking to reduce this hurdle.
- RIEs are not suitable for all projects; even if an RIE can theoretically be carried out for every project, it is not always the most expedient method. It is therefore important to weigh the advantages and disadvantages of various evaluation methods on a case-by-case basis.

RIE at KfW Development Bank

KfW Development Bank’s Evaluation Unit increasingly provides institutional and methodological knowledge to support implementation of RIEs. On the next few pages of this report, you can find impressions of KfW Development Bank’s evaluation designs. The Evaluation Unit adapts the use of RIEs – taking into account the methodological possibilities and limits consistent with the principle of “form follows function” – to the relevant content-specific question, the context and, the needs and capacities of its partners. Depending on needs, households can be surveyed, or analyses conducted with satellite or other secondary data. Ideally, RIEs are implemented in cooperation with other development banks such as the World Bank or the Agence française de développement as well as local or academic partners. This allows synergies in learning, both between development banks and between partners.

Experiences with experimental evaluations at KfW Development Bank

“Since 2005, KfW’s multisectoral Reintegration and Reconstruction programme has sought to improve living conditions in Liberia and contribute to consolidating the ongoing peace process. The programme is being carried out in cooperation with Deutsche Welthungerhilfe and other non-governmental organisations (NGOs).

When preparing the fifth programme phase, our team took the opportunity to initiate a rigorous impact evaluation in the form of a Randomised Controlled Trial (RCT). Our aim was to verify the impact logic, understand causal relationships and measure the actual impact of the project. We also wanted to better understand the implications of specific implementation aspects in order to incorporate them into the design of follow-up projects and ultimately achieve greater effectiveness.

The RCT is being conducted by external researchers in cooperation with the implementing NGO. The initial results of the RCT already offer exciting lessons learned for an effective continuation of the programme. For example, the RCT has shown that, despite the strong role of NGOs in project implementation, more trust can be placed in the government. We are already learning a lot about the impacts of our project on the social, health and economic situation of the programme participants. Our experience so far motivates me to continue conducting impact evaluations in the future – whenever possible.”

Alina Sennewald
“Governance” portfolio manager in the West Africa region

Burkina Faso –

the first 1,000 days count for a lifetime



Many children in Burkina Faso suffer from malnutrition in the first years of their lives. Yet proper nutrition is never as important as at the beginning of life. In this project in Burkina Faso’s south-western province of loba, young mothers receive transfer payments and training. The goal: Reduce malnutrition in newborns and young children, thereby setting a positive course for the future in the long term. The impact evaluation analyses the long-term impacts of these payments on mothers, children and the village communities.

Burkina Faso is one of the poorest countries in the world: the Human Development Index ranks the country 182 out of 189 nations. Intensified by climate change, regular droughts and land degradation pose significant challenges to the food security of the growing population. This applies in particular to rural areas and to pregnant women and small children. Particularly for children, nutrition is crucial in the first 1,000 days of life. Malnutrition and undernourishment in those days have a lasting impact on children’s cognitive development and physical health. According to data from the United Nations Development Programme (UNDP), 25 % of children under five in Burkina Faso suffer from impairments to their growth.

The project aims to improve the food security of as many as 15,000 mothers and mothers-to-be and their 18,000 infants in the province of loba in south-western Burkina Faso

The women receive quarterly payments for three years. Mothers of children under the age of two and mothers-to-be are eligible to participate. To ensure reliable transfers via mobile money, the women receive assistance to apply for ID cards and are provided with a mobile phone. These measures foster women’s financial inclusion and empower them to assert their economic and political rights. Mothers-to-be are also invited to participate in educational campaigns about nutrition, hygiene and health. The campaigns are conducted in local women’s groups, but also presented in movies or theatre performances and on the radio.

The project primarily focuses on participating women and their children, but also on their communities and local economies. The primary goal of the impact evaluation is to capture the – intended and unintended – causal impacts. To this end, both qualitative and quantitative data are collected at the level of the women, their children, communities, markets and health centres. The data are collected before the start of the project (baseline), one year later and three years later. The analyses provide reliable findings for possibly expanding the project to other regions. It allows to answer the following questions: did the payments reach the women, and did the targeting criteria actually work?

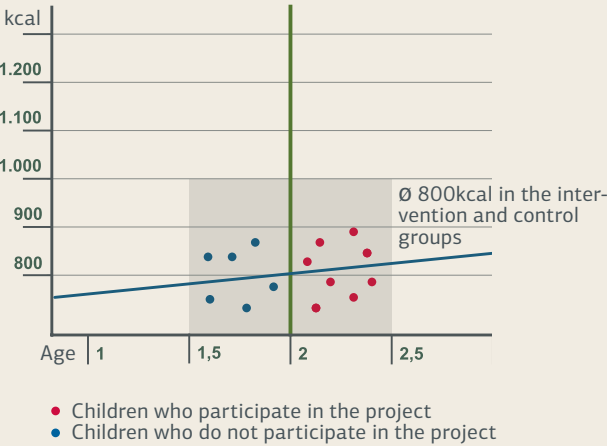


The first 1,000 days count – the effects of malnutrition in the first years of life have consequences that last a lifetime: the RIE measures the extent to which cash transfers can help in Burkina Faso.

What have the women learned about food security and nutritional practices, and how does this affect their behaviour? Has participating in the project improved the physical and mental health of the women and children – for example, have growth impairments among children been prevented? How has the project been received by families and communities, and how does it affect domestic violence and neighbourhood conflicts? Have there been negative impacts – for example on food prices – due to changes in demand patterns? Do impacts that are visible in the short term still persist after three years?

A “regression discontinuity approach” is used to answer all these questions

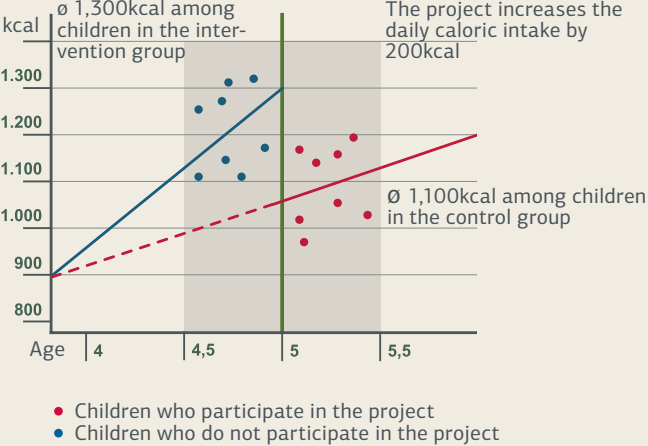
Let’s assume before the start of the project, the average caloric intake of children aged one to three years is 800kcal (kilocalories). Households with children under two years of age now receive transfer payments, while households with children over two years of age do not. This clear targeting criterion can be used



Measurement prior to the start of the project

to identify the payments’ impacts. It is assumed that mothers and their children between the ages of one and a half and two can be compared to children between the ages of two and two and a half. The households are surveyed before the project begins in order to obtain initial baseline information about these children and the households.

At the end of the project, the children are three years older. This means that mothers and their children aged four and a half to five years are compared with mothers and their children aged five to five and a half years. The households are surveyed again to collect this data. Taking into account the results of the previous baseline survey, the impact of the measure can be quantified in the fictitious example as follows: the children’s caloric intake increases with age and is 1,100kcal for children in the comparison group. For children participating in the project, the value also increases, but to 1,300kcal. This is 200kcal, or about 18 % more than for children who are not part of the project.



Measurement at the end of the project – 3 years later

Cash-for-Work –

a rigorous analysis of the impacts in Yemen



The Labor Intensive Works Programme is a programme run by the Yemeni Social Fund for Development (SFD) to support the poorest population groups in crisis-ridden Yemen. The country has suffered from armed conflicts for decades; a lack of job and income opportunities as well as weak basic infrastructure further increase the vulnerability of the population. The impact evaluation examines whether and to what extent the programme contributes to increasing the resilience of the beneficiaries and thus to improving their living conditions in the long term.

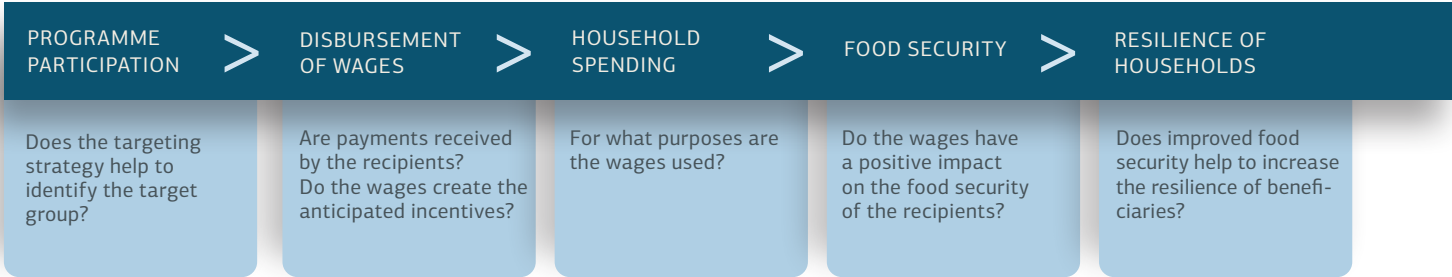
The Cash-for-Work programme constructs and maintains basic infrastructure and simultaneously addresses several bottlenecks: the additional income generated is intended to facilitate project participants' access to everyday needs such as food and medicine, as well as education. The new infrastructure measures aim to directly and indirectly improve living conditions within communities and to thereby reduce conflicts.

But do these expected impacts really materialise? There are only a few recent studies on the impacts of this kind of cash-for-work measure from Yemen or other countries in the Middle East.¹ This evaluation is therefore designed as a rigorous impact evaluation (RIE) and is intended to measure the impacts of the project, incorporate them

into the literature and generate evidence-based conclusions for the future.

The evaluation attempts to answer a wide range of questions: Does the income of vulnerable rural households increase? Does the additional income improve food security, and does this contribute to strengthening resilience? Does greater resilience improve future job prospects? Do the infrastructure measures contribute to a better quality of life?

The graph below broadly shows how the evaluation analyses the programme's Theory of Change using the main impact variable "resilience of beneficiaries" as an example; beneficiaries of the programme build the infrastructure and receive wages for their work. Once this payment is disbursed, the next step is to analyse the impact of the wages on household spending. If the wages are spent on essential goods such as basic staples, one can evaluate whether the increased spending on basic staples has a positive impact on the food security of the beneficiaries. If food security has been positively influenced, the impact of the project on increasing resilience can be determined in the final step.



Impact chain of the Labor-Intensive Works Programme (LIWP)



Cash-for-work – the Cash-for-Work Program aims to build infrastructure and improve food security in Yemen. How effective is it?

¹ Vgl. Lombardini, S., und Mager, F. (2019). Livelihoods in the Za'atari Camp: Impact evaluation of Oxfam's Cash for Work activities in the Za'atari camp (Jordan), Oxfam Policy & Practice.

Simiyu Climate Resilience in Tanzania –

an impact evaluation



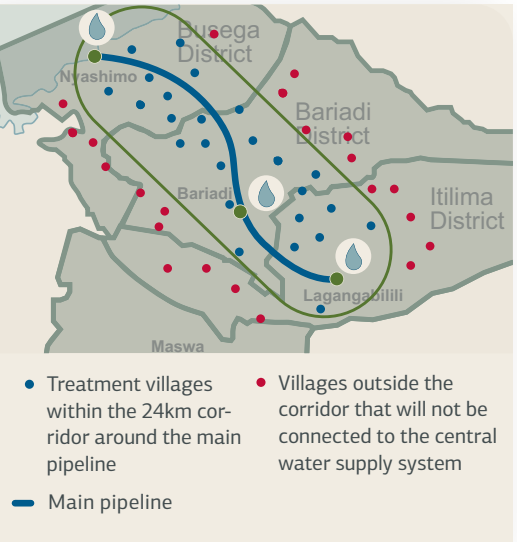
The Simiyu Climate Resilience project is cofinanced by the Green Climate Fund and promotes clean water supply, access to sanitation and climate-smart agriculture. The aim is to strengthen the resilience of the population in the Simiyu Region in northern Tanzania towards the impacts of climate change. But how can changes in the population be measured and analysed?

The Simiyu Climate Resilience Project supports the Tanzanian government in its aim to provide 90 % of the rural population with enough water by 2025. Supply structures will be created in three components – water supply, sanitation, agriculture – which will also strengthen the resilience of the population towards the impacts of climate change. The largest component of the project is the construction of an approximately 100km-long water pipeline from Lake Victoria to the country's interior. Connection to the supply system is intended to ensure year-round, reliable and safe access to clean water for villages within a 24km corridor around the main pipeline.

Construction of the water supply system is scheduled to start sometime in 2022. The construction schedule enables the impact evaluation to methodically measure the impacts of the project on the population in the Simiyu Region in the coming years:

The evaluation compares the water access of the households connected to the system within the corridor with that of the households outside the corridor. The latter cannot be connected to the pipeline. This makes it possible to determine the causal effect of the project on water access and people's health. As part of the impact evaluation, the intent is also to analyse the change in the population's water access during pipeline construction in order to already generate interim results during project implementation.

Shorter distances – at the moment, children and women in particular still have to walk long distances to fetch water from Lake Victoria (and other sources). This will change once the main pipeline is installed.





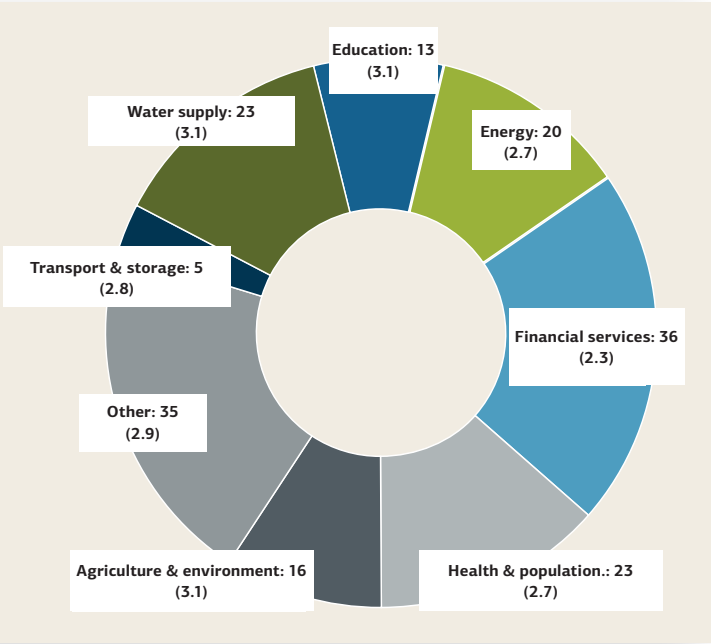
Examining our impact –
using new sources of data to collect
important information for project
evaluation.

»» MEASURE

Evaluation of completed FC projects

Changing perspectives – a stroll through sectors and regions

Which were the most important findings from evaluations in 2019/2020, and what are the most important lessons learned? On this stroll, you will gain insight into the wide variety of Financial Cooperation (FC) projects evaluated and the unique features of individual projects.



Number of projects evaluated by sector with average rating in brackets. Rating ranges from 1 (best) to 6 (lowest rating).



The financial sector comes out on top – This female Jordan entrepreneur used an FC credit to successfully start her own company.

Financial sector plays a prominent role with above-average success

Making up 36 of the overall 171 projects (21 %), the financial sector was subject to the most ex post evaluations in 2019/2020. Ten of the 36 financial projects addressed micro-, small- and medium-sized enterprises (MSMEs) as their target group, while a further eight projects were designed to increase energy efficiency in businesses. The comparatively large number of evaluations is closely linked to the Evaluation Unit’s pooling of evaluations of multi-phase projects; as soon as one project phase is included in the random sample from the FC portfolio, preceding phases are as well evaluated. This approach allows for a more comprehensive assessment of the overall project’s effectiveness. For example, six phases of the regional energy efficiency fund in south-east Europe and five phases of the MENA regional

fund for MSME financing were evaluated within one ex post evaluation report (EPE). When compared to other sectors, projects from the financial sector perform particularly well. On average, financial projects received an overall rating of 2.3, while the average rating across all sectors was 2.8.

There are a number of reasons why the financial sector scored so highly. On the one hand, major hurdles often have to be surmounted to implement these projects. For instance, stable regulatory conditions are needed to establish and disburse a financial product. This is one of the reasons why there are relatively few financial projects in Sub-Saharan Africa included in the overall random sample, while more projects are implemented in more politically stable regions, such as central and south-east Europe, and, as a result, more projects from these regions are evaluated. On the other hand, when projects are pooled, evaluations often detect an improvement to the rating: projects with a very good or good rating in the initial phases were also able to achieve at least the same rating in later phases.

Rising number of projects in fragile contexts

A high number of global conflicts results in more projects being implemented in conflict countries. This trend is also increasingly reflected in the evaluations. Almost ten years ago (2011/2012 period), 38 FC projects in fragile contexts were evaluated, a figure which has now almost doubled to 68. These projects are often executed in several phases. The education sector is taking on a particularly important role within the portfolio of fragile context projects. One example of this is the multi-phase EQUIP II education project in Afghanistan. This project has been extended three times and aims to improve primary education and contribute to the alleviation of poverty through the increased utilisation of educational opportunities. Not least due to the fragile circumstances in Afghanistan, it became evident that several aspects of this project required adjustment. For instance, in addition

to measures to increase capacities in schools, contributions to improving the quality of education are also required. In the future, improving the adjustment of measures to changing local security situations is recommended. A positive factor in this project was the involvement of communities in the administration of schools. The school committees set up by EQUIP II looked after the safety of school children and played an important role in the reopening of schools that had been closed due to the conflict. The importance of strong local partners for project implementation, particularly in fragile contexts, was also demonstrated by the example of the Social Fund for Development in Yemen. In this case, the strong partner was one of the factors that led to the project’s regular extension, thereby helping to improve living conditions for young people and children in Yemen.

New experiences with policy-based lending

Policy-based lending is becoming increasingly important within the FC portfolio; four such projects were evaluated in 2019 and 2020. The central findings are that project ownership is the most important factor for success. Furthermore, experience has shown that policy dialogues – in which specific measures and aspects of reform financing are clarified and agreed upon – are extremely important but also take up a great deal of time and other resources. Sufficient capacities, including on-site presence, are therefore essential components for a credible, ongoing and trust-based dialogue with partners. The example of a residential water management project in Peru shows that the likelihood of achieving a project’s goals increases if they are developed in close liaison with the government. Instead of defining an ultimate objective, individual stages in a reform process are agreed upon, leading to funds being disbursed on a pro rata basis (trigger). The evaluations reveal that, while the current flexibility in defining triggers should be maintained, criteria need to be formulated more precisely. In Peru, for example, due to their rather vague definition, the agreed targets tended to be the basis of an ongoing political dialogue as opposed to an instrument for releasing disbursements following a trigger event.

Upward trend – the number of projects taking place in a fragile context continues to rise.



Evaluations in the health sector – indications on fighting COVID-19

Against the background of the coronavirus pandemic, findings from the healthcare sector are also of particular interest. Programmes for fighting infectious diseases may provide valuable insights into dealing with COVID-19, for example. The assessment of past evaluations revealed that consideration of cross-border infection dynamics plays an important role for the success of a project. Due to the high mobility of people, including their movement across borders, regional approaches are needed to supplement national ones. This was demonstrated in a programme to tackle tuberculosis in the Caucasus area, a region where people have a high level of mobility. Due to the heavy focus on purely national aspects, the approach was not sustainable, which is why the project was rated as "not successful" on the whole. Furthermore, too few

qualified staff were willing to expose themselves to the risk of getting infected – an important finding for future projects of this type. In contrast, multinational cooperation was an integral feature of a successful multi-phase HIV prevention project in the CEMAC countries (Central Africa) and helped to increase the project's efficiency with simultaneous positive impacts on the population's reproductive health.

The pandemic is not over yet – findings from evaluations can help the fight in developing countries.



New findings and methods in the green sector

Involving local populations in the declaration of ecological conservation areas increases their acceptance and is an important factor in the success of projects in the environmental sector. Thanks to its participative approaches, a project to promote sustainable agriculture in Burkina Faso was able to achieve a solid developmental impact (impact rating of 2). As well as involving the project's beneficiaries in the implementation, management and maintenance of the measures, the process also incorporated a research institute, government agencies, and private stakeholders, such as banks. A 2020 cross-sectional evaluation on conservation areas also highlighted the effectiveness of projects that combine measures for improving socio-economic conditions with nature conservation measures. The combination of income-generating measures and training programmes to promote behavioural change was identified as a further factor for success.

Satellite data have proven to be a particularly effective element for enhancing evaluations in the agriculture and environmental sectors. High-resolution satellite images of relevant variables, such as total area of conserved forests, can help identify deforestation or other developments at an early stage. The use of publicly available satellite data on forest coverage, for example, enabled the analysis of forest loss and growth as part of a forestry project in Vietnam. Satellite images revealed that, although reforestation measures were successful in restoring forests in some project sites, other parts of the intervention areas suffered from high deforestation activities, especially after project end. Consequently, the project was rated as achieving "only limited success".

During the evaluation of another project in Laos, it was possible to track the movement of so-called deforestation hotspots, which are posing an ever-increasing risk to the promoted conservation areas. For example, a new method for measuring forest fragmentation revealed the construction of a major road, an issue that had not been mentioned during the visit to the area. At the same time, however, it was also possible to trace the positive impacts that the community patrol models are having on the maintenance of conservation areas. On the whole, projects from the agriculture and environmental sectors have achieved a positive impact – the sector takes the top spot when it comes to average rating for the DAC criterion of impact.



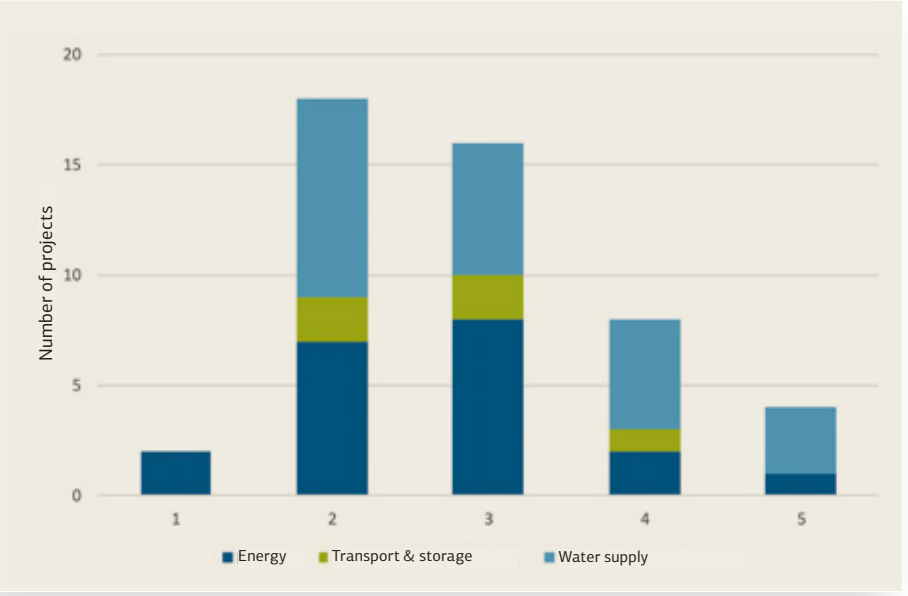
Important factor in the success of biodiversity projects – the combination of income-generating measures in conjunction with training programmes to promote behavioural change.

Heterogeneous results in the infrastructure sectors

In addition to the most frequently awarded overall ratings of 2, 3 and 4, on rare occasions particularly successful projects are awarded a rating of 1, while projects that are overwhelmingly disappointing can be rated as 5. In the 2019–2020 evaluation period, the only sector to contain projects receiving the entire range of possible overall ratings was the energy sector. Two projects in Cambodia related to rural electrification were able to exceed expectations. The construction of a new power line, a substation and a power distribution system led, among others, to lower electricity prices, meaning that the project significantly contributed to improving living conditions.

In the water sector, a rating of 5 was awarded to three projects. The "hazardous waste disposal sites" project in Tunisia confirmed an important finding: involving the local population – for instance, when selecting a project location – plays a decisive role in a project's success in this sector, too.

Over the past two years, the success of projects in the transport and storage sector was average, but once again emphasised the relevance of the DAC criterion of sustainability: despite some remarkable effects on traffic volume achieved by new infrastructure built to rectify flood damage in Cambodia, the project's sustainability is limited due to a lack of routine maintenance measures. And the importance of sustainability as a criterion is still growing. Securing maintenance is particularly important in projects that promote resilience to climate change among local populations.



Ratings in the infrastructure sector

Substation in Pakistan – improving living conditions.

From decentralisation to residential construction projects – the “other” sector



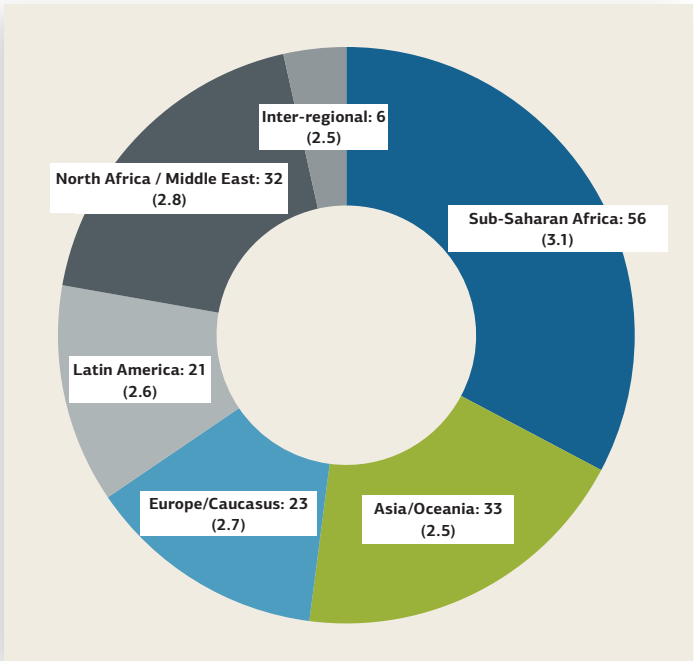
This sector contains projects without a specific overarching focus, including decentralisation projects, residential building projects and reintegration and recovery projects. One example of this type of project is the “Reintegration and Recovery Programme” (RPP) in Liberia, which was implemented during the Ebola epidemic. In this project, the possibility to deploy funds quickly enabled not only the implementation of the originally planned measures, but also construction of healthcare centers (overall rating: 3). A very positive, sustainable example in this sector is the construction of a very different type of infrastructure: Social housing projects in Honduras led to the establishment of a charitable foundation that constructs homes in rural areas without any further FC involvement (overall rating: 2).

This reporting period moreover mirrors FC’s decreasing involvement in decentralisation projects – only the first two phases of the FADeC project in Benin were evaluated in this area with an overall rating of 3. The project applied an innovative, incentive-based approach for transferring disbursements to municipalities.

Whether building homes or contributing to recovery – adapted solutions outside of the main sectors.

Sub-Saharan Africa: region with the most projects

As in the previous reporting period, Sub-Saharan Africa dominates the evaluation portfolio in terms of the number of projects (56 projects). In general, the regional distribution of FC activities also exhibits continuity over time. The only regions to have more projects evaluated in the 2019/2020 period are North Africa and the Middle East (32 vs. 15 projects in the previous period). On average, the overall rating awarded to projects in Sub-Saharan Africa is slightly lower than the overall rating across all regions. The hypothesis proposed in the 15th Evaluation Report, namely that the increasing number of projects in North Africa and the Middle East could have a negative impact on the overall result, has yet to be proven. As in previous reporting periods, projects in Asia/Oceania receive the best ratings. Over 90% of the projects are rated as “successful”.



Number of projects by region, average rating in brackets



Focus area within the evaluation portfolio – the majority of German FC projects are carried out in Sub-Saharan Africa.

New sources of data present new challenges to evaluation practices. At the same time, they allow for better measurement of impacts thanks to quantitative empirical analyses, including remote analyses. For example, satellite data can be used to record baseline information retrospectively, which improves the design of the evaluation. Combining traditional types of data with these new sources of data opens up new perspectives, increases the information content and improves the robustness of findings.



Verifying impacts with geodata – a cornerstone of modern evaluation work.

Impacts 2.0 – improving measurement

Evaluating impacts – applying new sources of data

Generating positive impacts and improving living conditions as a result is the most important objective of development cooperation yet also one of its biggest challenges. After all, even though every Financial Cooperation project leads to concrete changes, the resulting impacts are not always direct and obvious. An important function of the evaluation process is therefore to examine positive and unintended negative impacts and quantify these as reliably as possible. This task is covered by the impact criterion applied in ex post evaluations. Past ex post evaluations are primarily based on figures and information from internal project documents, supplemented by an in-depth study of relevant literature and interviews with as many local stakeholders as possible. This method is well suited to study as many different aspects and nuances of projects as possible.

However, these sources provide only limited in-depth, quantitative information to assess the impacts and sustainability of a project. For this reason, further secondary data and new data sources, such as satellites, are being used increasingly. These types of data can be a very useful source of supplementary information to enhance missing or inconsistent data during the project implementation phase and thereafter.

These approaches showcased their strengths in the context of coronavirus pandemic travel restrictions, enabling work to continue on a remote basis. In addition to satellite data, there are a number of other new (and old) data sources and forms of analysis. Over recent years, recording and application of these sources has become significantly easier thanks to the following developments:

- New information sources such as online surveys and open data portals are more and more available thanks to increasing digitalisation.
- The global spread of mobile phones is making it easier to collect primary data from target groups.
- Quick and, in some cases, free access to existing data, such as satellite data archives.
- The increasing automation of analyses and scalability allow wide-spread use at low incremental costs.
- The growing number of open source solutions and data communities creates networking effects and enables experience to be shared with others.
- A growing pool of executing agency data, which can also be used for evaluation purposes.
- Increasing availability of census data and other administrative statistics on the Internet.

Project example: forest development in Vietnam



Brief description of the project

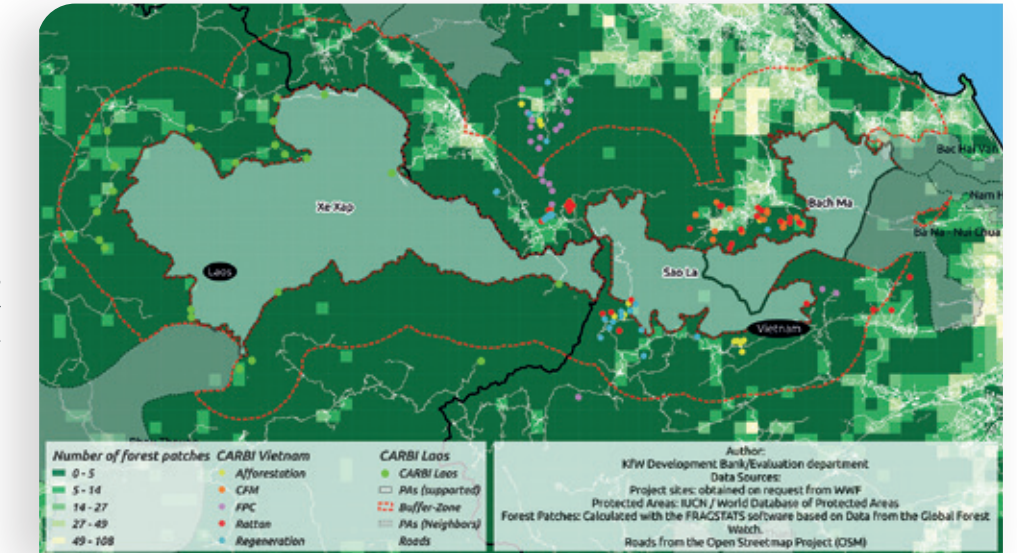
The projects involved the reforestation and sustainable management of state-owned areas of forest containing local species used for long-term timber trees. The goal was to restore ecologically degraded areas, including their key ecological role, while also improving living conditions of the local population.

Result

Household surveys revealed that water availability and quality had improved. Site visits confirmed these (subjective) appraisals to a great extent. A growing level of environmental awareness was also detected among the projects' beneficiaries.

The reforestation measures completed during the project led to the desired growth in forest within the project areas. However, geodata analyses helped to detect forest loss in other parts of the municipal forests. In terms of

Not just for evaluations, but for new measures, too – satellite data provides new insights.



area, forest loss exceeded growth. As a result, it became clear that, in addition to reforestation, land usage planning and controlled deforestation are also important to the sustainable recovery of ecosystems.

A unique method: geodata analysis

A variety of instruments were used to measure forest losses in this evaluation. Information from Global Forest Watch (GFW) was used to analyse large-scale forest loss; to detect early stages of forest degradation (e.g. removal of individual trees), data from the ESA's (Sentinel) mission for the Copernicus programme was evaluated with the help of the dNBR algorithm (delta Normalised Burn Ratio). Furthermore, forest loss and growth were classified on an object basis using Sentinel data. The Sentinel data combined with the dNBR algorithm proved to be a particularly precise method in this context.

All three methods suggest large-scale deforestation. This finding could be corroborated by comparison of findings from the three distinct data sources and methods. Quantifying forest growth and loss in this way would not have been possible without the use of satellite data.

In some cases, the results from the geodata analysis even indicated deforestation trends that contradicted the findings from the local visits. For example, information provided by the official authorities pointed towards a clear decline in illegal logging. However, the analysis of satellite images revealed deforestation and forest clearance that went beyond the intended scope. Without analysis of satellite data, this development would have remained undetected.

This evaluation clearly shows how new sources of data – satellite data in this case – can effectively support the evaluation process. The availability of suitable data for the relevant indicators plays a decisive role here.

Project example: power transmission and rural electrification in Cambodia



Electricity to boost energy and incomes – positive impacts of electricity supply on households income in Cambodia.

Brief description of the project

The construction of power lines and rural electrification were part of reconstruction processes in Cambodia after the first democratic elections. The goal was firstly to improve the transmission of power from Takeo to the capital city of Phnom Penh and secondly to bring electricity to selected rural areas. The measures included the construction of an electricity line and of a new substation and power distribution system. The project’s overarching goal was to reduce poverty and improve social and ecological sustainability.

Result

The measures led to significantly improved living conditions and simultaneously reduced the use of local diesel generators.

Energy prices have fallen in rural and urban areas, and the measure continued to have positive effects after its completion.

Unique sources of data

In addition to before-and-after comparisons and interviews in the field, secondary data was used to triangulate results for the evaluation of these projects. The secondary data included the project-executing agency’s annual report as well as the World Development Indicators.

Furthermore, it was possible to access household data collected in the region by the World Bank and use them for the evaluation. This evaluation revealed, for example, that the proportion of household expenditure spent on electricity costs was significantly lower in the project regions than in the rest of Cambodia.

The evaluation of this project shows that it is worth taking a broader view and that the collection and use of quantitative data from other donors improves the informative value of the evaluation.



Project example: irrigation at Mount Kenya



Group lending in farmland irrigation – a more diverse array of plants and more frequent harvests.

Brief description of the project

The project supported smallholders in the regions to the east/south-east of Mount Kenya in the transition from rain-dependent farming to irrigated agriculture. In addition to increasing agricultural production, the main goal of this multi-phase project was to improve the living conditions of rural households. The project made use of existing organised groups and cooperatives and issued loans according to a group lending principle – 50% of the measures were financed from grants, while the remaining measures were covered by loans to groups of smallholders.

Result

Thanks to the measures, it was possible to increase the area of irrigated farmland. According to household surveys, the option to use irrigation allowed for a diversification of agricultural produce. As a result, the smallholders were able to better adapt their production to market demand and, in some cases, switch to cash crops, i.e. agricultural produce with higher margins.

The project measures did not result in any significant change in biomass production. This developed similarly in areas that were structurally comparable to the project regions but that were not targeted by the project.

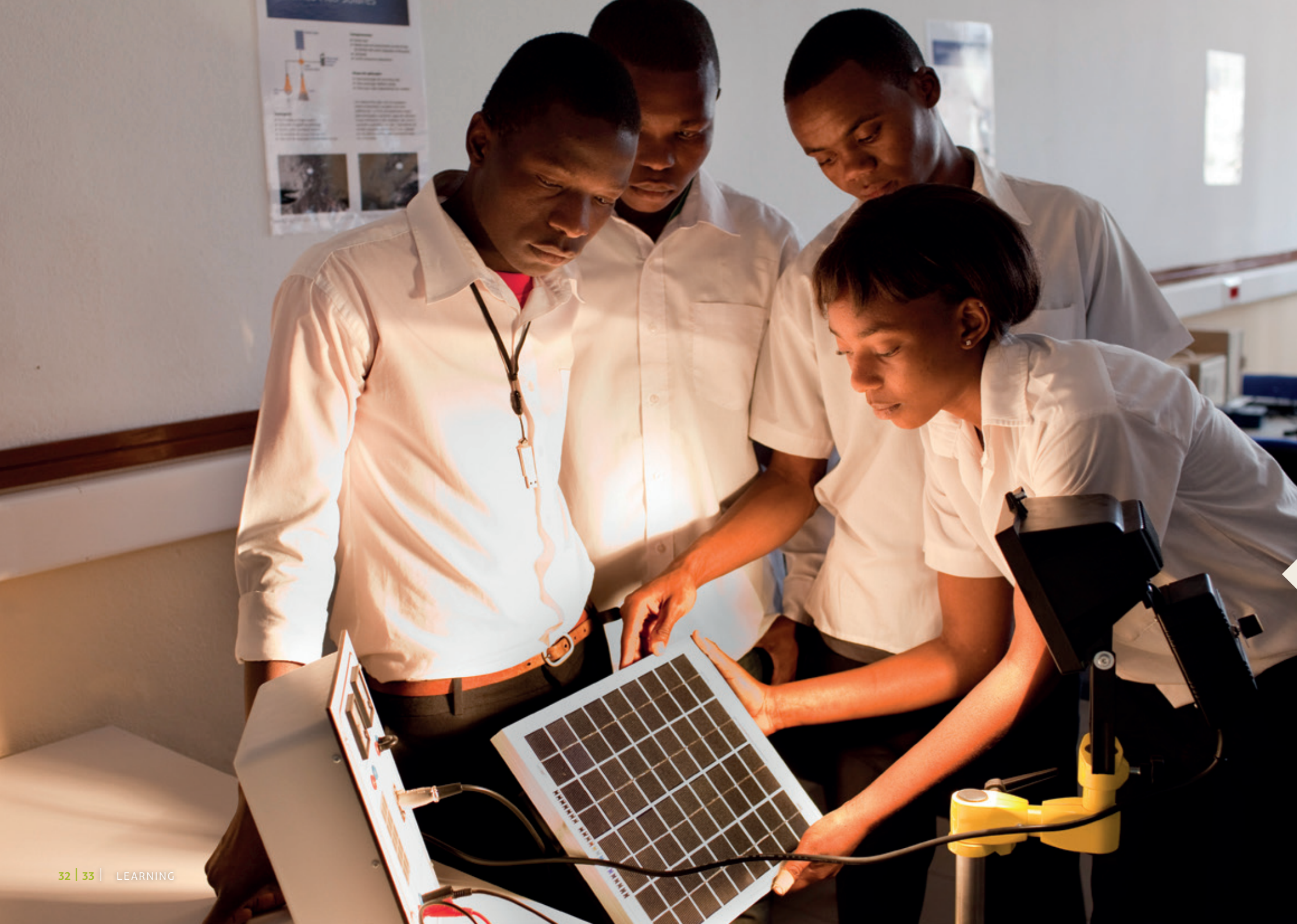
However, the irrigation measures did have a positive impact on harvesting cycles. In the project regions, smallholders were able to plant and harvest crops more regularly than those in other areas with similar attributes. According to some of the smallholders in the target group, the project helped to make agricultural work their main source of employment and secure a constant stream of income.



Unique sources of data

The difference-in-difference analysis revealed that the project did not have a significant effect on biomass production in the project regions. It was only after the additional analysis of satellite images that a change to the harvesting cycle could be identified – and thus an additional positive impact of the project.

In this case, the use of new data sources complements traditional evaluation methods – and facilitates a more robust and extensive evaluation of the project impact.



Learning from experience – promoting learning from completed projects is an important goal for the FC Evaluation Unit.

»» LEARN

Improving future projects



Current MAPME projects focus on monitoring the containment of illegal deforestation, desertification and the planning of infrastructure projects in the education and health sectors. Usually, operating departments and the FC Evaluation Unit are both involved in order to create synergies.



More information about the initiative is available here (QR code) or at: <https://www.mapme-initiative.org/>.

MAPME –

an initiative to promote the use of geodata for planning, monitoring and evaluation

MAPME (Maps for Planning, Monitoring & Evaluation) is an initiative launched by KfW Development Bank's Evaluation Unit in conjunction with Agence française de développement and MapTailor. The goal is to promote the wider use of geodata in FC projects: for planning, monitoring and evaluating projects, geodata can generate added value everywhere.

Open source and open data

MAPME is based on the consistent use of freely available satellite data and free software (open source). This enables cost-effective measures of various impact indicators and also ensures the wide-scale use of data at an institutional level. To give one example, the approach enables deforestation to be calculated before, during and after the implementation of a forest conservation project. As well as being low in cost, the standardised approach allows impacts to be recorded across a large number of individual projects.

Why use geo-information?

In evaluations, geo-information can be an important source for supplementing impressions and data from a project (triangulation). Its particular strength is the ability to finely break down data over time and space, or to put it more simply: satellites enable us to travel through time and space and observe various places on earth at different points in time – and measure any changes.

For example, in terms of methodology, this enables observation of project areas and comparable control areas before and after project implementation and, as such, facilitates tracking of developments and measurement of impacts. This delivers important data for the impact evaluation of FC projects. At the same time, the approach also enables recording of strong baseline data for better planning and creation of automatic monitoring systems.

From a stand-alone solution into a network

MAPME thrives on exchanges and connections between a variety of stakeholders – both within KfW Development Bank, but also with external institutions. For instance, the evaluation department at the French development bank Agence française de développement is an important partner in the initiative. The private sector is also involved through the start-up MapTailor Geospatial Consultants. Within KfW Development Bank, the Evaluation Unit works closely with the operating departments to tap into synergies and provide others with access to open source solutions. For example, it is collaborating with the Latin America department and the Competence Centre for Natural Resources to create a geo-referenced database containing the 600+ conservation areas promoted by KfW around the world. This is opening up some exciting opportunities for learning more about the effectiveness of projects. In the future, data regarding biodiversity and the threat status of conservation areas may be used to support planning and monitoring in impact management.

MAPME in practice: KfW's conservation area database

Conservation financing is an important area of work in FC and has been steadily grown since 2004. In 2019, KfW Development Bank financed 602 conservation areas in 66 countries with a current portfolio of EUR 2.6 billion. The FC Evaluation Unit is attempting to learn more about the promoted conservation areas by collaborating on a project with colleagues from the bank's operating departments. Both internal project data and open geodata are used for this work. The main objectives of conservation financing are subsumed under the three following areas:

- conserving biological diversity
- mitigating global climate change, e.g. by reducing deforestation
- improving the livelihoods of populations who rely on the natural resources

As the initial step, a geo-referenced database of conservation areas was created. To do this, internal project information was compared with the World Database on Protected Areas – WDPA (IUCN) and linked to external sources of data. The database currently consists of 433 conservation areas in 16 Latin American countries. The majority of these are terrestrial conservation areas, though 19 marine and 43 mixed marine/terrestrial conservation areas are also promoted. They cover a total area of 1.1 million km², an area roughly three times the size of Germany.

With this work, FC has contributed to protecting an area of forest covering over 0.9 million km². The majority of the promoted forests are located in the Amazon basin, where KfW Development Bank works with national partners from Bolivia, Brazil, Colombia, Ecuador and Peru.

Geodata supports the planning, monitoring and evaluation of conservation area projects

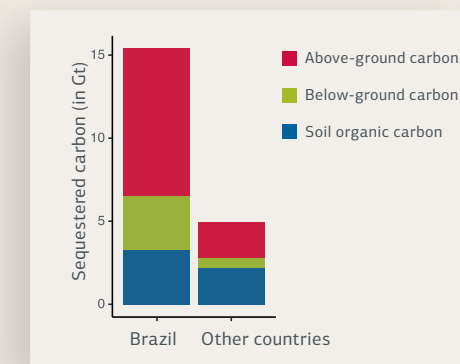
Geodata can shed light on the relevance of forest conservation for mitigating global climate change, for example by visualising carbon sequestration in the vegetation and soil of the supported conservation areas. Conservation can help to prevent this carbon from leaking into the atmosphere in the form of harmful greenhouse gases due to deforestation and forest degradation.

The network of promoted conservation areas in Latin America sequesters a total of 20.7 gigatonnes of carbon, an amount that is 103 times higher than Germany's annual greenhouse gas emissions. The majority of this carbon is stored in Brazil, an important partner country in international forest conservation.

By creating and analysing the conservation area database, the aim is to learn more about the effectiveness of FC projects and the acute potential threat in partner countries (and beyond). At the same time, this work also contributes to the planning of new projects and to satellite-based impact monitoring.



Locations and types of conservation areas supported in Latin America



Sequestered carbon in Latin America's conservation areas

The QUER app – digital and interactive access to more than 1,000 evaluation findings

Quick Evaluation Results (QUER) is the FC Evaluation Unit’s new digital knowledge tool. Staff from Financial Cooperation can use the QUER app to quickly access specific and tailored lessons learned from evaluation reports created since 2007, and apply these findings to new and current projects.

Promoting institutional learning

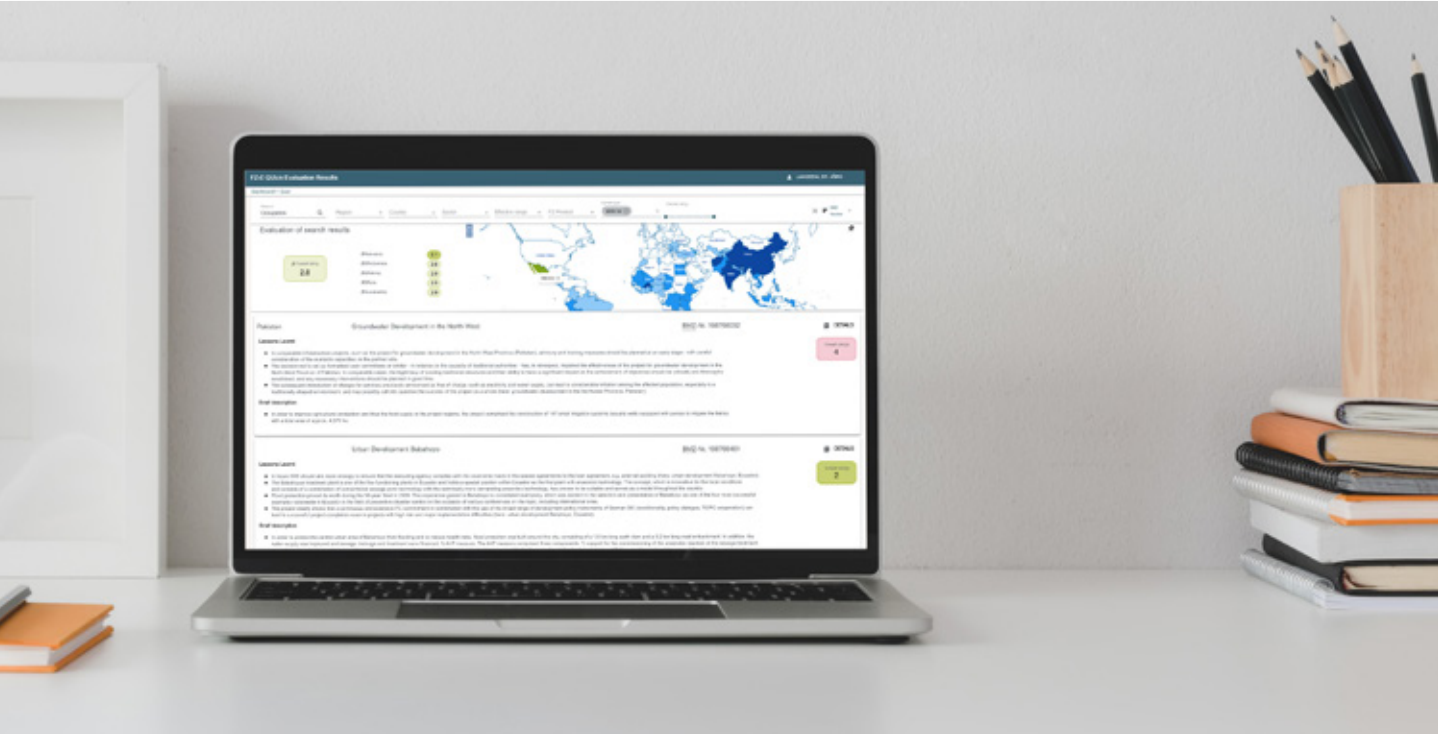
For many institutions, systematic institutional learning is an important goal at various levels. Two key elements are required to ensure that decisions are made on the basis of evidence: Firstly, evidence must be available in a format that is as systematically generated and structured as possible. And secondly, evidence must be processed in a way that makes it effective and constructive for the purpose of institutional learning (“evidence-to-practice”).

The work produced by the independent FC Evaluation Unit has been fulfilling the first requirement for the past 20+ years. Once completed, Financial Cooperation projects are systematically analysed in ex post evaluations based on OECD-DAC criteria. Since the introduction of an annual, representative random draw of completed projects in 2007, more than 1,000 evaluation results have been compiled in a structured manner.

To make use of this enormous wealth of evaluation findings, the FC Evaluation Unit asked itself how to meet the second requirement – “evidence-to-practice” – and what properties a modern knowledge base should have. The answer? Its contents should be focused and user-friendly, i.e. easily accessible and intuitive to understand, digital and interactive, and allow for demand-oriented searching and filtering.

Individually tailored evaluation findings at the touch of a button

This is how QUER – Quick Evaluation Results – came into being. The interactive app contains over 1,000 evaluation results from 2007 onwards. Knowledge that was previously only accessible by laboriously searching through the PDF versions of reports has now been digitalised.



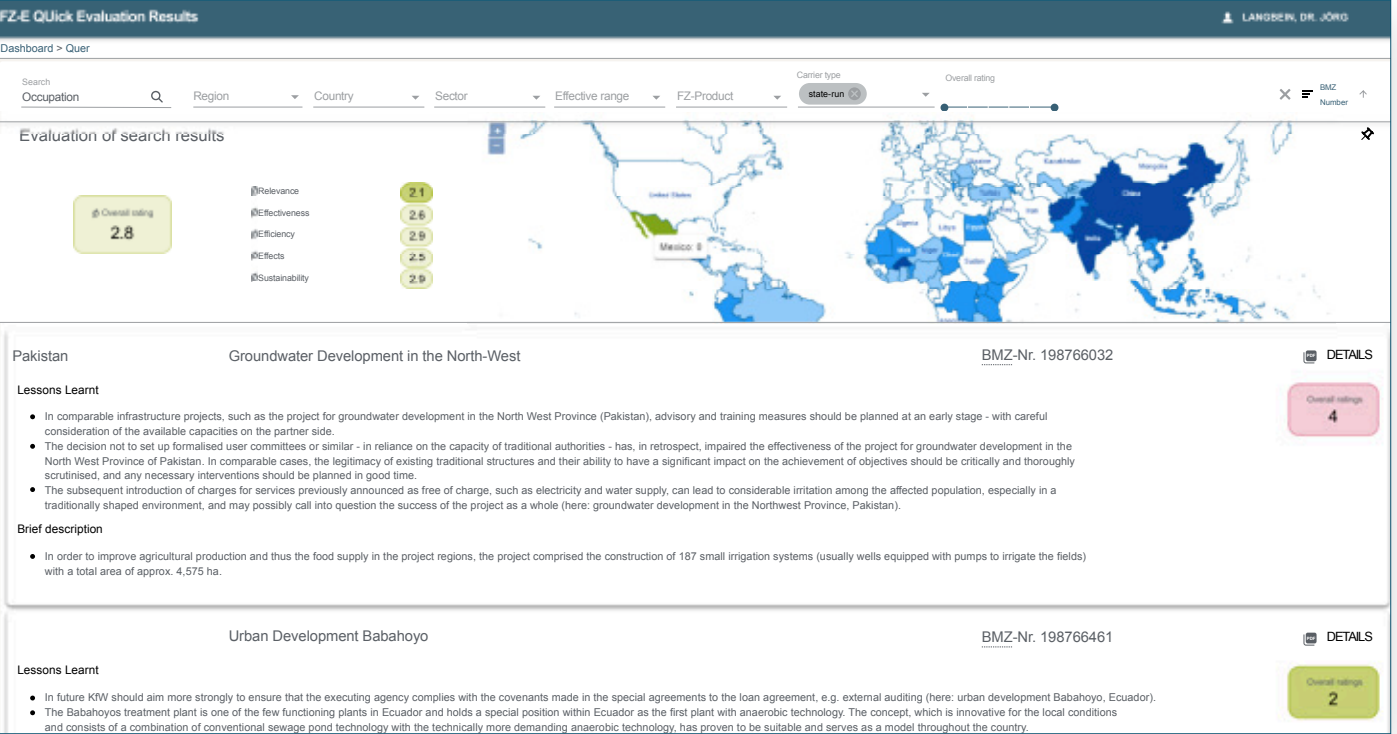
The interactive app – with all ex post evaluations by the FC Evaluation Unit since 2007.

This enables users to find exactly the evaluations they need to plan new projects in just a few clicks. Using the free text search function, users can enter key words and then filter reports by region, sector and type of executing agency. For example, when portfolio managers are designing a new energy project in Uganda, they can read up on experiences from similar projects in the country, region or even other places around the world. What is more, the most important

lessons learned from every evaluation have been processed separately and can be accessed directly. In addition to a brief description, a summary and the entire report are available at the touch of a button.

QUER therefore offers precise, uncomplicated and tailored responses to questions such as:

What are the factors for the success or failure of certain projects? Do projects fare better depending on whether the project-executing agency is a state agency, multilateral organisation or NGO? Which indicators were used? Which targets were formulated for the implemented projects? Which risks need to be taken into account? The app is part of KfW’s internal Portfolio Management Tool (PMT) and is continuously updated with new findings so that all FC staff have direct and quick access. QUER helps them to use the evaluation findings for the planning of new projects or the analysis of entire regions and portfolios.



In-depth view of an evaluated project’s data – lessons learned, risks, indicators, evaluation in line with OECD DAC.



COVID-19 and employment – suggestions for low- and middle-income countries



You will find results from previous evaluations here, along with information on how it can be applied to the coronavirus pandemic.



A rapid response is important – countries should be expanding healthcare services and providing businesses and employees with swift financial support.

Financial support should be designed to enable companies adversely affected by the crisis to continue to pay their staff. At the same time, these companies’ access to funds should be facilitated through the provision (and subsidisation) of new credit lines. Microfinance institutions can play a role in reaching informal enterprises and securing liquidity. In the past, subsidised wages and the temporary reduction of other labour costs, such as social insurance contributions, have proven helpful in preventing job losses. Reaching particularly at-risk businesses and employees in LMIC contexts, especially those in more rural areas, is challenging.

The FC Evaluation Unit also presents evaluation results in shorter formats, such as “Evaluierung aktuell”, in English “Evaluation update”, (see QR code in the box) and Policy Briefs. The following article is an abridged version of a Policy Brief authored by Jochen Kluve, Jörg Langbein (both KfW) and Michael Weber (World Bank), which was written on behalf of the Federal Ministry for Economic Cooperation and Development (BMZ) and published in October 2020. In the article, the authors argue that a tailored combination of labour market and social protection measures is needed to mitigate the economic and social effects of COVID-19. This approach is designed to be implemented on a step-by-step basis following the three phases of the crisis: assistance, reorganisation, and increasing resilience. The unique nature of the pandemic calls for flexibility and a willingness to try new things. When looking for proof of the effectiveness of new or modified programmes, consistent appraisals and evaluations can contribute to improved social protection and better employment policy.

COVID-19 has impacted economies, businesses and workers all over the world. What started out as a health crisis soon developed into a pandemic with serious consequences for economies and labour markets across the globe. By August 2020, around 70% of countries had introduced movement restrictions, impacting businesses and people’s livelihoods.¹ The global economy was hit extremely hard, much harder than during previous crises (ILO, 2020).² The number of hours worked fell, unemployment rates rose, income was lost and businesses were shut down. In countries with low and middle incomes (LMICs), the economic consequences are particularly severe for vulnerable populations, such as those in informal employment, women and young people, and also for small- and medium-sized enterprises (SMEs).

Providing short-term assistance

In theory, early support measures should dampen the direct consequences of state healthcare measures for otherwise competitive businesses by helping them to avoid liquidity shortages and to keep staff employed. Several options for short-term policies are outlined below:

Digital solutions such as mobile-phone-based payment systems or digital registration solutions can help in this regard.

Employees who lose their job need a secure income to preserve their livelihoods. One of the most common instruments for providing quick help to households in a crisis is social assistance programmes, particularly cash transfers. By September 2020, 156 out of 188 countries had planned, implemented or already concluded a monetary transfer programme due to the crisis.³ There is a wide base of evidence demonstrating the positive effects that cash transfers have on securing livelihoods; they can help to reduce poverty, for example, and improve health. People are given the opportunity to participate actively in the labour market, support is provided for tackling economic and climate-related crises, and economic multipliers are created to ensure consumption smoothing.⁴

Medium- and long-term reorganisation and building resilience

Emergency measures must be replaced by tailored support for workers and enterprises in the medium- and long-term phases of reorganisation and resilience building.

Here, the focus should be on jobs that support a sustainable structural shift that is in harmony with the environment. Furthermore, a large number of workers in LMICs are employed by microenterprises or are self-employed working from home. Targeted support measures are important for this group; microfinancing networks can be used to pass funds on to these enterprises.

Active labour market programmes – such as qualification measures – can support the adjustment of political measures in line with the changing circumstances of the crisis, from the mitigation of immediate effects to the restructuring phase. Public works, for example, played an important role in the recovery process after past shocks, such as the 2008–2010 financial crisis.⁵ Given the nature of Covid-19, adaptations to labour intensive public works are nonetheless needed. Programmes need to ensure that participants can maintain social distancing rules and wear appropriate personal protective equipment. Investing in skills, expertise and training, particularly those that focus on digital solutions, can therefore be a sensible long-term investment. Investing in (advanced) qualifications is particularly effective, especially during an economic crisis.⁶

Expanding programmes to include vulnerable members of the population is key for future social protection programmes. In LMICs, the range and scope of social protection measures needs to be expanded to ensure greater resilience for future crises. This is true for vulnerable workers, such as low-income, informal and low-skilled workers, as well as women. At the same time, policy-makers should pay particular attention to the new generation of young job-seekers – the so-called “COVID-19 generation” – to avoid any long-term negative effects.⁷

Summary and recommendations

A tailored combination of labour market and social protection measures is needed to mitigate the economic and social effects of COVID-19. This approach should be applied gradually, in line with the three phases of the crisis: assistance, reorganisation and increasing resilience. In order to gain proof of the effectiveness of new or modified programmes, social protection and employment policies can be rigorously and consistently evaluated.⁸

³ Source: Gentilini, U., M. Almenfi, P. Dale, A.V. Lopez and U. Zafar (2020). Social Protection and Jobs Response to Covid-19: A real time review of country measures. COVID-19 Living Paper. Version 13. 18 September. World Bank, Washington, DC.

⁴ For example: Garcia, S. and J. Hill (2010). The impact of conditional cash transfers and health: unpacking the causal chain. Journal of Development Effectiveness, 2(1): 117–137. or: Kabeer, N. and H. Waddington (2015). Economic impacts of conditional cash transfer programmes: a systematic review and meta-analysis. Journal of Development Effectiveness, 7(3): 290–303.

⁵ Source: Azam, M., C. Ferré and M.I. Ajwad (2013). Can public works programs mitigate the impact of crises in Europe? The case of Latvia. IZA Journal of European Labor Studies, 2(1): 10.

⁶ Source: Card, D., J. Kluve and A. Weber (2018). What works? A meta analysis of recent active labor market program evaluations. Journal of the European Economic Association, 16(3): 894–931.

⁷ Cf.: Gregg, P. and E. Tominey (2005). The wage scar from male youth unemployment. Labour Economics, 12(4): 487–509.

⁸ Source: Haushofer, J. and C. J. Metcalfe (2020). Which interventions work best in a pandemic?. Science, 368(6495): 1063–1065.

¹ Source: University of Oxford und Blavatnik School of Government (2020). Coronavirus Government Response Tracker. www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker (accessed 25 August 2020).

² Source: ILO (2020), ILO Monitor: “COVID-19 and the world of work”, ILO Briefing Note, 6th Edition, International Labour Organization, Geneva, www.ilo.org/wcmsp5/groups/public/@dgreports/@dcomm/documents/briefingnote/wcms_755910.pdf (accessed 20 September 2020).



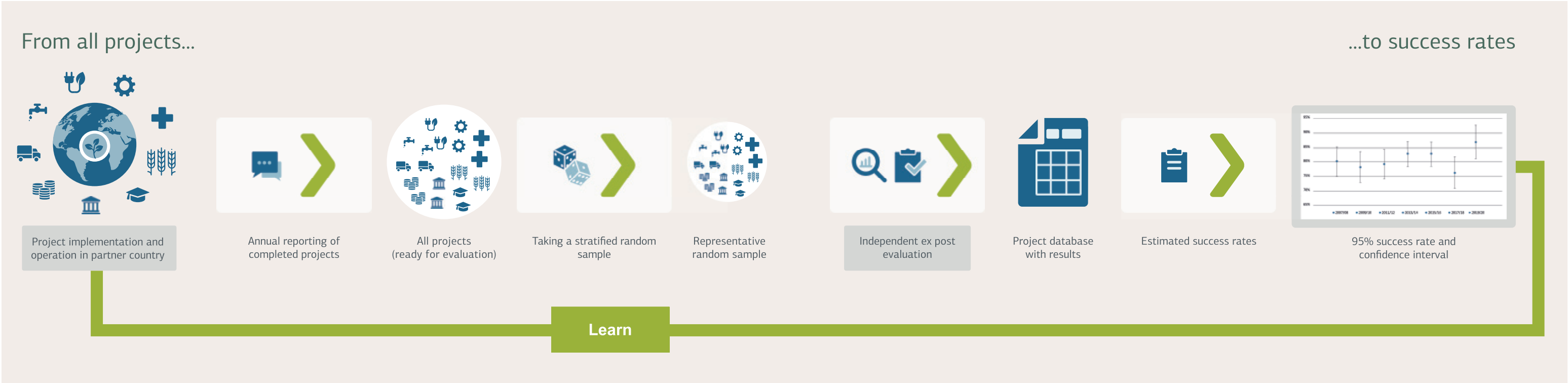
From projects to success rate - Every year FC E draws a representative sample from all FC projects. This sample forms the basis for calculating the success rate.

»» Methodology

Ex post evaluations 2019-2020

Working practices and methodology – how does the FC Evaluation Unit calculate the success rate?

Every year a representative random sample of projects is selected for evaluation, which the FC Evaluation Unit then assesses on the basis of OECD-DAC criteria. The results of these evaluations form the basis for calculating the success rate, in other words, statistically estimating the percentage of successful FC projects.



How is the overall rating calculated from the ratings of the individual OECD-DAC criteria?

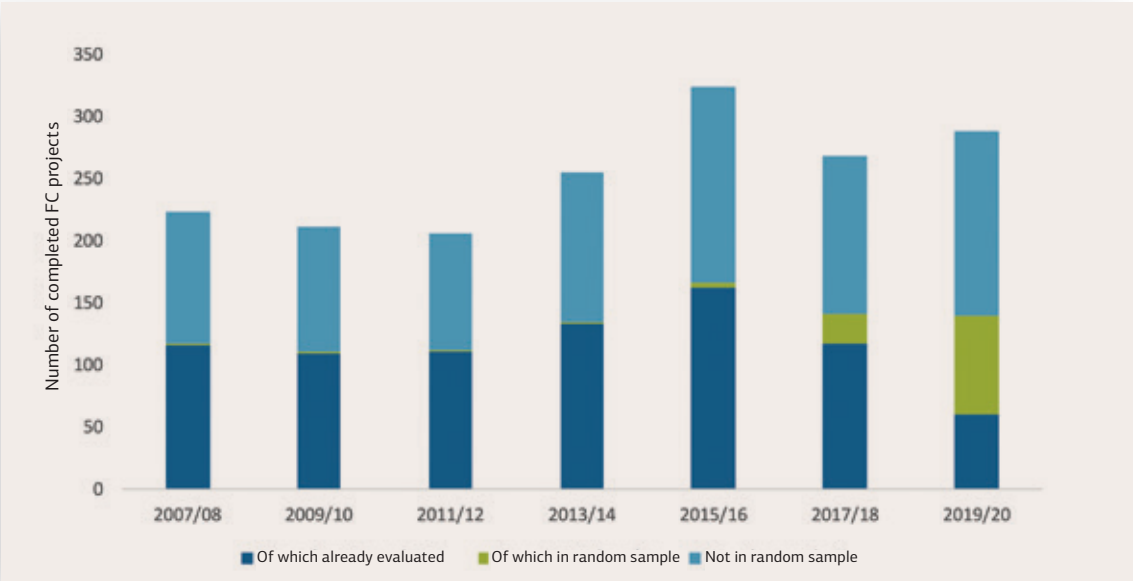
- Each evaluation rates a project along the six OECD-DAC criteria. To take into account that these multiple criteria differ in their importance, the overall rating of a project is not a simple average of the individual criteria (1–6 or 1–4 for sustainability). This ensures that particularly serious deficiencies in individual criteria cannot be offset by positive results in others.
- Concretely, poor results in the OECD-DAC criteria of effectiveness, impact and sustainability generally prevent a project from being classified as successful overall. The underlying idea is that projects which do not achieve their objectives, or whose impacts are not sustainable (in the sense of being long-lasting), do not deserve to be evaluated as successful, despite their potentially successful performance in other criteria.

- Every year, KfW Development Bank's Evaluation Unit takes a random sample (stratified by sector) from the universe of completed measures. An ex post evaluation is conducted for all projects in the random sample. This method ensures that the evaluations are statistically representative and informative for the FC portfolio and maintains independence in the selection of projects to be evaluated. In the two-year period of 2019/2020, a total of 149 projects were randomly selected from the universe of 292 completed projects.
- Further projects can be evaluated, even if they are not in the sample, for example because they are preceding or follow-up phases of sampled projects. In the calendar years 2019 and 2020, there were 42 such "pooled" projects. Together with 129 evaluations from the original random sample, 171 projects were evaluated in the years 2019 and 2020.
- Only projects that are part of the random sample are used to estimate the success rate (unbiased estimate). The success rate is defined as the percentage of projects that received an overall rating of 3 or above. Overall, 63 projects have been included in the success rate calculation so far for the basic population for 2019/2020.
- Note that success rates are estimates and therefore always subject to statistical uncertainty. The confidence interval provides information about the accuracy of the estimate.
- For example, the success rates from past two-year periods since the introduction of the random-sample approach in 2007/2008 are not statistically different from one another. Despite the average success rates fluctuating between 77 % and 87 %, the confidence intervals for the success rates overlap in the region of 81 % – it is therefore possible and seems likely that the underlying success rates of all FC projects are within this overlapping range.

Sample size and distribution of ratings – results for 2019/2020

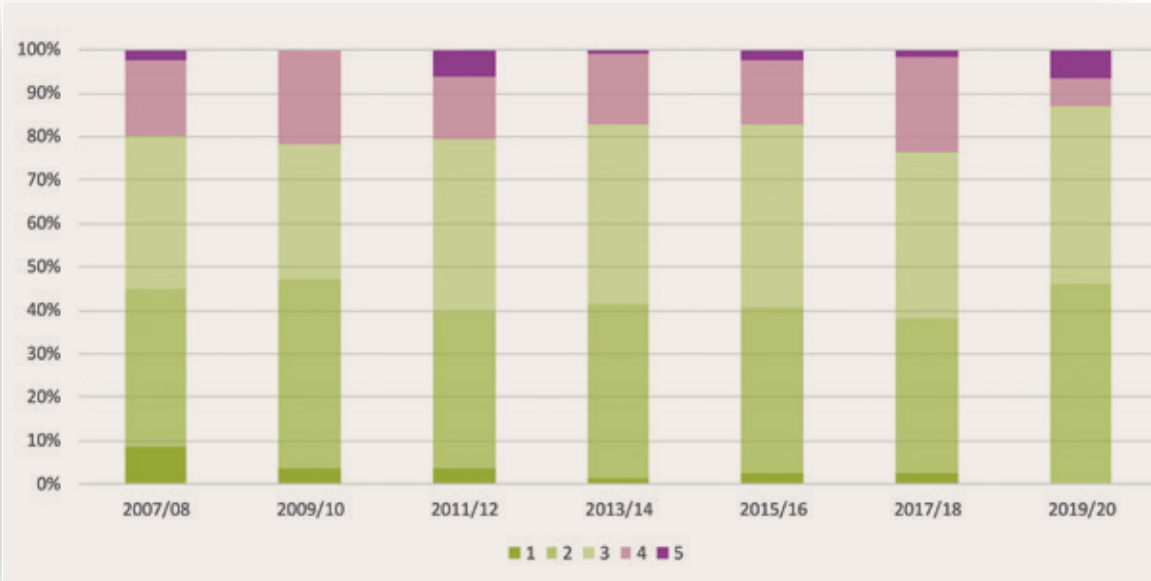
The results from this reporting period are based on a similar number of ex post evaluations as in previous years and also exhibit a similar success rate: almost 86 percent of all evaluated projects were rated as successful. This percentage is slightly higher than in the years before.

A similar number of ex post evaluations were performed in the calendar years 2019 and 2020 as in previous years.

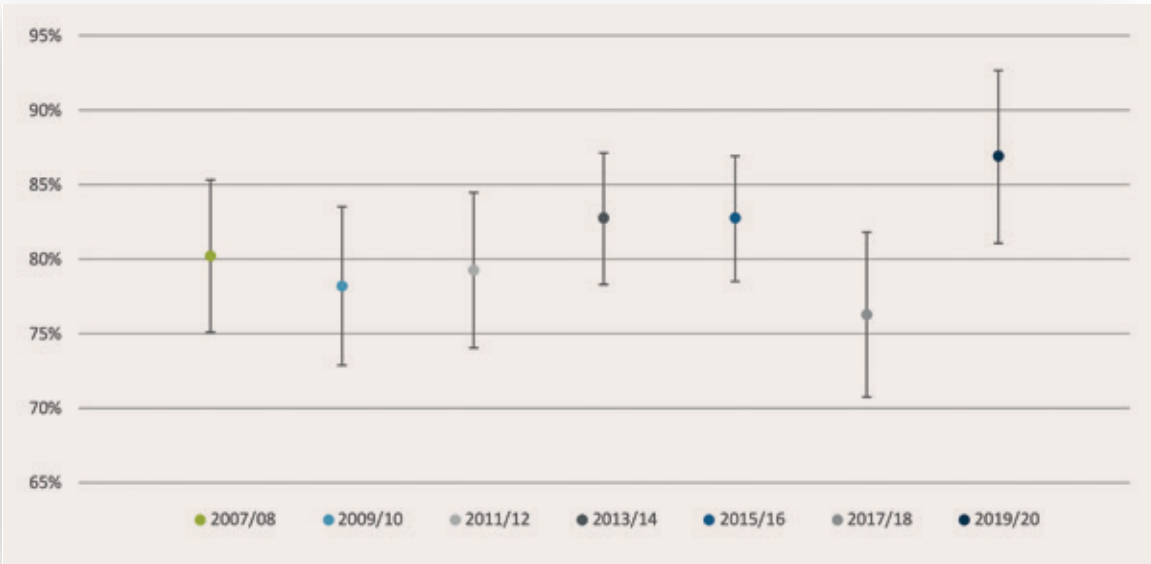


Number of evaluated projects from the random sample by project numbers

- Out of a total of 171 evaluations, 129 were included in the success rate calculation (only projects from the drawn sample), 63 of which originate from the 2019/2020 sample. The remaining 66 projects originate from the samples from previous years.
- The majority of the 63 projects from the 2019/2020 sample were rated as 2 or 3, though the percentage of projects rated as 2 was slightly higher. The remaining projects were rated as 4 and 5 to roughly the same extent (5 % and 6 %, respectively). No ratings of 1 were given. Projects are deemed successful if they receive a rating of 3 or above.
- The distribution of ratings is therefore similar to the distribution in previous periods but with some small changes: while there are significantly fewer projects with a rating of 4, the percentage of projects rated as 2, 3 and 5 has increased. The top rating of 1 remains an exception.



Distribution of ratings over time (top) and estimated success rate with confidence interval (bottom)



The success rate for the period of 2019/2020 is slightly higher than in previous years but does not exhibit statistically significant differences.

- On the whole, the success rate has been trending upwards since 2011/2012. However, this is not statistically significant, nor is the slightly lower value for the estimated success rate in the 2017/2018 period.
- The percentage of completed evaluations from the 2019/2020 random sample is still comparatively low. As such, the estimated success rate for the universe of evaluations from 2019/2020 has yet to be finalised.
- By evaluating remaining projects from earlier samples, it was possible to increase the accuracy of estimated success rates for previous periods. For example, 56 evaluations from the random sample 2017/2018 were completed in the 2019/2020 period, which meant that the range of the confidence interval could be reduced by around eight percentage points.

Country	Project
Health	
Afghanistan	Polio Control Programme, Phase II
Armenia	Supraregional Tuberculosis Control Programme II
Azerbaijan	Supraregional Tuberculosis Control Programme II
Burkina Faso	Basket financing, Healthcare I and II Basket financing, Healthcare I and II HIV/AIDS prevention and reproductive health
CEMAC - Central Africa	HIV Prevention in Central Africa Phase II
	HIV Prevention in Central Africa Phase II
	HIV/AIDS Prevention in Central Africa Phase III
	HIV/AIDS Prevention in Central Africa Phase IV
Georgia	Supraregional Tuberculosis Control Programme II
Yemen	Basic nutrition / maternal and infant health III
Kazakhstan	National Tuberculosis Control Programme I and II in Kazakhstan
	National Tuberculosis Control Programme I and II in Kazakhstan
Kyrgyzstan	Tuberculosis (TB) Control Programme III in Kyrgyzstan
	Tuberculosis (TB) Control Programme IV in Kyrgyzstan
Pakistan	CP- Basic Health Programme in FATA
Philippinen	Programme to support the Philippine Health Sector Reform Agenda
Sierra Leone	Combating HIV/AIDS (HAPP I)
Tansania	Combating sexually transmitted diseases and HIV
	Combating sexually transmitted diseases and HIV
	Active health



You can find all ex post evaluations online.

Imprint

Published by
KfW Group

Group Communications Department
Palmengartenstrasse 5–9
60325 Frankfurt am Main, Germany

Telephone: +49 69 7431-0
infocenter@kfw.de
www.kfw.de

Editorial team
KfW Development Bank
FC Evaluation Unit

Design
www.adesso-experience.de
www.forsch-design.com

Print
www.rehmsdruck.de
Zero net emissions were generated in
the production of this publication

Paper
www.antalıs.de
Cover – Invercote Creato 350gr. FSC
Content – Nautilus SuperWhite 160gr.
FSC, recycled

Photo credits

Cover KfW Group / Jonas Wresch;
P. 4 KfW-stock / ich.tv /Thorsten Thor;
P. 4 + 5 KfW KfW Group / Jonas Wresch; KfW
Group / two photos Photothek / Ute Grabow-
sky; KfW Group / Claudia Arce; P. 6 + 7 KfW
Group / Susanne Schröder; P. 8 KfW Group /
Jonas Wresch; P. 10 + 11 KfW Group / Frank
Blümner, FUNDASAL / Salomon Morales;
P. 12 + 13 KfW Group / Jonas Wresch;
P. 14 + 15 Studio Neuberg / Arben Llapashtica;
P. 16 + 17 KfW Group / Sabrina Behr;
P. 18 + 19 KfW Group / Anis Alshargabi;
KfW Group / Andrea Höltnke; P. 20 + 21 KfW
Group / Photothek / Ute Grabowsky; P. 22 + 23
KfW-stock / Finance in Motion / SANAD, KfW
Group / Photothek / Ute Grabowsky; P. 24 +
25 KfW Group / Photothek / Ute Grabowsky,
KfW Group / Jonas Wresch, KfW Group / Brit
Horschke; P. 26 + 27 KfW Group / Photothek /
Ute Grabowsky, KfW Group / Lotte
Westermann; P. 28 + 29 KfW Group / Johannes
Schielein; P. 30 + 31 KfW-stock / Joerg
Boethling, KfW-stock / Fred Hoogervorst;
P. 32 + 33 KfW Group / Photothek / Ute
Grabowsky; P. 34 RKive / Alamy Stock Photo;
P. 35 + 36 + 37 Adesso / Sarah Profitlich;
P. 38 + 39 KfW Group / Jonas Wresch; P. 40 +
41 KfW Group / Claudia Arce Backpage KfW
Group / Florian Kopp



KfW Group
KfW Development Bank
Palmengartenstrasse 5-9
60325 Frankfurt am Main, Germany

Telephone: +49 69 7431-0
info@kfw-entwicklungsbank.de
www.kfw.de

600 000 4847