KNOWING WHAT WORKS

Central project evaluation

Improvement of labour and social standards in the Pakistani textile industry, Pakistan
Project number 2016.2029.3

Evaluation Report

On behalf of GIZ by Felipe Isidor-Serrano (Mainlevel Consulting AG) and Khalid Mehmood

Published: August 2022

giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

Publication details

Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH is a federal enterprise and supports the German Federal Government in achieving its objectives in the fields of international education and international cooperation for sustainable development.

GIZ's Evaluation Unit reports directly to the Management Board. It is separate from GIZ's operational business. This organisational structure strengthens its independence. The unit is mandated to generate evidence-based results and recommendations for decision-making, to provide plausible verification of results and to increase the transparency of findings.

The Evaluation Unit commissioned external independent evaluators to conduct the evaluation. This evaluation report was written by these external evaluators. All opinions and assessments expressed in the report are those of the authors.

Evaluators:
Dr Felipe Isidor-Serrano, André Gersmeier (Mainlevel Consulting AG), Khalid Mehmood

Authors of the evaluation report:
Dr Felipe Isidor-Serrano, André Gersmeier (Mainlevel Consulting AG), Khalid Mehmood

Consulting firm:
Mainlevel Consulting AG
Ludwig-Erhard-Str. 14
65760 Eschborn
T: +49 69 153 240 920
E: +49 69 153 240 929
I: info@mainlevel.de

Coordination and management:
Claudia Kornahrens, GIZ, Head of Section
Vera Hundt, GIZ, Evaluation Manager
GIZ Corporate Unit Evaluation
Central Project Evaluation Section

Responsible:
Albert Engel, GIZ, Director
GIZ Corporate Unit Evaluation

Editing:
International Correspondents in Education (ICE)

Published by:
Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH

Registered offices:
Bonn and Eschborn
Friedrich-Ebert-Allee 32 + 36
53113 Bonn, Germany
T: +49 228 44 60-0
F: +49 228 44 60-17 66
E: evaluierung@giz.de
I: www.giz.de/evaluierung
www.youtube.com/user/GIZonlineTV
www.facebook.com/gizprofile
https://twitter.com/giz_gmbh

Design/layout:
DITHO Design GmbH, Cologne

Printing and distribution:
GIZ, Bonn

Printed on 100% recycled paper, certified to Forest Stewardship Council (FSC) standards.

Bonn, August 2022

This publication can be downloaded as a PDF file from the GIZ website:
www.giz.de/evaluierung.

Contents

List of figures

List of photos

List of tables

Abbreviations

BMZ	German Federal Ministry for Economic Cooperation and Development
CoP	Community of Practice
CPE	Central Project Evaluation
DAC	Development Assistance Committee
DfS	Dialogue for Sustainability
EFP	Employers Federation of Pakistan
EU	European Union
FABRIC	GIZ regional programme Social and Labour Standards in the Textile and Garment Sector in Asia
FES	Friedrich Ebert Foundation
GDP	Gross Domestic Product
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH
GSP+	Generalised System of Preferences
ILO	International Labour Organization
IRI	Industrial Relations Institute
LHRD	Labour & Human Resource Department
M&E	Monitoring and Evaluation
NGO	Non-Governmental Organisation
OECD	Organisation for Economic Co-operation and Development
OHS	Occupational health and safety
PESSI	Punjab Employees Social Security Institute
RBM	Results-Based Monitoring
SAA-CIWCE	Saeed Ahmad Awan Centre for Improvement in Working Conditions and Environment
SDG	Sustainable Development Goal
SMART	Specific, Measurable, Achievable, Relevant and Time-Bound
SP-SHP	GIZ project Support to Social Security Including Health Insurance in Pakistan
ToC	Theory of Change
TVET	Technical and Vocational Education and Training

The project at a glance

Pakistan/Punjab: Improvement of labour and social standards in the Pakistani textile industry

Project number	2016.2029.3
Creditor reporting system code(s)	16020 – Employment policy and planning (60%) 25010 – Public sector policies and institution support to the business environment (40%)
Project objective	The prerequisites for compliance with labour and social standards in the province of Punjab have improved.
Project term	January 2017 - December 2020
Project value	EUR 12,050,000
Commissioning party	German Federal Ministry for Economic Cooperation and Development (BMZ)
Lead executing agency	Labour & Human Resource Department (LHRD) of the province of Punjab
Implementing organisations (in the partner country)	• Directorate General of Labour Welfare • District Labour Offices of Lahore, Sialkot, Faisalabad and Multan • Punjab Employees Social Security Institution (PESSI) • Punjab Workers Welfare Board • Saeed Ahmad Awan Centre for Improvement in Working Conditions and Environment (SAA-CIWCE) • Industrial Relation Institute (IRI)
Other development organisations involved	International Labour Organization (ILO)
Target group(s)	Staff of LHRD as well as the management and employees of factories in the textile sector, with a focus on the province of Punjab

1 Evaluation objectives and questions

This chapter aims to describe the purpose of the evaluation, the standard evaluation criteria, and additional stakeholders' knowledge interests and evaluation questions.

1.1 Evaluation objectives

Functions of the evaluation

Central project evaluations (CPEs) of projects commissioned by German Federal Ministry for Economic Cooperation and Development (BMZ) fulfil three basic functions: they support evidence-based decisions, promote transparency and accountability, and foster organisational learning within the scope of contributing to effective knowledge management. The Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH structures the planning, implementation and use of evaluations so that the contribution the evaluation process and the evaluation findings make to these basic functions is optimised (GIZ, 2018a).

Selection of the evaluation

The project to be evaluated has been selected randomly following the guidelines for GIZ's CPEs.

Evaluation type

According to the Terms of Reference and GIZ's evaluation guidelines, this evaluation is a final evaluation, the project under evaluation having ended on 31 December 2020.

Limitations regarding feasibility of the evaluation

Cooperation with GIZ Corporate Unit Evaluation as well as with the project team was very fruitful, with all parties involved showing interest in and ownership towards the evaluation. Nevertheless, a remote evaluation mission had to be conducted due to COVID-19 restrictions.

1.2 Evaluation questions

The project is assessed on the basis of standardised evaluation criteria and questions to ensure comparability by GIZ. This is based on the Organisation for Economic Co-operation and Development (OECD)/Development Assistance Committee (DAC) evaluation criteria (updated 2020) for international cooperation and the evaluation criteria for German bilateral cooperation (in German): **relevance, coherence, efficiency**, **effectiveness**, **impact** and **sustainability**.

Specific assessment dimensions and analytical questions have been derived from this framework. These form the basis for all CPEs in GIZ and can be found in the **evaluation matrix** (Annex). In addition, contributions to the 2030 Agenda for Sustainable Development and its principles are taken into account, as are cross-cutting issues such as gender, the environment, conflict sensitivity and human rights. Also, aspects regarding the quality of implementation are included in all OECD/DAC criteria.

Use of results

Although the official follow-on project had already started during the evaluation, the findings should be used to steer the project based on lessons learnt and recommendations from the evaluation. A previous project financed by the German Federal Foreign Office is an indirect predecessor of this project. Because it is not an official predecessor, it was not part of the OECD/DAC criteria assessment (see section 4.1). The follow-on project is also not part of the OECD/DAC criteria assessment. Lessons learnt are considered in the recommendations (section 5.2) and included in the effectiveness (section 4.4), impact (section 4.5) and sustainability (section 4.7) criteria.

Additional evaluation questions

During the inception mission, a participatory exercise ('Wish Tree') was conducted with the project team members to understand their knowledge interests in the evaluation. In addition, interviews were conducted with GIZ's sectoral unit, BMZ and the project's community of practice (CoP). The knowledge interests of the stakeholders are largely covered by the questions from the evaluation matrix.

As Pakistan is considered to be a volatile/fragile country, additional evaluation questions as part of the OECD/DAC criteria assessment (see Annex) regarding context and conflict sensitivity (do no harm) were incorporated.

Table 1: Knowledge interests by main evaluation stakeholder groups

Evaluation stakeholder group	Knowledge interests in evaluation/ additional evaluation questions	Relevant section in this report
BMZ	• What does the project do that might already be covered by other sectors and global projects (not bilateral projects) and that might not be necessary? • Structural sustainability of the project: What would happen if project staff (e.g. the project manager) were to leave? • How did the project manage to adapt to changes (change offers)?	Included in **coherence criterion** Included in **sustainability criterion** Included in **relevance criterion**
GIZ's sectoral unit	• How effective is the capacity development of stakeholder groups that have a high turnover of staff? • Lessons learned from different approaches (e.g. Dialogue for Sustainability (DfS))	Included in **effectiveness criterion**
Key project partners, in particular the CoP	• How did the company culture change towards more cooperation?	Included in **impact criterion**
GIZ project team	• Overall lessons learned, and the outcome, impact and sustainability of activities	Included in **effectiveness, impact** and **sustainability criteria**
GIZ Corporate Unit Evaluation	• Accountability towards the public (success rate of GIZ's projects) • Learning to understand strengths and weaknesses of single projects, potentials for replications in other countries and lessons learnt • Informing key stakeholder who enquire about GIZ activities	Included in **all criteria**

2 Object of the evaluation

This chapter aims to define the evaluation object, including the theory of change (ToC), and results hypotheses.

2.1 Definition of the evaluation object

The main object of evaluation is the selected technical cooperation measure 'Improvement of labour and social standards in the Pakistani textile industry', identified by the project number 2016.2029.3, and henceforth called 'the project'.

Temporal delineation: The object of this evaluation is the project with an overall term from 01 January 2017 to 31 December 2020.

Financial delineation: The project was financed through funds from BMZ and was implemented by GIZ. The total budget of the project was EUR 12,050,000. There was no co-financing for the project and this is therefore not part of the evaluation. As part of the total budget, the project administered a financing agreement to the ILO of EUR 1,700,000.

Predecessor project: The project 'Implementation of Social Standards Support Programme to the Textile and Garment Industry of Punjab' (PN 2013.9062.4), financed by the German Federal Foreign Office, and the project 'Water Efficiency in the Textile Industry' (PN 2013.9773.6) are indirect predecessors of this project. The Water Efficiency in the Textile Industry project ended in 2018, but as part of the third change offer in 2019 (GIZ, 2019a), the project continued initiatives from the predecessor from November 2019 onwards.

Follow-on project: A follow-on-project, 'Improvement of labour, social and environmental standards in Pakistan's textile industry' (PN 2019.2141.0), with a project term from 01 January 2021 to 31 December 2023, had already started during the evaluation.

Geographical delineation and focus: The project focused on the textile and garment industry in Pakistan, with its large textile clusters in Punjab, including the cities of Multan, Sialkot, Faisalabad and Lahore.

Political and sectoral context and framework conditions: The textile industry in Pakistan contributes 54% of the total export revenue and provides employment to 40% of the total labour force (about 15,000,000 people (Board of Investment, 2020)). The textile sector contributes 8.5% of the country's gross domestic product (GDP). Punjab is the region of Pakistan with the biggest share of the textile sector, as almost 70% of the industry is based there. Faisalabad city is the hub of textile activities in Punjab. Despite the textile sector being the key contributor to the economy of Pakistan, its working conditions are poor in many spheres, such as wages, social security, workplace safety, gender equality, and employment terms and conditions (GIZ, 2019a: 5). Despite a minor decline in exports between May and July 2020, the export-oriented textile industry benefited from the COVID-19 pandemic, leading to overall growth of 23% at the end of 2020 compared with 2019 (GIZ, 2021a: 7f).

The key governmental department responsible for improving the working conditions in Punjab is LHRD, with its mission statement of 'promoting the welfare and protecting the rights of the labour force and workers' (Government of Punjab, 2021). In 2011, as part of the decentralisation process, the responsibility for implementing labour standards was transferred to the provinces.

The labour inspection system in Pakistan is considered by the European Commission (2020: 4) to be weak, and low on the agenda of priorities. There are major reservations about the state participants in the inspection system. Until now, private sector actors have viewed investments in improving labour standards as largely isolated from improvements in production processes and have thus far been unable to identify potential for boosting production through improved working conditions.

Pakistan was granted Generalised System of Preferences (GSP+) status by the European Union (EU) in 2014, guaranteeing the country access to European markets. In order to maintain GSP+ status, Pakistan is required to implement core labour and social standards in addition to relevant international conventions and agreements in the area of good governance and sustainable development. To comply with the GSP+ regulations, a joint effort by the government, companies, and employers' and workers' organisations is needed. Pakistan was granted GSP+ status for up to 10 years, provided it demonstrated successful efforts in the implementation of 27 international conventions and agreements pertaining to labour and environmental standards.

Conflict and fragile context of the project: One of the main factors of fragility in Punjab is the political culture, which is deeply rooted in the feudal tradition. As a result, the political system is highly resistant to change, particularly when it comes to governance. The private sector, especially the textile and garment industry in Punjab, is similarly rooted in feudal structures that provide room for exploitative relations between employer and employee. Closely related to the challenge of radicalisation is that of combating violent extremism, although anti-terrorist operations are not merely focused on religious extremism, but also follow the logic of political considerations and fears. Balochistan and Khyber Pakhtunkhwa were the most terrorism-affected regions, while the project region of Punjab remained relatively stable (GIZ, 2018e: 5–7).

2.2 Results model including hypotheses

Contribution analysis (following Mayne (2012)) forms the cornerstone of this evaluation. A project's ToC is central to a contribution analysis to make credible causal statements on interventions and their observable results. The ToC is essential for assessing the OECD/DAC criteria and selecting hypotheses for the contribution analysis. At GIZ, ToCs are visualised in a results model and complemented by a narrative including corresponding hypotheses. A results model is a graphical representation of the project's ToC. It describes the logical connection and interrelationship of results (assumptions), and how and why they contribute to the overall objective. A results model defines intended positive results within the project, change hypotheses, including multidimensional causalities, system boundaries, assumptions and risks of the project (see Figure 2).

Overall project structure

The **project's objective** was to improve the prerequisites for compliance with labour standards in the province of Punjab by strengthening LHRD with regard to its capacities in occupational health and safety (OHS) prevention and inspection. A basic assumption was that poor compliance with labour standards has negative economic impacts on workers, enterprises, state institutions and society at large, and that compliance with labour standards is a matter of human rights. Giving visibility to both economic benefits of compliance and the negative economic impacts of poor compliance was considered a lever for motivating the stakeholders involved – both state and private sector actors – to take action.

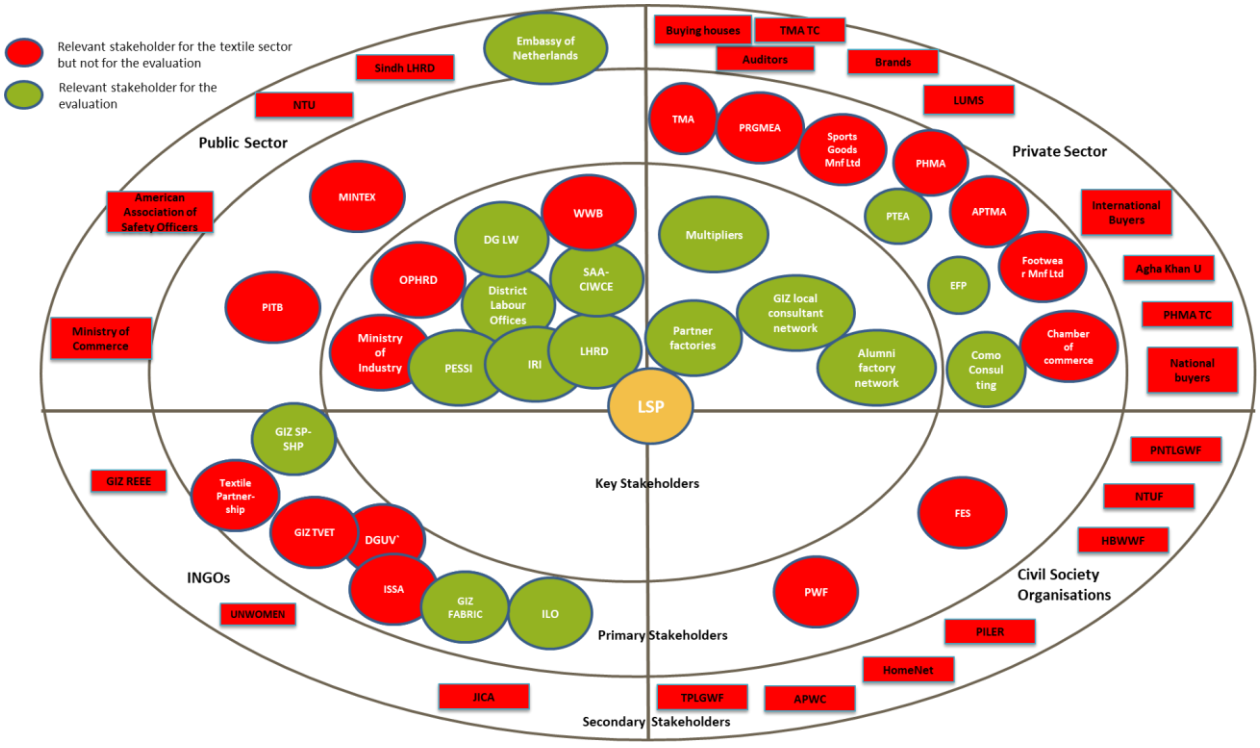
Based on the capacity development strategy of the project, the focus was on employees and the organisational capacity of LHRD as the lead executing agency of the project, as well as its downstream institutions and selected private sector actors (GIZ, 2018: 2). The Directorate General of Labour Welfare – including IRI and the Saeed Ahmad Awan Centre for Improvement in Working Conditions and Environment (SAA-CIWCE) – were the most important **implementing partners** at the public level. SAA-CIWCE operates at micro and meso

level and is responsible for education and training, research, information, and monitoring and advisory services (e.g. OHS audits). Other implementing partners were District Labour Offices of Lahore, Sialkot, Faisalabad and Multan. They are responsible for conducting labour inspections at micro level. In addition, the two institutions Punjab Employee Social Security Institute (PESSI) and the Punjab Workers Welfare Board (WWB), were involved in the implementation of the project. PESSI is the general health care provider and is responsible for issues regarding acute medical care and rehabilitation following work-related accidents and for compensation payments. The WWB is responsible for paying death grants (see public stakeholders on the stakeholder map Figure 1). The main actors from the private sector were (management of) partner factories and an alumni network of former partner factories. Business associations such as the Pakistan Hosiery Manufacturers Association and the Pakistan Textile Exporters Association were not part of the alumni network, but were relevant stakeholders at macro level. Cooperation with civil society was rather limited. Based on the Berlin Agreement, the Friedrich Ebert Foundation (FES) is working more closely with the trade unions (GTZ, 2004). Other stakeholders collaborating with the project were the GIZ project Support to Social Security including Health Insurance in Pakistan (PN 2015.2186.3), the Towel Manufacturers Association and the Employers Federation of Pakistan (EFP). The project also included service providers such as the German Social Accident Insurance and international consultants (Como Consult). Multipliers were four local consulting firms (Institute of Quality, NEC Consultants Pvt. Limited, Aftec Pvt. Limited Environmental Services Pakistan), one governmental service provider (Pakistan Institute of Management) and one non-governmental organisation (Pakistan Society for Training and Development).

Direct **target groups** of the project were LHRD as well as the management and employees of factories in the textile and garment industry, with a focus on the province of Punjab. Indirect target groups/final beneficiaries at impact level were employees in the textile industry.

The stakeholder map in Figure 1 represents the evaluation stakeholders in green and additional project stakeholders in red.

Figure 1: Graphical presentation of the project's stakeholder map (November 2020)



At **output level**, **output A** aimed to improve the capacities of LHRD and selected downstream authorities of the province of Punjab that are responsible for carrying out labour inspections. A basic assumption was that labour inspectors/officers[1] were trained only on labour laws and had merely a policing function. This would leave the inspectorate unable to identify hazards and risks and give hands-on advice on how to deal with these. The solution was seen as ensuring adequate qualification for inspectors, including training on risks and hazards, reflecting on their role, and building their advisory skills, including on-site training in partner factories (A1). Result A2, support to the SAA-CIWCE in developing campaigning material, was supposed to increase the outreach of SAA-CIWCE's prevention work, but also to reinforce efforts to support inspectors in their advisory role. These initiatives in A1 and A2 were intended to contribute to A3: the textile and garment industry is enabled to absorb advisory services pertaining to OHS.

Results under **output B** were equally a prerequisite for some of the results in output A. The organisational capacity of LHRD to uphold labour and social standards required capacity development. When suggesting that inspections should take place as often as necessary (more often in high-risk companies and less often in low-risk companies – A4), the downstream institutions under LHRD needed to build prevention and inspection on insights from institutions dealing with rehabilitation and compensation, the areas where the costs of poor working conditions were becoming apparent in the shape of accident victims. Therefore, informed decision-making as well as cooperation between the downstream institutions had to be strengthened. Establishing a system for monitoring accidents in the workplace at PESSI was supposed to allow labour inspectors to decide which are the high-risk industries that require more frequent visits than others; designing campaigns to address the most relevant risks and hazards was intended to give visibility to the costs of accidents and, as a result, produce an incentive to reduce these costs in the very long run; and an more low-hanging fruit was seen in the possibility to make labour inspectors and prevention experts realise the relevance, efficiency and effectiveness of their work for workers' lives as well as for the resources of state institutions, in this case PESSI's (B1). As well as simply monitoring accidents, there is a need to establish management information systems, in view of the numerous discrepancies and lack of coherence in the databases of the various LHRD downstream institutions (B2). An understanding of the interdependencies of the different downstream institutions was a requirement for encouraging PESSI, for example, to understand that their efforts in monitoring accidents would increase the impact of SAA-CIWCE's work and, in turn, be beneficial for PESSI. Sharing successes regarding efficiency and effectiveness in different forums will also help to build the reputation of state institutions and thus help to build stakeholders' commitment and support (B4). However, all downstream institutions needed to initiate a process of continuous improvement (B3). This would also entail management based on key performance indicators (KPIs), for which the collection of relevant information is a prerequisite. Continuous improvement of performance bottlenecks, based on cooperation, collection of information and monitoring, plus support from other stakeholders, would ultimately contribute to improving the organisational capacity of LHRD.

Regarding **output C**, partner factories had to see the benefits of improving social and environmental compliance for overall business performance. The Dialogue for Sustainability (DfS) was implemented as a tool to show companies how they could realise the benefits themselves, whether by mobilising internal resources and knowledge (C1) or knowing where external support was available (C3 and C4). Understanding the benefits of compliance from a business perspective was seen as the key element to increasing compliance overall. Workers with negotiation skills and the skills to assert their rights in a solution-oriented manner would further enhance opportunities and defuse existing deep-rooted conflicts (C2). For the past 12 months of the project term, and after a replenishment of the budget, result C5 was added to the result model. In view of the short time remaining, a realistic result was to develop the capacities of the private sector in the implementation of good environmental practices. Hence, the contribution to output C remains weak and is marked with a dashed

[1] For improved reader friendliness, in the following, the term 'labour inspector' is used. This also includes labour officers who are responsible for smaller companies, as opposed to labour inspectors. The project has involved both stakeholders in its activities.

line in the result model. Other results (C6 and C7) regarding good environmental practices were somewhat exploratory in nature, given that 12 months is too short a time to bring about real change.

At **outcome/impact level**, the project is a module of the development cooperation programme Promoting Sustainable Economic Development; the **programme objective** is to improve conditions for creating and securing employment and income opportunities that meet international environmental and social standards and that contribute to inclusive and sustainable economic growth. Both output A and output B were supposed to support state institutions in the good governance of labour standards. Hence, LHRD was supposed to improve conditions to maintain GSP+ status. The **programme's subobjective** – to promote social security, environmental and social standards through increasing compliance of employment and production conditions with environmental and labour standards – was supposed to be triggered by output C: private sector actors should succeed in increasing their compliance while boosting productivity. Both programme objectives should then contribute to Sustainable Development Goals (SDGs) 1, 3, 5, 6, 8 and 12, with SDG 8 on sustainable economic growth being the most prominent.

Additional information on the results model

System boundary

The system boundary is defined based on the scope of control of the project, i.e. results outside the system boundary are beyond the exclusive responsibility of the project and, indeed, are affected by other factors, stakeholders and interventions in the respective country. For the project, this meant that the programme objectives, the contribution to the SDGs, and the risks highlighted can be influenced to only a limited extent. The project's influence ended with its contribution to the prerequisites for compliance with social and environmental standards.

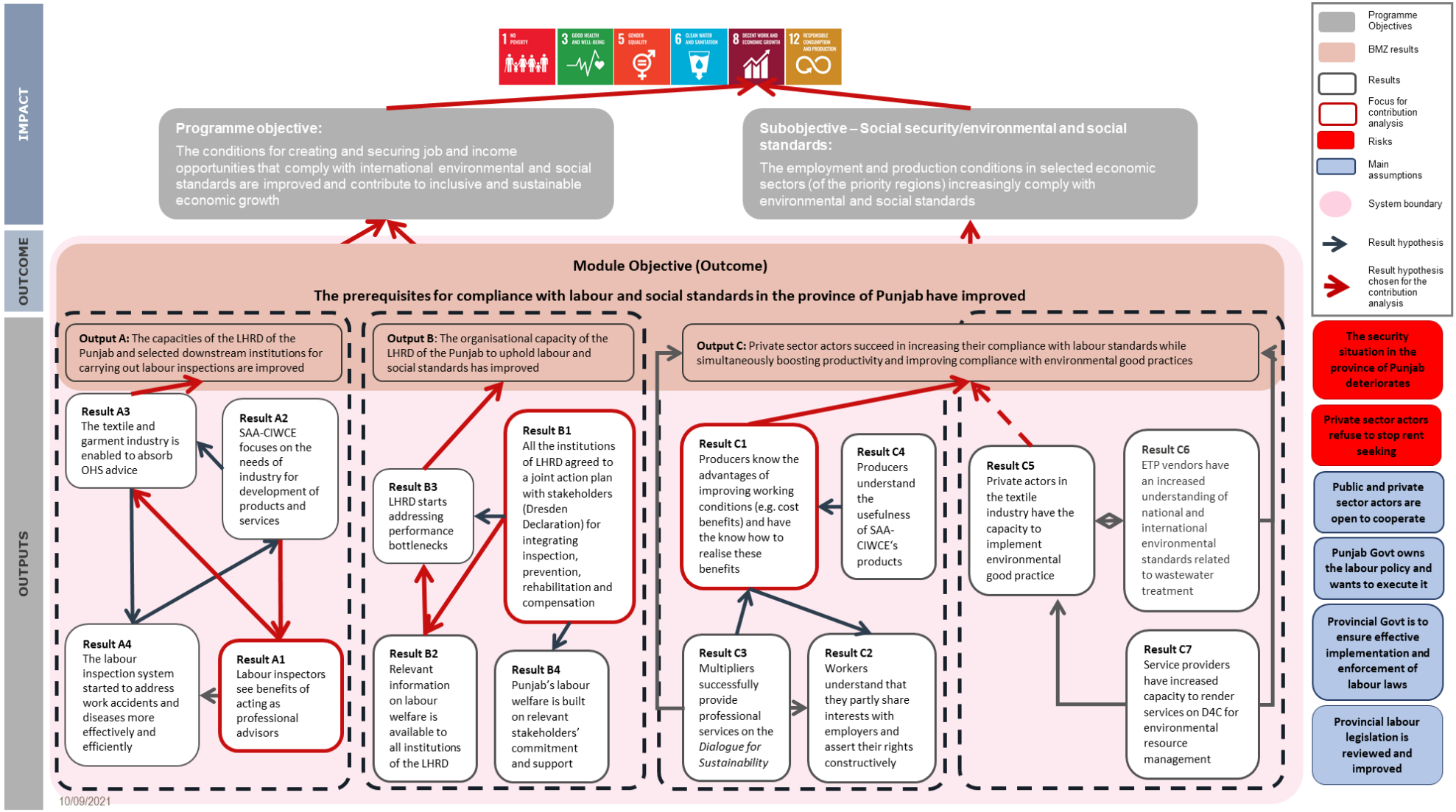
Unintended results and risks

Potential **unintended positive and negative results** at outcome/impact level were monitored by the project on an exploratory basis. Thus, unintended results are not yet included into the results model. The high staff turnover, especially at the managerial level of the provincial government, labour inspectors and factory management, has been identified as potentially impinging on the project's ability to cooperate with the partners and build capacity.

Potential interactions between social, economic and environmental results

Social, economic and environmental results are strongly related within the project structure and aligned with the SDGs. The project aimed to contribute to improved compliance with social, labour and environmental standards in the textile industry. Measures to increase productivity were supposed to contribute to improved social, economic and environmental results. As the factories have a poor efficiency into a final product, the project aimed to increase efficiency within the factories to reduce solid waste, save raw materials, improve working conditions and increase the competitiveness of the textile industry.

Figure 2: Adapted results model used for the evaluation (April 2021)



3 Evaluability and evaluation process

This chapter aims to clarify the availability and quality of data and the process of the evaluation.

3.1 Evaluability: data availability and quality

This section covers the following aspects:

- availability of essential documents,
- monitoring and baseline data including partner data, and
- secondary data.

Availability of essential documents

All relevant project documents, including the Capacity WORKS tools, were available to the evaluation team. Specific references to documents are made throughout the evaluation report and listed under the references. In particular, the project planning in the project proposal and the capacity development strategy as well as the reporting in the progress reports were compared with the actual implementation. Documents such as the stakeholder analysis and the results model were revised for the purpose of the evaluation, as described above. The operational plan and the governance structure were used to assess the quality of implementation.

Monitoring and baseline data including partner data

A results-based monitoring system (RBM) at project level was in place and well maintained. The project conducted monitoring at an operational and strategic level. All monitoring data was inserted into an Excel-based master sheet and uploaded to the GIZ online Results Monitor application. Monthly monitoring and evaluation (M&E) meetings were conducted with the project's technical team to track the activities and their expected output, including to check the status of the objective.

Each indicator was operationalised and well described through an Word document, 'Indicator Progress Update Sheet'. The templates consisted of a description and well-thought-through operationalisation of each indicator. A traffic light assessment showed whether indicators were on track, a forecast was given, and the status updated every three months. In addition, the source of verification, the challenges and risks, and the required current and future activities were described.

Baseline data

As monitoring data did not exist at partner level, one of the project's activities included setting up baseline data for the specific partners. Hence, the project's indicators are based mainly on a comparison of jointly developed partner KPIs between data collection at two points in time (baseline and endline). The monitoring was of the highest quality and provided a well-structured set of data for the evaluation team. There was no baseline study available or conducted for the project.

Partner and secondary data

The monitoring system depends on partner contributions. Project stakeholders participated in data collection (e.g. providing a 'List of recruitment' on outcome indicator 2). However, there was no written schedule showing when partners should provide data. Data was provided by the partners whenever the project asked for it. The process was perceived by the project to have worked well. Necessary partner monitoring data was already made available through the project's documents. As partners did not have reliable data, the project supported them to set up their own monitoring and quality control systems to increase evidence-based decision-making. Further national and secondary data was either not available or not considered trustworthy by the project. In addition, the indicators were too specific to use macro data from the national system.

3.2 Evaluation process

This section covers the following aspects:

- milestones of the evaluation process,
- involvement of stakeholders,
- selection of interviewees,
- data analysis process,
- roles of international and local evaluators,
- (semi-)remote evaluation (if applicable), and
- context and conflict sensitivity within the evaluation process (if applicable).

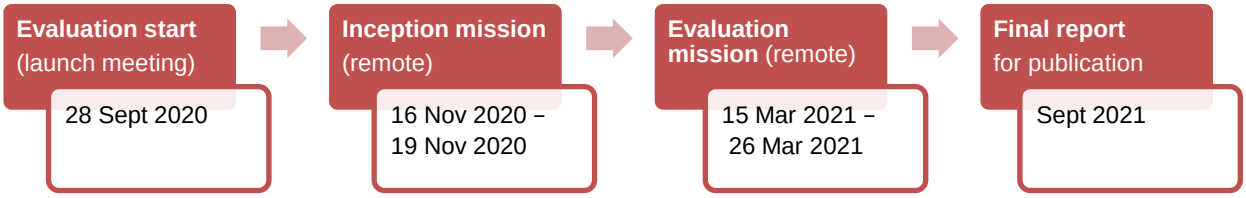


Figure 3: Milestones of the evaluation process

Involvement of stakeholders

The involvement of various stakeholders in the evaluation is central to CPEs. It strongly determines the success of the evaluation and acceptance of the evaluation findings and recommendations. The evaluation team initiated an activity with key project team members to map crucial stakeholders of the project and discuss their involvement in the evaluation. In addition, the evaluation team conducted remote interviews with representatives from BMZ and the GIZ sectoral unit as well as representatives from the project's alumni network to identify evaluation stakeholders. The red colour circles and boxes in Figure 1 show stakeholders who were not included in the evaluation process. The evaluation team also tried to involve civil society stakeholders to include the perspective of other external partners. However, these were either not available for interviews or absent from scheduled interviews. The final decision on who to involve in the evaluation was taken by the evaluation team, taking into account (i) the importance of the stakeholder, (ii) the value of (additional) information provided, and (iii) the feasibility of including stakeholders within the time frame/evaluation mission schedule. All in all, the number of stakeholders contacted during the evaluation mission was maximised. Table 2 lists the selected participants of the evaluation (disaggregated by gender). There was a special emphasis on focus group discussions with partner organisations (37 participants) and beneficiaries (18) with a view to collecting a wide range of opinions concerning effectiveness, impact and sustainability. Members of other GIZ projects and international stakeholders were included to assess the relevance and coherence criteria.

Selection of interviewees

Table 2: List of evaluation stakeholders and selected participants

Organisation/company/ target group	Overall number of persons involved in evaluation (including gender disaggregation)	No. of interview participants	No. of focus group participants	No. of workshop participants	No. of survey participants
Donors	0				
GIZ	9 (2 f, 7 m)	7	3	7	
GIZ project team, GIZ component managers, GIZ project management, GIZ country director, GIZ M&E, GIZ Support to Social Security Including Health Insurance in Pakistan (SP-SHP), GIZ regional programme Social and Labour Standards in the Textile and Garment Sector in Asia (FABRIC), GIZ Cluster Pakistan					
Partner organisations (direct target group)	43 (4 f, 39 m)	6	37	8	
Directorate General of Labour Welfare, IRI, PESSI, SAA-CIWCE, District Labour Offices in Lahore and Sialkot, partner factories, local consultants, multipliers, alumni factories network					
Other stakeholders (e.g. public actors, other development projects)	2 (2 f)	2			
ILO, Embassy of the Netherlands					
Civil society and private sector actors	3 (3 m)	3			
Pakistan Textile Exporters Association, EFP, Como Consultants					
Universities and think tanks	0				
Final beneficiaries/ indirect target groups (sum)	18 (2 f, 16 m)		18		
Factory workers	15 (2 f, 13 m)		15		
Labour inspectors	3 (3 m)		3 (3 m)		
Note: f = female; m = male					

Data analysis process

For efficient **data management and analysis**, the evaluation team compiled all qualitative findings from the documents and interviews using qualitative data analysis software (MaxQDA®). In the first step, field notes were taken during the actual interviews. The evaluation team used the on-site paper and pencil technique to identify first insights and recommendations. Once the interview was over, i.e. at the end of each day while impressions were still fresh, further notes were added. To analyse different data sources, a category system for evaluation questions, as per the evaluation matrix, was developed. Information from several data sources regarding a certain evaluation dimension could thus be retrieved and compared and the findings summarised. Quantitative monitoring data from the project's monitoring system was analysed mainly descriptively. Preliminary findings were then discussed with the project management during a validation workshop after the

official evaluation mission when data collection was completed. No online survey at partner factory level was conducted, despite being initially planned. Due to the availability of monitoring data and external monitoring reports, data was already available. Furthermore, project partners were reported to have participated in many online surveys before, so the risk of a survey fatigue led to the decision to use focus group discussions at partner factory level.

Roles of international and local evaluators

Mainlevel's evaluation team consisted of Felipe Isidor-Serrano (international expert), André Gersmeier (backstopping & technical support), and Khalid Mehmood (local expert). In general, tasks were divided as follows:

International evaluator: Felipe Isidor-Serrano (team leader):

- evaluation design and instruments,
- focal point for GIZ and the project team,
- responsible for successful implementation of inception and evaluation mission; implementation of virtual interviews with project team and stakeholders,
- data collection and analysis, and
- presentations and reporting.

Backstopping and technical support to the team leader: André Gersmeier:

- support for project management tasks,
- support for evaluation design and instruments,
- support for the implementation of data collection,
- researcher triangulation of results, and
- quality control.

Local evaluator: Khalid Mehmood

- technical expert in evaluation,
- regional expert in Pakistan in terms of understanding national legislation, policies, frameworks and international conventions or standards related to social and labour standards, OHS, environmental standards, and inclusion of persons with disabilities,
- planning of evaluation mission schedule,
- recruitment of interview partners,
- implementation of virtual interviews with (national) stakeholders and final beneficiaries,
- interpretation and triangulation of results with international evaluator, and
- reporting.

Teamwork, constant support and collaboration are an integral part of Mainlevel's company culture. During the evaluation mission, the evaluators were in regular contact to validate the data retrieved from interviews and discussions as well as the evaluation findings. Researcher triangulation helped to ensure a common interpretation and analysis of the available data.

Remote evaluation

The COVID-19 pandemic required the evaluation team to conduct the evaluation differently from the established procedures of on-site evaluations. While international travel continued to be restricted and quarantine rules were in place, travelling to Pakistan was not feasible. In addition, travelling within Pakistan was deemed to be too dangerous, according to the local evaluator's assessment of infection risks. The circumstances required (field) work to be conducted completely remotely. The above-mentioned infrastructure designed for the evaluation mission strengthened cooperation and quality assurance. Additionally, in order to

set up processes for virtual data collection, the consultants made use of communication software such as Microsoft Teams. It should be mentioned that remote data collection is always a challenge in terms of the openness of interview partners and focus group participants. Furthermore, not being on site makes it harder to read between the lines of discussions. The evaluation team tried to solve these challenges by holding regular exchange and debriefing sessions after interviews and discussions.

Context and conflict sensitivity within the evaluation process

As a result of the conditions during the COVID-19 pandemic, the entire evaluation process was conducted remotely. Therefore, despite Pakistan's rather conflictual national setting, the evaluation team was at no time at risk of violent, extremist-motivated attacks or political power conflicts. The evaluation team reflected on conflict sensitivity (do no harm) by closely reflecting on risks and potential unintended results with the project team. The evaluation team was always introduced to stakeholders by the project and was transparent about the evaluation process. Where needed, interviews with final beneficiaries were conducted in Pashto.

4 Assessment according to OECD/DAC criteria

In the following sections, information is provided on how the evaluation team evaluated the project against the OECD/DAC criteria. The sections should be considered in relation to the evaluation matrix. The tables present the relevant section of the individual criteria.

Given that most of the project's efforts were devoted to capacity building as a prerequisite for compliance with labour and social standards, the evaluation team used the Kirkpatrick framework on evaluating training effectiveness. The Kirkpatrick framework allows existing data to be analysed and to be combined with additional data collection on behaviour change and impact. The framework has four levels: **1. Reaction**, **2. Learning**, **3. Behaviour** and **4. Results** (Kirkpatrick, 2016). The first two levels mainly used existing data from training evaluations (for labour inspectors, factories) as an annex of the project's monitoring system. This was complemented with insights from qualitative interviews and focus group discussions. The levels mainly assessed the criteria of effectiveness (section 4.4) and impact (section 4.5).

4.1 Impact and sustainability of predecessor projects

No predecessor project was part of the evaluation. However, two previous projects – Implementation of Social Standards Support Programme to the Textile and Garment Industry of Punjab (PN 2013.9062.4), financed by the German Federal Foreign Office, and Water Efficiency in the Textile Industry (PN 2013.9773.6) – were considered predecessor projects by the project itself. How the project reacted to change is briefly considered in section 4.2 of the evaluation. However, it was not part of the OECD/DAC criteria assessment of the project and will therefore not be discussed in this section.

Photo 1: Second change management team workshop with companies (Source: GIZ Pakistan)

4.2 Relevance

This section analyses and assesses the relevance of the project Improvement of labour and social standards in the Pakistani textile industry, Pakistan (PN 2016.2029.3).

Summarising assessment and rating of relevance

Table 3: Rating of OECD/DAC criterion: relevance

Criterion	Assessment dimension	Score and rating
Relevance	Alignment with policies and priorities	30 out of 30 points
	Alignment with the needs and capacities of the beneficiaries and stakeholders	27 out of 30 points
	Appropriateness of the design	17 out of 20 points
	Adaptability – response to change	20 out of 20 points
Relevance total score and rating		Score: **94 out of 100 points** Rating: Level 1: highly successful

Belonging to the priority area of sustainable economic development, the project is fully aligned with the priorities of the BMZ country strategy for Pakistan. The project's strong focus on private sector development and synergies between public services is further aligned with Pakistan's Vision 2025. Through the project activities, shortcomings highlighted by the EU regarding the GSP+ were addressed. The project contributes to several SDGs, with SDG 8 on sustainable economic growth being the most prominent.

The project always built its measures on participative needs assessments and the self-defined KPIs of partners. Hence, private and public stakeholders' needs – such as capacity development for improved performance and compliance with labour, social and environmental standards – were met to a large extent. Participation of final beneficiaries (workers) and vulnerable target groups (female workers) was addressed. Human rights were an integral part of private sector collaboration, though less so at the public sector level, leaving room for improvement.

The project's design had a very holistic yet realistic and smart approach in combining productivity improvements with labour, social and environmental standard compliance. The design did not include environmental standards at public level due to the limited time for implementation and did not include international textile brands and buyers. A scaling-up and exit strategy can be considered initial steps that will need to be continued or taken up in the follow-on project. The project had to deal with a constantly changing environment of high staff turnover at project partner level, increasing demands from BMZ, contextual changes such as consideration of the abolition of the labour inspection system and political conflicts with India, and the effects of the COVID-19 pandemic. The project remained flexible and progressive despite the challenging environment and turned challenges into further opportunities (increased cooperation with the private sector, digitalisation).

In total, the relevance of the project is rated as Level 1: highly successful, with 94 out of 100 points.

Analysis and assessment of relevance

Relevance dimension 1: Alignment with policies and priorities

The relevance criterion was assessed mainly through analyses of secondary project data. Additional strategy documents and interviews with stakeholders were also considered. The analysis followed the analytical questions from the evaluation matrix (see Annex). Dimension 1 refers to the extent to which the objectives and the design of the project are aligned with global, country and institution-specific requirements. An assessment was conducted on the extent to which the project is aligned with BMZ's country strategy for Pakistan 2016–2020 (BMZ, 2016), Pakistan's long-term development strategy Vision 2025 (Ministry of Planning, Development & Reform, 2013) and Pakistan's strategy for Agenda 2030 (Ministry of Planning, Development & Reform, 2019). Further documents from BMZ (Bündnis für nachhaltige Textilien, Gemeinsame Fortschritte im Textilsektor and Zukunftspapier: Gute Arbeit weltweit) were considered relevant strategy documents by the project (BMZ, 2015, 2015a, 2018). In addition, priorities of the European Commission regarding compliance with the GSP+ were reviewed (European Commission, 2020).

Based on the BMZ country strategy for Pakistan (2016–2020), governance, energy and sustainable economic development are the priorities of German international cooperation with the country (BMZ, 2016: 3). The textile sector is highlighted as having a key role in sustainable economic development and in moving towards achieving the status of an upper middle-income country (BMZ, 2020: 6). The project belongs to the priority area of sustainable economic development. Based on the project's proposal (GIZ, 2019a) and the progress report (GIZ, 2020d), the project is fully aligned with the priorities of the BMZ country strategy. However, it was highlighted that the country strategy was developed retroactively on the basis of existing programmes and projects and therefore did not act as a guiding document of the project (Int_1).

At national level, following its development strategy Vision 2025, Pakistan aspires to join the league of upper middle-income countries. The strategy is further operationalised through seven pillars encompassing energy, water and food security, increased competitiveness, private sector and entrepreneurship-led growth, and democratic governance. The project's strong focus on private sector development includes several of the pillars mentioned (Ministry of Planning, Development & Reform, 2013: 10). Good governance occupies a prominent position in Pakistan's Vision 2025. The Vision is being implemented through sector strategies, largely by the provinces. The project's strategy embraced this; it included the different District Labour Offices as implementing partners for labour inspection and supported the implementation of the Punjab Labour Policy (2015 and revised in 2018). In addition, the project supported the process of decentralisation and contributed to synergies between public institutions (e.g. Dresden Declaration, monitoring of accidents in the workplace (accident monitoring system), quality management system).

At the EU level, Pakistan was granted GSP+ status by the EU in 2014 for up to 10 years, provided the country demonstrates successful efforts in the implementation of 27 international conventions and agreements pertaining to labour and environmental standards. The GSP+ provides access to European markets. In order to maintain its GSP+ status, Pakistan is required to implement core labour and social standards alongside relevant international conventions and agreements in the area of good governance and sustainable development. The European Commission considers the labour inspection system in Pakistan to be weak and to be low on the agenda of priorities (European Commission, 2020:4). Factories confirmed the need to improve the labour inspection system, in particular concerning the job description and role of labour inspectors (FGD_5). The project was aligned with improving the labour inspection system through capacity building. In addition, private sector development towards increased compliance with labour and social standards and productivity was connected to increasing textile exports to EU countries (FGD_10). Many other GSP+ priorities (e.g. freedom of expression, prevention of torture, and reducing the scope of domestic violence and forced, bonded and child labour) were less part of the project's design (Int_11).

At the international level, the project focused on the contribution to several SDGs (1 – no poverty, 3 – good health, 5 – gender equality, 6 – clean water and sanitation, 8 – good jobs and economic development and 12 – responsible consumption). There was focus on SDG 8 – to 'promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all' – with direct reference to Goal 8.8 on protecting labour rights and promoting safe and secure working environments for all workers. The 2019 added component on environmental standards (see dimension 4) also contribute to SDG 6 (clean water and sanitation) as well as SDG 12 (responsible consumption and production). Based on Pakistan's strategy for the 2030 Agenda the textile sector is especially mentioned in relation to economic growth and to GSP+ status boosting trade and productivity (Ministry of Planning, Development & Reform, 2019: 53).

Figure 3: The project's contribution to the SDGs

Overall, the evaluators concluded that the project is closely in line with international and national strategic reference frameworks and the objectives of Agenda 2030. Therefore, the maximum score is given.

Relevance dimension 1 – Alignment with policies and priorities – scores **30 out of 30 points.**

Relevance dimension 2: Alignment with needs and capacities of the beneficiaries and stakeholders

To understand the project's relevance for the target groups, perceptions were gathered through interviews and discussions. For triangulation, these were then contrasted with the assessed needs from the project documents. The analysis follows the analytical questions from the evaluation matrix (see Annex).

The project followed a multilevel, multiactor approach. The partner system included government structures at the macro level; textile and garment factories, downstream governmental institutions, regional departments and local consultants at the meso level; and labour inspectors and factory workers at the micro level. The Directorate General of Labour Welfare of LHRD, including IRI, SAA-CIWCE and PESSI, were the most important implementing partners at the public level. SAA-CIWCE is responsible for education and training, research, information, monitoring and advisory services (e.g. OHS audits). IRI is responsible for the training of the labour inspectors of the District Labour Offices of Lahore, Sialkot, Faisalabad and Multan, which conduct

labour inspections at micro level. PESSI is responsible for issues regarding acute medical care and rehabilitation following work-related accidents, and for compensation payments.

Overall, all public stakeholders confirmed that OHS is one of their weak areas where support from outside was needed (Int_11, Int_9, Int_2). Capacity development was regarded as a focus area by public stakeholders as it had been seen as low priority in the past (Int_12). Stakeholders confirmed that training curricula on labour inspections were outdated and lacked environmental assessment, and therefore needed revision. Other project approaches, such as the Vision Zero Campaign and the 'return on prevention' concept, were praised and said to have been highly relevant for connecting with the private sector (Int_9). Public stakeholders confirmed that they were not fully aware of their core problems and how to approach them in a targeted manner before cooperating with the project. The project conducted needs assessments to assess the capacity of staff and identify potentials for improvement so that capacity development measures could be implemented in a targeted manner. The participative needs-based approach was therefore praised for identifying blind spots (Int_2). Thus, the project was perceived as progressive compared with long-standing policies and habits that had not changed at factory level, such as the relationship between the management and employees (Int_4). Overall, the project had a stronger focus on the needs of the downstream departments responsible for implementation (e.g. IRI, SAA-CIWCE, PESSI) than on the administrative and policy level of LHRD (FGD_10). Frequent changes in personnel at LHRD led to challenges between individual priorities and long-term strategies. Improvements aimed at maintaining GSP+ status were appreciated by textile factory representatives, yet stakeholders highlighted that the administrative level lacked awareness of the importance of the GSP+ (Int_12).

The main partners from the private sector were (the management of) partner factories and an alumni network of former partner factories. Factories can be divided into three categories: small, medium and large. The majority of factories (80%) produce for local and regional markets, while only a small share of big factories export to European countries (Int_14). Medium-sized factories in particular were considered by the project to have potential for growth and for increased exports through international buyers (Int_1). Overall, factories confirmed that the industry lacks compliance with labour, social and environmental standards. Compliance has always been considered as a burden involving increased costs and efforts for factories (FGD_5). Furthermore, it was confirmed that the factories lacked capacities and strategies to improve their compliance and their productivity. While implementing the methodology of DfS, factories were supported to increase their productivity and become more compliant with labour, social and environmental standards (FGD_10; GIZ, 2018: 3). Governmental partners such as labour inspectors were perceived to lack skills and to be unable to provide vital support for capacity building to the private sector. Again, the project's participative and needs-based approach (measuring success against self-defined KPIs) was highlighted by all the factory representatives interviewed (FGD_2, FGD_4, FGD_5, FGD_6). Nevertheless, the project worked mainly with selected factories where motivation and potential for improvement were observed. Smaller factories without international exports were perceived as being committed less to compliance and more to increased productivity (FGD_6). Yet as the project focused on compliance through increased productivity, the approach was relevant to all types of factories.

Cooperation with **civil society** was rather limited. Civil society in Pakistan was considered by the project to be weak: few trade unions exist and are often subject to restrictions influencing their activities. Although usually a relevant partner for development cooperation projects, based on the Berlin Agreement (GTZ, 2004), FES was working more closely with the trade unions, leaving the project to focus on the private and public sectors to avoid duplication of activities.

Indirect target groups/final beneficiaries were employees in the textile industry and labour inspectors. The project focused on the province of Punjab (including the cities of Multan, Sialkot, Faisalabad and Lahore). Factory workers were part of the project's approach, participating in change management teams. They thereby improved their own capacities and were also regarded as final beneficiaries of improved compliance with

labour, social and environmental standards. As the evaluation had to be conducted remotely, the evaluation team could make contact with only a small number of factory workers. Based on that sample, it was observed that workers were not aware of their rights and especially lacked knowledge of how to treat chemicals. It was confirmed that the project met their needs. In addition, many workers are to some extent illiterate, and the project's practical approach using graphical posters and animated videos and including the posters in practical discussions was appropriate (FGD_1).

Disadvantaged and vulnerable beneficiaries (leave no one behind): According to the project proposal, the project aimed to enhance the expertise of female inspectors and experts and to address gender discrimination. Gender is a cross-cutting social and cultural issue in Pakistan. The project always tried to include women, closely monitored the potential for gender mainstreaming and tried to increase the gender balance. The needs of female labour inspectors and workers were taken into account in new strategies and documents (revision of curriculum) and female participation in change management teams was addressed. However, the project realised that the influence at the gender level was limited due to the low number of female employees and labour inspectors overall (GIZ, 2018c).

Human rights: Pakistan has ratified seven out of nine fundamental human rights agreements, but is not discharging its human rights obligations – including economic, social, cultural and political/civil rights – effectively. Human rights, especially work-related human rights, is a cross-cutting issue within the textile industry. The project tried to improve work-related human rights through strengthening the supervisory and regulatory role of state institutions and improved compliance with social and labour standards by textile and garment factories (Int_1, FGD_10). Furthermore, through change management teams in textile factories, employees were informed about their rights, and relationships between employees and employers were promoted. Factories confirmed that the project put a high value on the participation of vulnerable groups and women in change management teams at factory level (FGD_6). Cooperation with the private sector in developing procedures and methods for meeting its responsibility for human rights was geared towards international standards and guidelines on minimum standards and human rights (e.g. UN Guiding Principles on Business and Human Rights, UN Global Compact). The project was based on international ILO conventions and Articles 6 and 7 of the United Nations International Covenant on Economic, Social and Cultural Rights.

The evaluation team concluded that needs of the stakeholders were met. Different stakeholders expressed appreciation for the needs-based participatory approach. Stakeholders were not fully aware of their needs, and the project supported them in streamlining and prioritising their needs. Human rights were an integral part of the private sector cooperation, though less so at the public sector level, leaving room for the implementation of human rights conventions.

Relevance dimension 2 – Alignment with the needs and capacities of the beneficiaries and stakeholders – scores **27 out of 30 points.**

Relevance dimension 3: Appropriateness of the design

The basis for the assessment of the appropriateness of the design is the revised results model (see above) as well as the capacity development strategy (GIZ, 2018) and interviews with the project team. The analysis follows the analytical questions from the evaluation matrix (see Annex). The updated results model formed a solid base for the evaluation and contribution analysis. The results model was last updated in March 2020, which is why the results model was largely adopted for the evaluation. In a participatory exercise, the results model was discussed with the project team to identify areas for improvement. Ultimately, only minor changes had to be made to specify results hypotheses.

Overall, the project's objective to improve the prerequisites for compliance allowed for a broad target achievement but was further operationalised through specific indicators. The results model expresses the project's system boundary adequately, as the project's influence at the public level ends with organisational and individual capacity development. Further results at the behavioural level (e.g. behaviour of labour inspectors) to improve sustainable economic development were not part of the project. Module objective indicator (MOI) 1 reflects the multilevel approach to include state, private and civil society actors. Moreover, the assumptions and results hypothesis in the results model (see section 2.2) represent the dependency structures between different stakeholders. Based on the capacity development strategy, the focus of capacity development was on employees' capacity and the organisational capacity of LHRD as the lead executing agency of the project, as well as its downstream institutions and selected private sector actors (GIZ, 2018: 2). The project's bottom-up approach focused on the operational level to improve the results at the administrative level. However, the project team and external stakeholders considered the influence on wider administrative decision-making to be limited. In particular, changes in personnel and a rather project-oriented culture remained a challenge (Int_1, Int_14).

The private sector played a prominent role in the project design (MOI 3 and 4 and output 3). The project design already incorporated increased compliance (in terms of behaviour and results) combined with increased productivity as a prerequisite. The project built on good practices in the private sector as a lever to achieve an increased recognition of the need for action on the public side (in pursuing an improved labour inspection system) as well as among further private sector actors. Hence, the project was designed to create interlinkages between private and public sector actors and to foster trust (Int_1). Interaction between the social, environmental and economic dimensions was at the core of the project design. Improved compliance with labour standards, combined with measures to increase productivity, were supposed to contribute to the competitiveness of the project's partner companies and ultimately to secure jobs and improve sustainable economic development. Thus, the strategy was always aligned with economic benefits (FGD_10). Project partners regarded that approach as the only working strategy for sustainable change (Int_4): 'From Day 1 the message was given that we are not doing any audit and we want real efforts in the factories. If we work together, we have real benefits based on real implementation' (FGD_5). Environmental good practices were not only a cross-cutting issue but also a project objective (output C, indicator C3) that was addressed with the private sector. In particular, the high consumption of energy and resources in the textile industry (e.g. raw material, water, energy), the use of (toxic) chemical substances and the inadequate wastewater and sewage sludge treatment in textile production are highly damaging to public environmental goods (GIZ, 2018c: 3). As the factories have a poor efficiency into a final product, the project aimed to increase efficiency within the factories to reduce solid waste and save money. The environmental aspects were included only towards the end of the project and were focused on the private level. Therefore, the design did not include public sector engagement on improved compliance with environmental standards.

Other international projects highlighted the realistic objectives of the project (Int_14, Int_6). Nevertheless, within the project, phase scalability and an exit strategy at the private sector level can be considered as first steps that will need to be continued or taken up in the follow-on project (Int_1). The project design was not intended to be considered fully sustainable without further follow-on activities. The project was praised for its holistic approach to addressing the working environment, workers' occupational health and safety, accident prevention, improved monitoring of work-related accidents, and overall governance, for the potential benefit of all organisations and stakeholders involved (Int_16). Other influences outside the project's sphere of responsibility (e.g. international textile brands and buyers) were considered to a lesser extent, and public–private partnerships, such as develoPPP partnerships, were not pursued. The project perceived the absence of international brands and buyers as an advantage that allowed for an increase in intrinsic motivation of productivity and a unique selling proposition, in contrast to extrinsic motivation, which is very dependent on brands' requirements. Furthermore, the potential for 'green washing' by companies are perceived to be very high (Int_1). Nevertheless, other develoPPP projects have taken a similar approach to that of the DfS on

cooperation with international brands. This allows project activities to be further tested in an existing supply chain. As international brands are one of the most important stakeholders for export-oriented markets and the German as well as a European Supply Chain Transparency Act are approaching, further cooperation will need to be considered.

Overall, the evaluation team concluded that the project's design was highly appropriate due to its holistic, realistic and smart approach in combining productivity improvements with labour, social and environmental standard compliance. However, the design lacked further strategies for public-level engagement, and its scaling-up and exit strategy needed further attention. Only three points were deducted as the design did already provide for a follow-on project, in which the sustainability of results will play a bigger role (Int_1).

Relevance dimension 3 – Appropriateness of the design – scores **17 out of 20 points.**

Relevance dimension 4: Adaptability – response to change

The assessment was based on three change offers (GIZ, 2017a, 2018b, 2019a), yearly progress reports (GIZ, 2018f, 2019b, 2020d) and interviews with project management and project staff. The analysis follows the analytical questions from the evaluation matrix (see Annex).

The project submitted three change offers. In 2017, at the request of BMZ, the project was extended by a financing agreement of EUR 1,900,000 with the ILO. This did not result in any changes to the project's content or procedures (GIZ, 2017a). In 2018 a second change offer was based on new commitments and included major conceptual changes (new module objective, results and indicators) as well as an extension of the term to December 2020 and an additional EUR 6,000,000. The duration of the project was therefore extended to four years and the volume increased to EUR 10,050,000 (GIZ, 2018b). In November 2019, BMZ increased the volume of the project again to include initiatives from the project Water Efficiency in the Textile Industry (PN 2013.9773.6), which was completed in 2018. This change was made to include the environmental component of labour and social standards. Thus, output C was extended to include the use of best environmental practices. Accordingly, indicator C3 was added to the results matrix and the results model was updated (GIZ, 2019a).

Due to the constant changes and increasing requirements of BMZ, the project was forced to apply a very iterative approach (Int_17, Int_1). Nevertheless, the project managed to update relevant strategic documents (e.g. results model and stakeholder map) and remained very confident of always having an alternative strategy (Int_1). Although it was very demanding for the project, the constant revision of strategies led to a more extensive monitoring of indicators, risks and opportunities than is usually needed for project steering and reporting. However, the project also reported challenges, as the constant replanning required more flexibility and effort than would be needed for a longer-term strategy (Int_1, Int_17).

Keeping track of the steering structure was a challenge not only because of the change offers but also because of a high turnover in secretaries and commissioners at Directorate General of Labour Welfare. The project provided updated documents on steering structures and committee meetings and was reported to have updated relevant stakeholders (FGD_10, Int_1). Project partners confirmed that they have been updated well.

From 2019 to February 2020, the government gave consideration to abolishing labour inspections in Punjab. The project responded by engaging in discussions with export-oriented private sector (partner) companies, the Pakistan Employers Federation and the Ministry of Commerce to prevent the abolition of labour inspections. This rather informal approach led to the signing of the so-called Dresden Declaration as a plan of action to improve the labour inspection system. Although discussions on abolishing labour inspections decreased, they remain a risk (GIZ, 2019b: 9).

From mid-March 2020, due to the COVID-19 pandemic, activities could only be implemented remotely, so events were organised virtually. Furthermore, the implementation of individual activities slowed down due to delays in the issuing of visas for international consultants, including the consultant for the introduction of a quality management system (indicator B2). The international experts therefore reduced their travelling and used digital tools to cooperate with their project partners in Pakistan. The costs saved from the reduction in travel by international consultants were used for further training of the project team in digital moderation skills (Int_8). Project partners appreciated that the project was still active, increased their virtual interaction and adjusted methods to their needs (e.g. strategies for factories if no orders came in (FGD_4, FGD_2, FGD_10, Int_4)). The project's approach of relying on local consultants and further investing in their capacities was also noted (Int_17).

Overall, the evaluation team concluded that the project remained flexible and progressive despite the challenging environment. Weaknesses in terms of loss of efficiency or reduced implementation because of the constant changes remain plausible, but were not reported. The project has built up a good relationship with the private sector in response to the unstable staffing situation in the public sector. Full points are given.

Relevance dimension 4 – Adaptability – response to change – scores **20 out of 20 points.**

Methodology for assessing relevance

Table 4: Methodology for assessing OECD/DAC criterion: relevance

Relevance assessment dimensions	Basis for assessment	Evaluation design and empirical methods	Data quality and limitations
Alignment with policies and priorities	• BMZ country strategy for Pakistan (2016–2020) • Pakistan Vision 2025 • Documents on EU GSP+ with Pakistan	**Evaluation design:** The analysis follows the analytical questions from the evaluation matrix (see Annex) **Empirical methods:** Document review and criteria-led analysis	• Quality of data: good • No limitations
Alignment with the needs and capacities of the beneficiaries and stakeholders	**Direct target group:** • Staff of LHRD and subsidiary institutions (IRI, SAA-CIWCE, PESSI) • Implementing partners from the District Labour Offices • Management of factories in the textile sector **Indirect target group** • Employees in the textile sector, with a focus on the province of Punjab	**Evaluation design:** The analysis follows the analytical questions from the evaluation matrix (see Annex) **Empirical methods:** Document analysis, interviews and focus group discussions	• Quality of data: moderate • Due to remote evaluation design, access to final beneficiaries was limited
Appropriateness of the design*	Results model (including results hypotheses) Capacity development strategy	**Evaluation design:** The analysis follows the analytical questions from the evaluation matrix (see Annex) **Empirical methods:** Document analysis, interviews	• Quality of data: strong • No limitations

Relevance assessment dimensions	Basis for assessment	Evaluation design and empirical methods	Data quality and limitations
Adaptability – response to change	Three change offers from 2017, 2018 and 2019	**Evaluation design:** The analysis follows the analytical questions from the evaluation matrix (see Annex) **Empirical methods:** Document analysis, interviews	• Quality of data: strong • No limitations

* The project design encompasses the project's objective and ToC (GIZ results model, graphic illustration and narrative results hypotheses) with outputs, activities, instruments and results hypotheses as well as the implementation strategy (e.g. methodological approach, capacity development strategy, results hypotheses).

Conflict sensitivity in the project design

Within the integrated context analysis, the project is considered within category B, which means that the context is characterised by fragility, conflict and violence (GIZ, 2018e). Risks were frequently monitored by the project, reported in the specific progress reports, and outlined in the project's PCA-Matrix (GIZ, 2018e) and the brief political economy analysis Pakistan 2020 (GIZ, 2020e). Although the general security situation was assumed to have improved during the project period, an increasing risk of escalation in the conflict between Pakistan and India was perceived (Int_17).

Table 5: Dividers/escalating factors in the project context

Which dividers/escalating factors were identified in the project context?	Addressed by the project? (yes/no)	If addressed, how is it considered by the project design?
Deeply rooted in the feudal tradition – political system is resistant to change	Yes	Through a bottom-up approach in cooperation with LHRD
Feudal traditions in the private sector provide room for exploitative relations between employers and employees	Yes	By fostering dialogue between workers and employers and promoting labour standards
Tension with India and Afghanistan	No	
The experience of exploitation increases receptiveness to radical ideologies; lack of adequate education and employment opportunities, particularly for young men, combined with radical views and poverty, create long-term potential for violent extremism	Yes	By improving the performance of private sector actors, leading to sustainable economic growth and jobs

Table 6: Connectors/deescalating factors in the project context

Which deescalating factors/connectors were identified in the project context?	Addressed by the project? (yes/no)	If addressed, how is it considered by the project design?
Stabilisation of the security situation in the affected regions and provision of security for all citizens	Yes	• By supporting increased compliance with international labour standards to increase and reduce experience of deprivation • Through support to increased productivity as a basis for increased wages • By supporting dialogue between employers and workers to solve conflicts on the production floor

4.3 Coherence

This section analyses and assesses the coherence of the project Improvement of labour and social standards in the Pakistani textile industry, Pakistan (PN 2016.2029.3).

Summarising assessment and rating of coherence

Table 7: Rating of OECD/DAC criterion: coherence

Criterion	Assessment dimension	Score and rating
Coherence	Internal coherence	35 out of 50 points
	External coherence	40 out of 50 points
Overall score and rating		Score: **75 out of 100 points** Rating: Level 3: moderately successful

The project was designed somewhat in isolation, as other projects in Pakistan did not focus on the textile industry or on textile workers. Nevertheless, internal coordination at operational level was running well and no parallel structures existed that led to problems with partners. Synergies were achieved in the area of monitoring occupational accidents, and knowledge sharing took place as part of the regional project FABRIC. However, the lack of strategic coherence resulted in the absence of further synergies in the field of sustainable economic development in Pakistan.

International partners are not very active in Punjab, though they are in other regions. External coherence at donor level is based on occasional discussions but lacks systematic coordination. No negative issues or parallel structures were found and communication between international stakeholders was working well. Nevertheless, no further synergies were realised for the project, although other projects were able to use the project's experiences and lessons learn for their own project design.

In total, the coherence of the project is rated as Level 2: successful, with 75 out of 100 points.

Analysis and assessment of coherence

Coherence dimension 1: Internal coherence

Internal coherence addresses the synergies and division of tasks between the intervention and other interventions of German development cooperation and also the intervention's consistency with the relevant international norms and standards to which German development cooperation adheres. The assessment was

based on interviews with bilateral and regional GIZ projects in Pakistan as well as programme documents (GIZ, 2017b). The analysis follows the analytical questions from the evaluation matrix (see Annex).

As previously reported, the project was designed somewhat in isolation, as other projects in Pakistan did not focus on the textile industry or on textile workers (Int_1). Due to vacancies at BMZ and in the German Embassy, the alignment of the different project designs was perceived as weak, and synergies were not well considered in the planning (Int_1, Int_5). Projects appeared to be fitted into a country strategy, and a joint programme was retrospectively rather than proactively planned for. As such, synergies in the training of skilled workers to strengthen the competencies of future employees were not pursued (Int_13).

However, project managers confirmed that internal coordination at the operational level was running well, and no parallel structures existed that led to problems with partners (Int_17). In particular, the cooperation with the social protection project worked well as the project could build on existing partners such as the Agha Khan University (Int_15). Both this project and the social protection project were cooperating with PESSI and confirmed that coordination was running well (Int_15): synergies were achieved in the areas of monitoring occupational accidents and illnesses and access to health care services (Int_13).

As part of the regional dialogue and knowledge sharing with FABRIC, the project team presented different cases and examples of good practice at a number of conferences. They presented on their approach to working on the business performance and energy efficiency of textile factories at a regional conference in Myanmar; on activities for the development of management information systems at a joint workshop in Cambodia; and at a conference on labour inspectorates with stakeholders from Bangladesh, Pakistan and Cambodia (GIZ, 2020d: 13; Int_5). Although other projects were positive about the internal coherence, the project itself failed to achieve further synergies and alignment (Int_1).

Internal coherence with the global and sector projects Promotion of Multi-Stakeholder Projects for Sustainable Textile Supply Chains, and Sustainability in Textile Supply Chains were not a focus area of the evaluation. However, project documents and interview partners did not mention any cooperation or synergies. The follow-on project proposal already puts a stronger focus on internal coherence and synergies (GIZ, 2020g: 3ff).

Overall, the evaluation team concluded that the project managed to identify synergies based on upcoming opportunities, and the internal coordination and communication was highlighted by other GIZ projects (Int_13, Int_15, Int_17). Nevertheless, the lack of strategic coherence has resulted in the absence of further synergies in the field of sustainable economic development in Pakistan (e.g. technical and vocational education and training (TVET)). Further potential synergies were already mentioned in the first project proposal in 2016 (GIZ, 2016), but were not further realised. Therefore, the evaluation team deducted 15 points from the score for internal coherence.

Coherence dimension 1 – Internal coherence – scores **35 out of 50 points**.

Coherence dimension 2: External coherence

External coherence considers the intervention's complementarity, harmonisation and coordination with the interventions of other partners, donors and international organisations. The criterion relates both to the intervention's design and to the results achieved. The assessment was based on interviews with staff of external development cooperation projects and other stakeholders. The analysis follows the analytical questions from the evaluation matrix (see Annex).

Overall, other international partners are not very active in Punjab, thought they are in other regions. Other relevant projects include the Buyers' Forum, funded by the Government of the Netherlands as a platform in

which international brands and buyers from Pakistan and Bangladesh participate, together with the ILO and the Ministry of Commerce. In addition, the EU-funded project International Labour and Environmental Standards Application in Pakistan's SMEs is being implemented by ILO and the World Wide Fund for Nature, but is focusing on Sindh.

Overall, the external coherence at donor level is based on occasional discussions and lacks systematic coordination (Int_1). The project was not a member of the Buyers' Forum, although it did present its approach and results to the forum (GIZ, 2018f: 17). However, the evaluation team is not aware of any follow-on activity or result based on this presentation. The project has established good working relations with the ILO project working in Sindh, which led to close coordination at operational level. Other projects are reported to have made use of the project's experience and built their activities on lessons learned from Punjab (Int_6, Int_14). As other external projects work in different regions, no competition or any other negative issues were raised to the evaluation team; rather, activities had led to a mutual exchange of experiences (Int_1, Int_14).

Regarding the extent to which the project complements and supports partners' own efforts (principle of subsidiarity), it should be mentioned that, as explained above, the project built its activities based on the needs of the partners themselves. However, partners were not always aware of their challenges and how to address them, so downstream departments brought together different stakeholders to create a joint action plan (Dresden Declaration, Int_11, Int_9, Int_2). Still, the project built its activities to a great extent on downstream departments' existing responsibilities, improving its partners' capacities to implement existing policies such as the Punjab Labour Policy.

As part of an agreement between BMZ and ILO, the project administered a financing agreement with ILO of EUR 1,700,000. The agreement aimed to compensate those affected by the fire at the Ali Enterprises textile factory in Sindh. The agreement had no effect on the results or procedures of the project design. In addition to the financing contract, the service package for ILO included the personnel costs and other expenses necessary for its implementation (GIZ, 2017a). It should be mentioned that the financing agreement is not related to the ILO project International Labour and Environmental Standards Application in Pakistan's SMEs mentioned above.

Overall, due to the lack of international stakeholders in Punjab, there was no significant coordination effort at the operational level. The coordination with FES (as described in section 4.2) set clear boundaries, but further synergies between international actors were not pursued. Nevertheless, the project remained active in communicating and exchanging with international stakeholders at an operational level.

Coherence dimension 2 – External coherence – scores **40 out of 50 points.**

Methodology for assessing coherence

Table 8: Methodology for assessing OECD/DAC criterion: coherence

Coherence: assessment dimensions	Basis for assessment	Evaluation design and empirical methods	Data quality and limitations
Internal coherence	Other GIZ projects within the textile cluster: • Supporting Technical and Vocational Education and Training Reform (PN 2016.2042.2) • Support to Social Security including Health Insurance in Pakistan (PN 2015.2186.3) • Regional programme Social and Labour Standards in the Textile and Garment Sector in Asia (FABRIC, PN 2014.2279.9)	**Evaluation design:** The analysis follows the analytical questions from the evaluation matrix (see Annex) **Empirical methods:** Interviews with GIZ projects in Pakistan	• Data quality: strong • No limitations
External coherence	Government of the Netherlands • Buyers' Forum • EU/ILO project International Labour and Environmental Standards Application in Pakistan's SMEs	**Evaluation design:** The analysis follows the analytical questions from the evaluation matrix (see Annex) **Empirical methods:** Interviews with external projects	• Data quality: strong • No limitations

4.4 Effectiveness

This section analyses and assesses the effectiveness of the project Improvement of labour and social standards in the Pakistani textile industry, Pakistan (PN 2016.2029.3). It is structured according to the assessment dimensions in the GIZ project evaluation matrix (see Annex).

Summarising assessment and rating of effectiveness

Table 9: Rating of OECD/DAC criterion: effectiveness

Criterion	Assessment dimension	Score and rating
Effectiveness	Achievement of the (intended) objectives	29 out of 30 points
	Contribution to achievement of objectives	24 out of 30 points
	Quality of implementation	18 out of 20 points
	Unintended results	19 out of 20 points
Overall score and rating		Score: **90 out of 100 points** Rating: Level 2: successful

All module objective indicators were fully achieved by the end of the project. Based on a new curriculum for training labour inspectors, capacity development – for example in labour law, communication skills and consulting – was achieved. Labour inspectors showed increased ownership and motivation to act as advisors for textile and garment factories in addition to their policing role. Further behavioural change could not yet be observed, and further practical experience is needed. Framework conditions are a threat to future behavioural change. Increased ownership for mutual collaboration and shared responsibility can be observed in the project's success in its role as moderator and mediator, setting the stage for collaboration that would not otherwise have happened. The DfS with related change management teams became a very successful

methodology for identifying problems, finding solutions, and defining action plans. Capacity WORKS has been an integral part of the project steering, communication and project management as RBM played an important role. As an unintended result, the informal exchange between private sector partners on best practices are proof of increased ownership and motivation.

In total, the effectiveness of the project is rated as Level 2: successful, with 90 out of 100 points.

Analysis and assessment of effectiveness

Effectiveness dimension 1: Achievement of the (intended) objectives

As a first step, the evaluation team assessed the extent to which the agreed project objective (outcome) has been achieved, measured against the objective indicators. This required a comparison between the current status and the targets of the outcome indicators. The evaluation team built on monitoring data and other primary data sources. During a qualitative content analysis, key project documents and relevant external documents were reviewed and examined for evidence regarding the indicators. The evaluation team further collected and triangulated perceptions from key stakeholders, including (i) the project team management and team members, and (ii) key partners and further project stakeholders. All indicators fulfil the SMART criteria (specific, measurable, achievable, relevant and time-bound), no adaptation was needed, and no further indicators were added. Table 10 provides an overview of indicator achievement, including baseline, current and target values. Further description is given below.

Module objective indicator 1: A plan of action for improving the state labour inspectorate, agreed with state, private sector and civil society actors is in place. A first agreement between multiple stakeholders – the Dresden Declaration – was signed in 2019 (GIZ, 2019c). The declaration contains four pledges: 1. collaboration between stakeholders; 2. recognition of the importance of OHS (Vision Zero); 3. action-oriented approach; and 4. trust, transparency and solution-oriented discussion to harmonise efforts to maintain Pakistan's GSP+ status and improve OHS (GIZ, 2019c: 2f). In 2020, the Dresden Declaration was further operationalised by a technical working group to draw up an action plan on the agreement and future proceedings (GIZ, 2020h). The Joint Plan of Action – Implementation of Dresden Declaration for Promoting Culture of Prevention contains four specific themes: 1. curriculum on labour inspection; 2. curriculum for safety officers on OHS; 3. joint plan of action; and 4. digital evidence-based decision-making solution (GIZ, 2020h). The curriculum for labour inspectors (1) and the digital monitoring system (4) were realised. Plans of action (3) for three thematic areas were developed and implemented. The fourth thematic area – the curriculum for safety officers on OHS (2) – was shifted to the follow-on project. The evaluation team concluded that the indicator has been fully achieved.

Module objective indicator 2: 75% of newly recruited labour inspectors from July 2018 have, within the first six months of their work, undergone the introductory curriculum developed within the framework of the project. The curriculum was developed by IRI with the support of the project and was accepted by LHRD in 2019 (indicator A1). Since July 2018, only 13 labour inspectors have been recruited; 2 left shortly after their recruitment; 9 out of the remaining 11 were trained on the basis of the revised curriculum. In March and May 2018, 12 labour inspectors were recruited. They were also trained on the basis of the new curriculum but were not included in the target group due to the cut-off date defined in the indicator. Only one of the newly recruited labour inspectors is female. In 2020, as a result of COVID-19, no further recruitment of labour inspectors took place. The evaluation team concluded that the indicator has been fully achieved.

Module objective indicator 3: 60% of 50 surveyed textile or garment companies confirmed that the products of SAA-CIWCE for OHS (e.g. campaigns, informational material, guidelines, training courses) were used. Based on the project's monitoring data, 61 companies were contacted; 50 factories replied to a survey and 34 companies (68%) confirmed that they had received products and services from SAA-CIWCE and that the

products met their industrial needs. In addition, 29 companies confirmed that they were currently using the products. The project outsourced an assessment to a local consultant to gauge the availability of SAA-CIWCE products in the targeted companies. In total, only 17 out of 53 factories responded to the survey (32%). Based on the sample, the majority of factories confirmed they had received products and services from SAA-CIWCE, and about half of the factories confirmed they were still using products and services from SAA-CIWCE. Out of 12 participants, all had received training, while two-thirds (8) had received guidelines and more than half (7) had received information materials/investigations. When asked which products and services were most useful, all of the 12 participants mentioned training, 9 mentioned information materials and 9 mentioned guidelines. The least useful were investigations (1 respondent), visits for technical assistance (2) and risk assessment facilities (4) (GIZ, 2021: 54–60). Due to the limited number of respondents, the project carried out a second assessment of objective indicator 3 in which the above-mentioned monitoring data showed that the indicator is almost fully achieved (to 97%). The evaluation team checked for data reliability and confirmed the achievement.

Module objective indicator 4: Five textile and garment producers have concluded contracts with private sector and/or public service providers (producer associations, chambers of commerce, consulting firms, universities, etc.) for advisory services based on the project's dialogue methodology, to ensure compliance with social standards while also increasing productivity. The indicator deals with the institutionalisation, dissemination and scaling of the project's methodology of DfS through private sector partners as multipliers of the project. The project implemented DfS in support of local consultants with 24 partner factories. In addition, two of the contracted external local consultants were able to conclude six contracts in total with private sector factories (based on contracts shown to the evaluation team).

The evaluation team concluded that project objective indicators 1, 2 and 4 were fully achieved by the end of the project, and that project objective indicator 3 was partly achieved (97%) by the end of the project.

Effectiveness dimension 1 – Achievement of the (intended) objectives – scores **29 out of 30 points.**

Table 10: Assessed objective indicators for specific modules (outcome level)

Project's objective indicator according to the (last change) offer	SMART* criteria assessment
1. A plan of action for improving the state labour inspectorate, agreed with state, private sector and civil society actors, is in place. Base value (2018): 0 Target value (2020): 1 Current value (2020): 1 Achievement in % (2020): 100 Source: GIZ (2019c), GIZ (2020h)	The indicator fulfils all SMART criteria.
2. 75% of newly recruited labour inspectors from July 2018 have, within the first six months of their work, undergone the introductory curriculum developed within the framework of the project. Base value (2018): 0 Target value (2020): 75% of newly recruited labour inspectors from July 2018. Current value (2020): 9 of 11 (82%) Achievement in % (2020): 109 Source: project monitoring data	The indicator fulfils all SMART criteria.
3. 60% of 50 surveyed textile or garment companies confirm that the products of SAA-CIWCE for OHS are used. Base value (2018): 0 Target value (2020): 60% of 50 surveyed textile or garment companies (i.e. 30). Current value (2020): 29 out of 50 respondents (i.e. 58%) Achievement in % (2020): 97 Source: project monitoring data	The indicator fulfils all SMART criteria.

Project's objective indicator according to the (last change) offer	SMART* criteria assessment
4. Five textile and garment producers have concluded contracts with private sector and/or public service providers for advisory services based on the project's dialogue methodology, to ensure compliance with social standards while also increasing productivity. Base value (2018): 0 Target value (2019): 5 Current value (2020): 5 Achievement in % (2020): 100 Source: Five contracts with textile and garment producers that were shown to the evaluation team	The indicator fulfils all SMART criteria.
* SMART: specific, measurable, achievable, relevant and time-bound	

Effectiveness dimension 2: Contribution to achievement of objectives

In a second step, a contribution analysis was conducted to assess the extent to which the activities and achieved results (outputs) of the project substantially contributed to the achievement of the project objective. Following Mayne (2012), a contribution analysis is based on six steps. The validated results model, including risks and assumptions, guided the analysis (Step 1). During a participatory exercise with the project management, three hypotheses (Table 12, Table 13, Table 14) from output to objective were identified (Step 2). Selection criteria for the hypotheses comprised key interests of the project team as well as the feasibility of implementing the contribution analyses in the given time frame. In Step 3, evidence for the outputs, influencing factors and conflicting explanations was collected and a contribution story (Step 4) was compiled. Qualitative data collection instruments comprised semi-structured interviews and focus group discussions with project partners at operational level. Collecting further evidence for alternative hypotheses (Steps 5 and 6) was not feasible within the scope of the evaluation. However, a validation workshop at the end of the data collection with the project team supported the validation of findings and revealed explaining factors. As described above, the Kirkpatrick framework comprising **four levels of analysis** at organisational level **– reaction, learning, behaviour change** and **results** – was used as an additional basis for analysis.

This section focuses on the analysis of the results hypotheses rather than on the actual results. Nevertheless, a brief overview of the results (see Table 11) is a good starting point. The achievement of results at the output level is a central prerequisite for testing the assumptions of the results hypotheses from output to outcome level. Therefore, based on the project's monitoring system, a triangulation of achieved outputs is needed at the level of partners and participants of the intervention. More than half of the output indicators (A1, B1, B2, C1, C3) were fully achieved. Three indicators were partly achieved (A2, A3, C2). No indicator was not achieved.

Table 11: Assessment of output indicator achievement

Indicators	Baseline	Target value	Indicator achievement
Output A: The capacities of the Labour & Human Resource Department and selected downstream authorities of the province of Punjab for carrying out labour inspections are improved.			
A1. The Industrial Relations Institute of the Labour & Human Resource Department of the province of Punjab has revised the curriculum for training labour inspectors.	0	1	Achieved: Revised curriculum with letter of endorsement by the secretary of LHRD is available.
A2. In a competency survey, 16 out of 20 labour inspectors of 4 selected labour authorities at district level achieve results 1 level better on a 5-level scale than 10 labour inspectors from a comparison group.	0	80% of 20	Partly achieved (50%): 5 out of 20 (40%) achieved 1 level better. Averages: • Post-assessment of the intervention group: 4.1 • Pre-assessment of a comparison group: 3.43

Indicators	Baseline	Target value	Indicator achievement
A3. The Labour & Human Resource Department's Saeed Ahmed Awan Centre for the Improvement of Working Conditions and Environment has conducted four campaigns on risks in the area of occupational health and safety in the textile and garment sector, of which one directly relates to the results of the analysis of work-related accidents by the Labour & Human Resource Department's PESSI.	0	4 campaigns, of which 1 relates to the analysis of work-related accidents	Partly achieved (50%): 2 out 4 campaigns launched: • Return on prevention study launched in 2018 • Vision Zero (focuses on work-related accidents) launched in 2018 • Prevention Campaign on Chemical and Mechanical Safety with SAA-CIWCE in January 2021 (delayed due to COVID-19) • Campaign on electrical and building safety under preparation (delayed due to COVID-19)
Output B: The organisational capacity of the Labour & Human Resource Department of the province of Punjab to uphold labour and social standards has improved.			
B1. Representatives of all relevant departments of the Labour & Human Resource Department have agreed on five joint recommendations for improved cooperation in the areas of inspection, prevention, rehabilitation, and compensation, as part of an overall concept on the issue of occupational health and safety	0	5 recommendations	Achieved: • Ownership of OHS by employers, employees and government level/political level • Prioritising goals for prevention • Managing interfaces within the different sections • Dealing with limited resources in terms of personnel • Attractiveness of work for own personnel
B2. 3 downstream institutions of the Labour & Human Resource Department in the province of Punjab have improved their performance capacity – as measured by 2 of their KPIs.	0	3 downstream institutions with performance measurement (based on 2 KPIs)	Achieved: Only baseline assessment conducted in 2019, plus the first three workshops to set up a quality management system in 2020. No second performance assessment conducted due to COVID-19. However, performance improvement is plausible (IRI: improved training of labour inspectors; PESSI: improved performance through accident monitoring system; SAA-CIWCE: improved campaigns and training capacities) (see hypothesis 2).
Output C: Private sector actors succeed in increasing their compliance with labour and social standards while simultaneously boosting productivity and improving compliance with environmental standards.			
C1. 8 out of 10 of the participating companies achieve the self-imposed target values at 70% of the 8 indicators for measuring labour and social standards and productivity defined in the context of the factory interventions, whereby 1 objective contributes to gender equality.	0	Target values achieved for 70% of the 8 indicators in 8 companies	Achieved (100%): 8 out of 8 of the participating factories achieved the self-imposed target values at 70% of the 8 indicators (monitoring data).
C2. 5 out of 7 (micro)enterprises or sole proprietors who received advice reach their self-defined target values in 70% of the 6 defined indicators for measurement of labour and social standards and productivity within 5 pilot measures that aim to facilitate the use of the dialogue formats in other parts of the value chain (such as subsectors and informal economies, e.g. home-based work, microenterprises).	0	Target values achieved for 70% of the 6 indicators in 5 out of 7 companies	Partly achieved (80%): 4 out of 5 (micro)enterprises who received advice reached their self-defined target values in 70% of the 6 defined indicators (monitoring data).

Indicators	Baseline	Target value	Indicator achievement
C3. 5 out of 7 (micro)enterprises or sole proprietors who received advice reach their self-defined target values in 50% of the 10 defined indicators for measurement of environmental good practices (wastewater, waste, energy, and chemical management) through the project's dialogue methodology.	0	Target values achieved for 50% of the 10 indicators in 5 out of 7 companies	Achieved (120%): 11 out of 11 enterprises (220%) who received advice reached their self-defined target values in 50% of the 10 defined indicators (monitoring data).

As part of the contribution analyses, the project team verified **hypothesis 1:** Through capacity building of LHRD and downstream institutions, inspectors gain adequate qualification and advisory skills that enable factories in the textile and garment industry to improve OHS measures.

1. Reaction and **2. Learning:** The overall assumption was that labour inspectors were trained only on labour laws and had merely a policing function. Therefore, inspectors were supposed to reflect on their role, as well as building their advisory skills**.** As described above, indicator A1 was achieved (see Table 11), and the curriculum for the training of labour inspectors was revised and was implemented by LHRD. Relevant project stakeholders confirmed that the training content shifted the focus from a policing to an advisory role for inspectors (Int_9). Training of trainers of IRI took place and IRI conducted training activities based on the new curriculum. Overall, 110 labour inspectors (9 newly recruited and 101 existing inspectors) were trained. As per the project's monitoring data, 11 cycles of training were realised on the basis of the revised curriculum. Based on indicator A2 (see Table 11), a self-assessment exercise between untrained labour inspectors' skills before the intervention and a different group of labour inspectors after the intervention showed a difference of 0.67 points on a 5-point scale (4.1 for the post-assessment intervention group, 3.43 for the pre-assessment comparison group (GIZ, no year)). The quality of data is very limited. The assessment is neither a pre/post assessment, as this would need the same group to be assessed before and after, nor a quasi-experimental assessment with a comparison group, as it would require comparable data from both groups before and after the intervention. Therefore, it is important to triangulate the numbers with further qualitative data. Evaluation stakeholders criticised the skills of labour inspectors before participating and, in comparison, praised newly trained inspectors for their improved skills by other public partners (Int_4). Therefore, it can be concluded that an increase in skills is very likely to have occurred.

3. Behaviour change: During group discussions, labour inspectors and District Labour Officers showed their ownership to working as advisors instead of supervisors. They also confirmed that factories lack sufficient knowledge about labour law, describing the strong believe in the private sector that compliance would be connected to increased costs and margins would drop (FGD_9, Int_16). Inspectors and relevant public stakeholders actually showed increased motivation to work on documents such as the training curriculum for labour inspectors (Int_9, Int_16). Nevertheless, factory representatives reported not having experienced any change in the behaviour of labour inspectors (FGD_2). It is not clear whether this is because labour inspectors who have not yet received training on the new curriculum have been working in the factories, because no inspection happened due to COVID-19, or because the change in behaviour has not occurred. Yet individuals were pessimistic about a long-term change in labour inspectors' behaviour as long as this function is not outsourced from the department (FGD_5). Moreover, representatives of several factories mentioned a lack of practical experience to the evaluation team (FGD_5, FGD_2, FGD_4). There are still not enough inspectors for the number of factories, and working as an advisor is perceived to take even more time than simply acting in a policing role. This is especially the case if inspectors have to build trust with factories and make several visits to a factory before being able to work properly (FGD_9, Int_4). There are not enough inspectors to fulfil the role of advisor, given the number of factories (FGD_5). Moreover, some of the current inspectors have not yet been trained, so do not carry out the advisory role (FGD_9).

4. Results: One of the major success factors for labour inspectors was seen as persistence. Inspectors explained that they had to visit factories three or four times before being able to act as advisors (FGD_9). Once trust was built up, inspectors reported having been able to also advise factory workers about their rights in informal meetings (FGD_9). Overall, the changed mindset among trained inspectors and related departments could be observed (Int_4). Furthermore, increased understanding and harmony among inspectors and factories, trust-building measures, and the sharing of knowledge and information were perceived as the main success factors in the long run (Int_16). Thus, inspectors confirmed that the campaigning material developed by SAA-CIWCE supported their advisory role because they were able to build on a strategy (seven golden rules) and have something to which they could refer (FGD_9, Int_4).

Overall, the evaluation team concluded that the hypothesis can be partly confirmed. Although capacity building can be regarded as mostly successful, the feedback from factories does not yet allow behaviour change to be confirmed. A number of hindering factors were identified that reduce the likelihood for long-lasting behaviour change, despite the changing mindset (see section 4.7 – sustainability dimension 3).

Table 12: Selected results hypothesis for effectiveness: hypothesis 1

Hypothesis 1 (activity – output – outcome)	Through capacity building of LHRD and downstream institutions, inspectors gain adequate qualifications and advisory skills that enable factories in the textile and garment industry to improve OHS measures.
Main assumption	• Labour inspectors will act as professional advisors for the industry.
Risks/unintended results	• The inspectorate is prone to corruption. • Factories are not open to cooperation.
Confirmed/partly confirmed/not confirmed	Partly confirmed

As part of the contribution analyses, the project team verified **hypothesis 2:** Through increased targeted collaboration at all institutions of LHRD (B1) and improved capacities for decision-making (management information systems – B2) a process of continuous improvement of performance bottlenecks will increase the organisational capacity of LHRD.

As capacity development in terms of improved individual skills did not play a major role for hypothesis 2, the Kirkpatrick model for training evaluation was not applied. Therefore, reaction, learning and behaviour change were omitted.

4. Results: All institutions related to LHRD confirmed that the breakthrough starting point of targeted collaboration and continuous improvement was the signing of the Dresden Declaration in 2019 (Int_2, Int_9, Int_11). Related stakeholders appreciated that stakeholders from different levels participated equally. The project engaged export-oriented private sector (partner) companies, the Pakistan Employers Federation and the Ministry of Commerce. The process especially strengthened the role of downstream departments and increased their ownership. As LHRD was initially hesitant to sign the declaration and wanted it to go through further administrative processes, other stakeholders stepped forward and signed first (Int_10, Int_2). As such, the Dresden Declaration showed how bottom-up processes can push each other upwards. Because of this, it was asserted that the Dresden Declaration should be considered a remarkable achievement (Int_7, Int_2). Stakeholders also pointed out that the project became a mediator between stakeholders and that the Dresden Declaration could not have been initiated by internal stakeholders. Furthermore, the project guided the development through the establishment of working groups and therefore contributed to a very targeted approach (Int_10, Int_2). The success was said to be based on encouraging stakeholders to change their mindset and on showcasing examples of good practice for improving occupational health in other countries,

such as Germany and Myanmar. Due to the project's efforts to form working groups (e.g. on revising the curriculum to train labour inspectors) and further institutional changes such as the Vision Zero Guidelines (see indicator A2 in Table 11), the communication gap between public institutions was perceived to have reduced, and capacities to have developed and mutual support increased to address problems (Int_4).

Overall, the evaluation team concluded that the hypothesis can be confirmed based on the main assumption of improved cooperation between downstream institutions.

Table 13: Selected results hypothesis for effectiveness: hypothesis 2

Hypothesis 2 (activity – output – outcome)	Through increased targeted collaboration at all institutions of LHRD (B1) and improved capacities for decision-making (management information systems – B2) a process of continuous improvement of performance bottlenecks will increase the organisational capacity of LHRD.
Main assumption	Downstream institutions had to establish cooperation to be able to take informed decisions. Therefore, informed decision-making as well as cooperation between the downstream institutions had to be strengthened.
Risks/unintended results	Downstream institutions do not see the mutual benefit of their work and do not stick to the joint action plan.
Confirmed/partly confirmed/not confirmed	Confirmed

As part of the contribution analyses, the project team verified **hypothesis 3:** Through increased capacities and mobilising internal resources and knowledge, private sector actors increase their compliance with labour (and environmental) standards.

1. Reaction and **2. Learning:** As part of the cooperation with the private sector, companies were supported in successfully addressing labour standards and productivity as interrelated factors for improving their competitiveness. The main method was the DfS, with related change management team workshops that were implemented by the project with local consultants. Furthermore, partner factories reported having learned through visiting other factories and discussions on good practices (FGD_4). Overall, the reaction of factory representatives was that motivation had increased in terms of how productivity and compliance are connected (Int_11). The involvement in change management teams of participants from workers to top management proved to be more of a success factor than focusing solely on the management and owner level (FGD_6, Int_12): 'Our workers have taken a keen interest and are giving supervisors recommendations, e.g. if we do it this way, then more problems are resolved' (FGD_4). Companies reported that the most significant change has been an improved relationship with factory workers due to the use of change management teams (FGD_2).

3. Behaviour change and **4. Results:** The methodological approach of change management teams was described as being a crucial success factor for behaviour change. As such, the participative approach and using SMART principles in identifying KPIs at the factory level divided into productivity and labour standards was perceived as having improved the ownership not only of the factory management but also of other participants in the change management teams, such as factory workers (FGD_4). Some stakeholders even mentioned that the institutionalisation of change management teams in factories was a greater achievement than improved productivity or labour standards because a systemised process also enables future change (FGD_6): 'We are continuously improving, and we review performance on twice-weekly basis now' (FGD_6). Factories appeared to be especially interested in low-cost solutions, such as training staff on OHS. Yet the costs involved in further infrastructural improvements remain a challenge to some companies and more support for improving the supply chain and sales strategies was deemed necessary (Int_11, FGD_3).

Other hindering factors were the lack of commitment of factories' top management and the difficulty of convincing stakeholders who do not see the potential for improvement or do not see the problems in labour standards (FGD_10). CEOs usually assigned their general manager to the change management teams, which led to challenges in the leadership's commitment and ownership (Abacus, 2020: 10). In addition, worker turnover remained an issue, as people trained through change management team workshops left in the middle of training (Abacus, 2020: 10).

Overall, the evaluation team concluded that the hypothesis can be confirmed for the projects' partner factories. However, it should be mentioned that the partner factories were selected by the project and are not a random sample. Therefore, the results are not representative of all small, medium and large factories in Punjab.

Figure 4: Poster for the Dialogue for Sustainability approach.

Table 14: Selected results hypothesis for effectiveness: hypothesis 3

Hypothesis 3 (activity – output – outcome)	Through increased capacities and mobilising internal resources and knowledge, private sector actors increase their compliance with labour (and environmental) standards.
Main assumption	Partner factories see the benefits of improving social and environmental compliance for overall business performance.
Risks/unintended results	Existing conflicts between workers and factories will further increase.
Confirmed/partly confirmed/not confirmed	Confirmed

Overall, the evaluation team concluded that the project's contributions, from capacity building to behaviour change and other results, are outstanding. Although not all the hypotheses can yet be confirmed, ownership, motivation and mutual collaboration have increased. The project's role as a moderator and mediator can be perceived as very successful based on the assessment of effectiveness. The DfS and related change management teams have become a very elaborate methodology for capacity development. Nevertheless, not all output indicators (A2, A3, C2) were fully achieved, and six points were therefore deducted.

Effectiveness dimension 2 – Contribution to achievement of objectives – scores **24 out of 30 points.**

Effectiveness dimension 3: Quality of implementation

Under this dimension, the quality of implementation in the context of GIZ's Capacity WORKS management model for cooperation was analysed. The assessment was based on strategic documents as well as interviews with the project's management, the project team and related key partners at public level.

RBM was well established in the project and was used for evidence-based decisions and risk management. Unintended positive and negative results are part of the project's Excel-based master sheet. Monthly M&E meetings were conducted with the project's technical team to track the activities and their expected output, including to check the status of target achievement (Int_3). As described above, civil society was a lesser part of the project. Based on the Berlin Agreement, FES is working more closely with the trade unions (GTZ, 2004). The evaluation team regarded the collaboration with FES as rather weak, and further potential to work with unions has not been maximised. However, no interview with FES was possible, limiting further the value of the

information (Int_1). With regard to a binding strategy, an implementation agreement was signed in 2017 with LHRD. A steering structure was regularly updated, but due to the high turnover in the management of LHRD – up to three or four secretaries in a year is not uncommon, and the same applies, for example, to the commissioner of PESSI – it was perceived to be difficult to follow this up (FGD_10). With regard to steering, the project prepared a plan of operations, including a Gantt chart and milestones, to show responsibilities to the executing partner and increase transparency. Steering committee meetings happened, though on more of a needs basis, and the project team tried to involve partners with frequent presentations about the current status (FGD_10). Regarding relevant knowledge management and institutionalisation, the high turnover of staff in the partner institutions was a challenge for the project. As a result, LHRD's contribution depended very much on the individuals involved.

Overall, the evaluation team concluded that the project's steering was of the highest quality, taking into account the above-mentioned environmental conditions and changes. Capacity WORKS has been an integral part of the project management. Nevertheless, further collaboration with civil society could have been maximised.

Effectiveness dimension 3 – Quality of implementation – scores **18 out of 20 points.**

Effectiveness dimension 4: Unintended results

To identify unintended and not formally intended positive and negative results, the evaluation team included exploratory questions in focus group discussions and interviews. Using the most significant change technique, several unintended results were identified. Special attention was paid to the motivation and ownership of project partners for achieving results jointly.

None of the interviewed stakeholders reported any major unintended negative results. Some activities were reported to have been less successful, such as the collaboration with smaller factories (FGD_10), a study tour to Indonesia (FGD_4), and costs for employees' social security cards (Int_2). In contrast, the project was praised for its skills in troubleshooting when needed (Int_4), its digital skills for remote collaboration (FGD_6), and its assistance even regarding topics that were not directly linked to its intended results (Int_9).

The most significant change was perceived to be a changing company culture. Companies reported that it was previously very unusual for other companies to visit one another: 'Now we have developed friendship and we are visiting each other and solving problems with each other. Sometimes we don't even involve GIZ, but have an exchange without them' (FGD_5). At the private sector level, due to increased ownership and motivation, informal exchange became much more frequent, and in this way the private sector started to push the public sector for further support and increasingly demanded government accountability. Considering the political feudal traditions in Pakistan and the limited influence on decision-making at public level, this is a remarkable achievement. Therefore, the evaluation team decided to deduct one point.

Effectiveness dimension 4 – Unintended results – scores **19 out of 20 points.**

Photo 2: Meeting of project manager and component manager with Minister of Labour of Punjab (Source: GIZ Pakistan).

Methodology for assessing effectiveness

Table 15: Methodology for assessing OECD/DAC criterion: effectiveness

Effectiveness: assessment dimensions	Basis for assessment	Evaluation design and empirical methods	Data quality and limitations
Achievement of the (intended) objectives	• Indicator Progress Update Sheets and RBM system • Perception of key partners, perception of project team members • SMART* criteria have been met	**Evaluation design:** The analysis follows the analytical questions from the evaluation matrix (see Annex) **Empirical methods:** Interviews, review of monitoring data, analysis of progress and endline reports	• Data quality: moderate • Limitations: low number of survey participants (17 out of 53 factories, i.e. 32%)
Contribution to achievement of objectives	• Results hypothesis 1 • Results hypothesis 2 • Results hypothesis 3	**Evaluation design:** Contribution analysis **Empirical methods:** Interviews, validation workshop	• Data quality: strong • No limitations
Quality of implementation	Capacity works considerations: • RBM system • Plan of operations • Involvement of all relevant stakeholder	**Evaluation design:** The analysis follows the analytical questions from the evaluation matrix (see Annex) **Empirical methods:** Document analysis, interviews	• Data quality: good • No limitations
Unintended results	• Assessment of the mindset and ownership of project partners (e.g. factories)	**Evaluation design:** Most significant change **Empirical methods:** Focus group discussion, interviews	• Data quality: good • No limitations

* SMART: specific, measurable, achievable, relevant and time-bound

4.5 Impact

This section analyses and assesses the impact of the project Improvement of labour and social standards in the Pakistani textile industry, Pakistan (PN 2016.2029.3).

Summarising assessment and rating of impact

Table 16: Rating of OECD/DAC criterion: impact

Criterion	Assessment dimension	Score and rating
Impact	Higher-level (intended) development changes/results	30 out of 30 points
	Contribution to higher-level (intended) development results/changes	30 out of 40 points
	Contribution to higher-level (unintended) development results/changes	30 out of 30 points
Impact score and rating		Score: **90 out of 100 points** Rating: Level 2: successful

Regarding the GSP+ and SDG 8 to promote sustained, inclusive and sustainable economic growth, several achievements in terms of improved working conditions were found. Prerequisites for compliance improved at the public level and at the level of downstream departments. Thus, they contributed to maintaining GSP+ status. A number of organisational capacities at downstream governmental institutions increased and were institutionalised, for example through a digital system for monitoring work-related accidents, a revised curriculum for training labour inspectors, campaigns, information material, guidelines and training material and the quality management system. Nevertheless, contributions from higher-level governmental actors were missing, and this hindered further improvements. Employment and production conditions for employees improved. Workers' participation in particular could be identified as a strong factor for targeted solutions, increased productivity and increased motivation among workers. Regarding unintended results, the project became a role model for other projects of international organisations. Furthermore, the motivation of representatives of alumni factories and their ownership led to the implementation of a CoP that is facilitating learning between factories.

In total, the impact of the project is rated Level 2: successful, with 90 out of 100 points.

Analysis and assessment of impact

Impact dimension 1: Higher-level (intended) development changes/results

During the reconstruction of the results model (seeFigure 2), (intended) overarching development results were identified. At higher outcome/impact level, results comprised contributions to (i) creating and securing job and income opportunities (programme objective), and (ii) improved employment and production conditions that comply with environmental and social standards (subobjective). Both of these goals go hand in hand with improving the conditions for maintaining GSP+ status as well as the promotion of decent work and economic growth (SDG 8). The evaluation team therefore assessed which results could plausibly be considered towards the GSP+ and SDG 8. The initiatives added in 2019 on environmental standards and good practices also contribute to SDG 6 (clean water and sanitation) as well as SDG 12 (responsible consumption and production). However, as these activities are relatively new, no plausible assessment can yet be made of their higher-level results. The evaluation team based its analysis on the findings of a research report on the DfS (Abacus, 2020), the project's aggregated results data, and the key project partners' singular perceptions.

The project proposal made a direct reference to SDG 8 – to 'promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all' – with a specific reference to target 8.8, to 'protect labour rights and promote safe and secure working environments' (GIZ, 2017b: 19). Although the project's primary objective was not trade development, the project was assigned the marker TD 2 (trade development). The project's measures focused on capacity building for implementing labour and social standards were considered to be a main driver for maintaining GSP+ status.

Figure 5: Labour standards achievements (Abacus, 2020: 13)

The research report (Abacus, 2020) on the impact of the DfS in six partner factories highlighted a number of results. However, the extent of success varied from one factory to another, depending on the level of leadership commitment and resources allocated to the programme. As shown in Figure 6 and Figure 7, several achievements in productivity and labour standards were identified. Based on the project's monitoring data, due to increased productivity, the salaries of approximately 7,000 workers increased by 9–10%. In addition, more than 1,316 jobs were created in 10 assessed partner factories. Overall, the project estimated that 20,000 people had benefited from improved working conditions in terms of improved ventilation systems, more environmentally friendly chemicals or improved chemical management, improved light systems, ergonomic chairs, personal protection equipment, safe drinking water, hygienic and subsidised lunches, regular evacuation drills, and the provision of first aid boxes and medical attendants (aggregated results data). In terms of economic outcomes, export-oriented factories experienced 6–10% growth (Abacus, 2020: 14), compared with an average GDP growth of 4.46% between 2014 and 2019 (Economic Survey of Pakistan, 2019).

The above-mentioned improvements were all confirmed by factories participating in a focus group discussion with the evaluation team. However, the numbers of people reached are estimates rather than robust monitoring data. The evaluation team assessed the improvements as plausible, as many of the participants interview confirmed the increased capacities, behavioural change and increased productivity, and highlighted that awareness and knowledge were simply absent before the cooperation with the project started (FGD_2, FGD_4, FGD_5, FGD_6, FGD_7, Int_12).

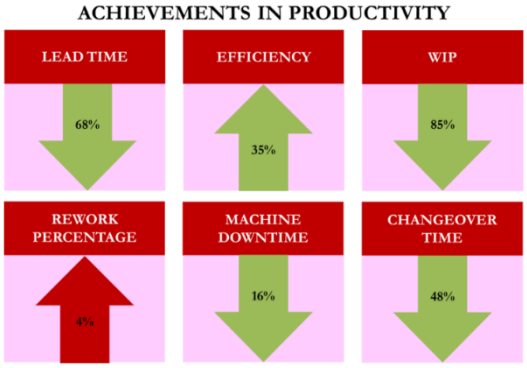


Figure 6: Achievements in productivity (Abacus, 2020: 13)

As described in the section on effectiveness, the implementation of Vision Zero, the quality management system in LHRD's downstream institutions, the accident monitoring system, the revised curriculum for training labour inspectors, and prevention campaigns can be considered a successful improvement of prerequisites for compliance with labour and social standards at public level. An impact at policy level – such as through the National Textiles Policy 2020–2025 or the Punjab Labour Policy, which was revised at the end of 2018, emphasising the strengthening of the labour inspection system – is not plausible, but it was also not a target of the project. Nevertheless, by improving the prerequisites, the project was able to support participating downstream departments in their implementation, and thus contribute to GSP+ status.

Impact dimension 1 – Higher-level (intended) development changes/results – scores **30 out of 30 points.**

Impact dimension 2: Contribution to higher-level (intended) development results/changes

More time is required to fully establish the project's contribution to overarching development results. The evaluation was conducted too soon, limiting the information value of the impact criterion assessment. Further ex-post assessment is needed to assess development results in greater detail. Given that the attribution of impacts, estimated through (quasi)experimental designs, was not possible due to time constraints and the remote set-up of the evaluation, the evaluation team applied contribution analyses. Two hypotheses from the results model were examined in more detail to explain causal relationships between the project's outcomes and impacts.

As part of the contribution analyses, the project team verified **hypothesis 4:** Through increased organisational capacities, LHRD is contributing to maintaining GSP+ status, which is then contributing to inclusive and sustainable economic growth.

Among other topics, the European Commission criticised the lack of a coherent approach and of broad-based approaches to improving labour standards. The labour inspection system in Pakistan in particular is perceived to be weak and to be low on the agenda of priorities (European Commission, 2020: 4).

Improving the coherence between public sector actors (see hypothesis 2 and the Dresden Declaration) would enable a growing sense of responsibility and ownership to be fostered, both at downstream level and among private sector actors (Int_16). The Dresden Declaration did not simply remain a signed paper but was further operationalised by a technical working group to draw an action plan on the agreement and future proceedings (GIZ, 2020h). Therefore, the evaluation team considered the following increased organisational capacities to be the main results of the project (GIZ, 2020h):

- the digital system for monitoring work-related accidents and social security cards by PESSI,
- the revised curriculum for the training of labour inspectors by IRI,
- campaigns, information material, guidelines and training material by SAA-CIWCE, and
- the quality management system based on the common assessment framework for LHRD.

The monitoring data made possible the calculation of the costs of occupational accidents by PESSI as well as a prevention campaign by SAA-CIWCE (GIZ, 2020d: 12). SAA-CIWCE has improved its support to the private sector by providing various booklets and signboards, for example on health and safety issues in local languages, on training in risk assessment and on mechanical topics (FGD_6).

Nevertheless, stakeholders also pointed out other hindering factors. Although private sector representatives acknowledged an initial change at public level, they reported that this has not yet sufficiently reached their working environment (FGD_2). Moreover, the lack of contribution from higher-level governmental actors is hindering further improvements at factory level (FGD_5).

Overall, the evaluation team concluded that the hypothesis can be partly confirmed. Given that the downstream institutions are pushing LHRD and collaborating with each other, ownership and perceived self-efficacy can be observed. Organisational capacities have clearly increased, but to what extent they will contribute to maintaining GSP+ status remains uncertain. If the risk of suspension of labour inspectors materialises, or if compliance with the ILO core labour standards is not follow up, GSP+ status might be lost, regardless of the increased organisational capacities. Moreover, it remains plausible to the evaluation team that the above-mentioned improvements at private level can be explained less by the improved prerequisites at public level than by the DfS.

Table 17: Selected results hypothesis for impact: hypothesis 4

Hypothesis 4 (outcome – impact)	Through increased organisational capacities, LHRD is contributing to maintaining GSP+ status, which is then contributing to inclusive and sustainable economic growth.
Main assumption	Good governance of state institutions in labour standards is the most crucial prerequisite for compliance with labour standards.
Risks	The labour inspection system is suspended.
Confirmed/partly confirmed/not confirmed	Partly confirmed

As part of the contribution analyses, the project team further verified **hypothesis 5:** Through increasing compliance with labour and environmental standards, private sector actors improve the employment and production conditions in selected economic sectors.

As described above, many improvements in productivity and efficiency were noted. In addition, many aspects of labour and social standards were observed to have improved. For this reason, the following contribution analysis focuses more on the impact on factory workers.

Companies participating in the DfS have increasingly involved their employees in the change management process. Workers frequently highlighted that attending meetings and talking to the management and owners had not happened previously. Several quotes are a strong indication of the impact: ‘It improved our self-confidence to talk to our manager and owner – that is best thing I learned’; ‘Our shyness and fear is gone’; ‘(…) that was the biggest motivation, sitting with owners and management – that helped a lot to improve our working’ (FGD_1).

Participants from alumni factories also pointed out that the social dialogue gap needed to be reduced as unions are not strong in Pakistan and thus the workforce lacks a strong voice (FGD_5). The project gave workers the opportunity to articulate their challenges on the one hand and to provide valuable input for increasing efficiency on the other. As workers pointed out: ‘Due to GIZ we sit with our management and we tell them our problems and then these are solved, like light and water problems, ventilation problems’ (FGD_1); ‘First we were working on contractual basis (per piece) and now we are working on salaries, and owing to improvements in line/layout we are making 60 pieces per worker as compared to 45 pieces previously’ (FGD_1). Nevertheless, the workers wish to have further training. Reported impact is perceived to have occurred through the consultation by the project and the consultants. In addition, information material was perceived to be very helpful (Int_12). Yet further training by the management and owners of factories is limited (FGD_1). Moreover, the DfS was not able to resolve issues of workers’ unions and other human rights issues. Nevertheless, worker councils were created in some factories (Abacus, 2020: 9).

Overall, the evaluation team concluded that relevant private sector actors had improved the employment situation. Based on workers’ perceptions, the hypothesis can be confirmed. However, workers’ rights depend very much on the conviction of management as to whether improvements will lead to greater productivity.

Table 18: Selected results hypothesis for impact: hypothesis 5

Hypothesis 5 (outcome – impact)	Through increasing compliance with labour and environmental standards, private sector actors improve the employment and production conditions in selected economic sectors.
Main assumption	Private sector actors see the economic benefit of increasing compliance with labour and environmental standards.
Risks	Private sector actors refuse to stop rent seeking.
Confirmed/partly confirmed/not confirmed	Confirmed

Overall, the evaluation team concluded that the project contributed to a large extent to increased organisational capacities and institutionalised prerequisites for compliance. Furthermore, an improved employment situation and production system in partner factories is evident. Nevertheless, improvements remain fragile as they depend on the factories' conviction. In areas that are less lucrative or that generate higher costs for the factories, the situation might not improve. Furthermore, the absence of contributions from higher-level governmental actors is hindering further improvements.

Impact dimension 2 – Contribution to higher-level (intended) development results/changes – scores **30 out of 40 points**.

Photo 3: Senior management team member shares his view during a change management team workshop (Source: GIZ Pakistan).

Impact dimension 3: Contribution to higher-level (unintended) development results/changes

To identify unintended and not formally intended positive and negative results at impact level, the evaluation team included exploratory questions in interviews with key partners. As with the analysis of the effectiveness criterion, the evaluation team made use of the most significant change technique.

One of the most significant unintended results were contributions to SDG 17 (partnership). As already mentioned within the unintended results at the effectiveness level, other unintended results regarding informal exchange can be observed at impact level. The project and related factories were to some extent considered role models. As such, external project partners (Int_14, Int_6) started building their activities on what was learned from the project's cooperation with the private sector. Furthermore, factories became role models that inspired other factories. Informal exchange increased significantly, leading to factories sharing information.

Approaches were reported to be replicated by other factories and the overall culture became more open (FGD_5, FGD_6). Hence, the CoP (see 4.7) was born out of a demand from alumni factories who wanted to continue the capacity building and share lessons learned (FGD_5). Although a turnover of staff at factory level was perceived as a challenge, factories also reported that this led to cross-fertilisation between factories. New staff at management level would report on activities and improvements as well as the CoP, so the informal network would grow further on its own. Moreover, factories acknowledged the processes and improvements happening at public level. PESSI held its first interprovincial coordination committee meeting in Khyber Pakhtunkhwa, at which best practices were shared (Int_2). Moreover, the professional reputation of LHRD and the downstream institutions are considered to have improved to some extent within the private sector (FGD_2). Further contributions to negative unintended development results were not reported to the evaluation team. Therefore, no harm was identified as a result of the project's activities.

Impact dimension 3 – Contribution to higher-level (unintended) development results/changes – scores **30 out of 30 points**.

Methodology for assessing impact

Table 19: Methodology for assessing OECD/DAC criterion: impact

Impact: assessment dimensions	Basis for assessment	Evaluation design and empirical methods	Data quality and limitations
Higher-level (intended) development changes/results	Programme objective indicators Further impact on: • SDG 8 • GSP+	**Evaluation design:** The analysis follows the analytical questions from the evaluation matrix (see Annex) **Empirical methods:** Interviews with public sector representatives (LHRD, ILO) and the project team, document analysis	• Data quality: good • Limitations: programme is limited to Punjab • Anecdotal evidence, as robust statistical data was missing
Contribution to higher-level (intended) development results/changes	• Hypothesis 4 (GSP+) • Hypothesis 5	**Evaluation design:** Contribution analysis **Empirical methods:** Interviews, validation workshop	• Data quality: good • Limitations: too early for development results to be established; plausible assessment rather than robust data analysis
Contribution to higher-level (unintended) development results/changes	• Do-no-harm analysis	**Evaluation design:** Most significant change **Empirical methods:** Focus group discussion, interviews	• Data quality: good • Limitations: anecdotal evidence; no data on negative unintended results

4.6 Efficiency

This section analyses and assesses the efficiency of the project Improvement of labour and social standards in the Pakistani textile industry, Pakistan (PN 2016.2029.3).

Summarising assessment and rating of efficiency

Table 20: Rating of OECD/DAC criterion: efficiency

Criterion	Assessment dimension	Score and rating
Efficiency	Production efficiency (resources/outputs)	65 out of 70 points
	Allocation efficiency (resources/outcome)	27 out of 30 points
Efficiency score and rating		Score: **92 out of 100 points** Rating: **Level 1: highly successful**

There are no robust indications that outputs A, B or C could have been maximised with the same volume of resources by considering a different setting or structure. Nevertheless, as some of the budget remained (residuals), the number of factories involved in output C could have been increased. Overarching costs for risk management, travelling and administration remained low (2%). Planned expenditure was regularly reviewed and adjusted accordingly. The project's use of resources was mostly appropriate with regard to achieving its objectives. Outcomes could have been maximised if scaling-up had been considered earlier in the project design. Human resource management was highlighted for its systematic approach in response to demands (COVID-19). Overall, the project had to deal with three change offers demanded by BMZ. The use of resources remained appropriate and led to satisfactory results.

In total, the efficiency of the project is rated Level 1: highly successful, with 92 out of 100 points.

Analysis and assessment of efficiency

The key question addressed by the efficiency criterion is whether the project's use of resources is appropriate with regard to achieving both the outputs and the outcome (project objective). The evaluation examined whether the level of resourcing (e.g. funding, expertise) led to satisfactory results. Combining information on both the project's costs and its results provides more insights than looking at these two components separately. Focusing on results alone would limit the use of data in strategic decision-making. Focusing on costs alone may distract from the recommendations that aim to ensure quality in the results.

A distinction is made between two types of efficiency: production and allocation. While the former evaluates the transformation of inputs to outputs, the latter evaluates the transformation of inputs to results at outcome level. This includes the analysis of the extent to which even more could have been achieved at output level with the same overall use of funds. It is therefore a question of investigating not only how costs could have been saved but also how existing resources could have been better used to achieve the desired results.

Following GIZ's guidelines on assessing efficiency, this central project evaluation applied the follow-the-money approach as a standard method for analysing the project's production efficiency.

The evaluation team used an Excel tool developed by GIZ's Corporate Unit Evaluation to standardise the efficiency analysis of the project. The tool (henceforth called the 'efficiency tool') takes into account GIZ's recommendations on analysing a project's efficiency. It refers to sources that are available in the project.

Efficiency dimension 1: Production efficiency

The following assessments were based on information extracted from the cost commitment report and further discussions with the project team and project management using the follow-the-money approach (Palenberg, 2011: 46). The costs and commitments of the project are presented in Table 21.

Table 21: Overview of costs

Module objective	**The prerequisites for compliance with labour and social standards in the province of Punjab have improved.**
BMZ costs	€10,073,780.60
Cofinancing	€0.00
Partner contribution	€0.00
Total costs	**€10,073,780.60**
Residual	€795,011.64

Data received in the cost commitment report were dated August 2020, so residuals remained. Based on the feedback received by the project's management and the information provided to the evaluation team, the project ended up with residuals of EUR 795,011.64. Of this sum, EUR 650,000 was transferred to the follow-on project, leaving the project with EUR 145,000.64 that was spent on output A (35%), output B (25%) and output C (40%). The project was commissioned before the Gemeinsamen Verfahrensreform (GVR) and is a follow-on project. Thus, the budget–actual comparison, which has been required for new projects since 2017 – was not assessed. Therefore, deviations cannot be analysed.

ZAS costs, which were external professional and administrative services that were booked for the project, represented a relatively high share of almost 8%. One of the main reasons was the need for risk management: EUR 140,000 was spent on the risk management office (Int_1).

Maximum principle and reallocation of funds: Although not all indicators were fulfilled (see Table 22) the evaluation team concluded that there was still a high likelihood that the outputs have been maximised. Some of the output indicators (A2, C1, C2, C3) measure not only the completion of activities but also the success of the capacity building. Regarding output C, it has been argued that the overall number of partner factories could have been increased, as for C1 and C2 the numbers of participating factories (10 for C1 and 7 for C2) were not reached. However, the project had to develop the DfS methodology further based on the partners' needs and capacities, so the number of manageable partner factories was considered to have been maximised (FGD_10). The achievement of indicator A3 was delayed due to the effects of COVID-19 and will therefore be transferred to the follow-on project. All indicators were considered very ambitious by the evaluation team, though not overambitious. According to the project's management, project planning was very dynamic due to the number of change offers (Int_1).

Table 22: Overview of output achievement

Output Indicators A	A1) The Industrial Relations Institute (IRI) of the LHRD of the province of Punjab has revised the curriculum for training labour inspectors.	A2) In a competency survey, 16 out of 20 labour inspectors of four selected labour authorities at district level achieve results 1 level better on a 5-level scale than 10 labour inspectors from a comparison group.	A3) The LHRD's SAA-CIWCE has conducted 4 campaigns on risks in the area of occupational health and safety in the textile and garment sector, of which 1 directly relates to the results of the analysis of work-related accidents by LHRD's PESSI.
Achievement	100%	50%	50%

Output Indicators B	B1) Representatives of all relevant departments of LHRD have agreed on 5 joint recommendations for improved cooperation in the areas of inspection, prevention, rehabilitation and compensation, as part of an overall concept on the issue of occupational health and safety	B2) 3 downstream institutions of LHRD in the province of Punjab have improved their performance capacity, as measured by 2 of their KPIs.

Achievement	100%	100%

Output Indicators C	C1) 8 out of 10 of the participating companies achieve the self-imposed target values at 70% of the 8 indicators for measuring labour and social standards and productivity defined in the context of the factory interventions, whereby one objective contributes to gender equality.	C2) 5 out of 7 (micro)enterprises or sole proprietors who received advice reach their self-defined target values in 70% of the 6 defined indicators for measurement of labour and social standards and productivity within 5 pilot measures that aim to facilitate the use of the dialogue formats in other parts of the value chain.	C3) 5 out of 7 (micro)enterprises or sole proprietors who received advice reach their self-defined target values in 50% of the 10 defined indicators for measures of environmental good practices (wastewater, waste, energy and chemical management) through the project's dialogue methodology.
Achievement	100%	80%	220%

As shown in Table 23, costs are unevenly distributed between outputs A, B and C. Output C ranks as the most expensive output with 33%, followed by output A with 29% and output B with 21%. In general, the relatively higher costs for output C can be explained by the fact that this is the focus area of the project and the only output relating to the private sector development. Given the private sector stakeholders and consultants involved, the evaluation team considered the costs to be relatively low. Costs relating to outputs A and B overlap as a clear division between the activities was not always possible. Although output B accounts for the lowest share, the evaluation team perceived it to be relatively more expensive, as international consultants (Como Consulting) were involved. Nevertheless, output A and B have a combined share of 50% of the costs, reflecting the number of stakeholders from downstream departments involved.

Overarching costs appear to be very high, at 17% of total costs. However, of this amount, more than 15% of total costs can be explained by the financial agreement with ILO. If these costs are deducted, only 2% of total costs remains, and this was spent on risk management, travelling and administration (Int_1). A regular review of the resources used by the project, focusing on the economical use of resources, was regularly updated, as the change offers made it necessary to adjust project planning (Int_1).

Table 23: Overview of costs allocated to outputs

	Output A	**Output B**	**Output C**	**Overarching costs**
Outputs	The capacities of Labour & Human Resource Department and selected downstream authorities of the province of Punjab for carrying out labour inspections are improved.	The organisational capacity of the LHRD of the province of Punjab to uphold labour and social standards has improved.	Private sector actors succeed in increasing their compliance with labour and social standards while simultaneously boosting productivity and improving compliance with environmental good practices.	
Cost including committed costs	€2,886,519.77	€2,106,116.70	€3,319,351.60	€1,761,792.54
Cofinancing	€0.00	€0.00	€0.00	€0.00
Partner contributions	€0.00	€0.00	€0.00	€0.00
Total costs	**€2,886,519.77**	**€2,106,116.70**	**€3,319,351.60**	**€1,761,792.54**
Total costs in %	**29%**	**21%**	**33%**	**17%**
BMZ total in % without cofinancing	**29%**	**21%**	**33%**	**17%**

As described in Table 24, and as mentioned above, output C worked mostly with national staff and national consultants as short-term experts. Project staff in Germany represented just 2% of the staff costs, while national staff accounted for 36% and international staff 62%. Seconded staff were mainly one cluster coordinator and the portfolio manager responsible for steering and strategic coordination of national partners and BMZ. As such, they had a greater share in output A and B. National staff overall consisted of 16 technical advisors and 9 support staff involved in supporting and implementing the activities.

Other GIZ projects pointed out that the project was managed efficiently and did not see potential for an alternative use of inputs (Int_5, Int_13, Int_15, Int_17). External stakeholders highlighted the project's use of inputs, especially its focus on the private sector and downstream departments as the most appropriate design (Int_14, Int_6).

Table 24: Distribution of personnel on outputs

	Output A	Output B	Output C	Overarching costs
International staff (AMA/PMA)	40%	38%	18%	3%
National staff	27%	27%	45%	1%
Project staff in Germany (PMI)	33%	33%	33%	0%

Monitoring system and handling of risks: As mentioned above, a monitoring system at project level was in place and was well maintained. Risk mitigation was found to be part of the monitoring at project level and was frequently discussed.

Consideration of lessons learned: As mentioned above, the project is built on the experiences and lessons learned of the two projects Implementation of Social Standards Support Programme to the Textile and Garment Industry of Punjab (PN 2013.9062.4) and Water Efficiency in the Textile Industry (PN 2013.9773.6). Thus, project staff who worked in the predecessor projects were already known in the sector and had a large network of national stakeholders and staff of international projects. Learning experiences were considered very important by the project. As such, approaches and methods were constantly reviewed and updated (e.g. DfS).

Outsourcing of activity packages: The project assigned a number of local and international consultants to activity packages such as the implementation of the DfS at partner factory level and quality management in LHRD. The evaluation team found no direct indications that activity packages could have been further outsourced to increase efficiency. Project partners highlighted that working with local consultants at factory level in particular was very cost efficient and effective (Int_8, Int_11).

Overall, according to the evaluators' analysis, there are no robust indications that outputs A, B or C could have been maximised with the same volume of resources by considering a different setting or structure. The project had to deal with many administrative tasks due to the change offers, which could have hindered the implementation further, but as the indicator achievement shows, the project managed to stay focused. Nevertheless, as residuals remained it could be argued that the number of factories involved in output C could have been maximised. Moreover, outputs were not all produced in time due to the impact of the COVID-19 pandemic.

Efficiency dimension 1 – Production efficiency – scores **65 out of 70 points**.

Efficiency dimension 2: Allocation efficiency

The evaluation team assessed the extent to which the project's use of resources was appropriate in relation to its objective based on the GIZ efficiency tool analysis. Other qualitative findings considered are plausible assumptions and anecdotal evidence. Nevertheless, this evidence provides indications on how the outcomes

could have been maximised. In contrast to production efficiency, allocation efficiency describes the transformation of inputs to outcomes. At module objective level, indicators MOI 1, MOI 2, MOI 3 and MOI 4 were fully achieved. Table 25 summarises the results already described in more detail in the effectiveness section (4.4).

Table 25: Overview of outcome achievement

Module outcome indicators	MOI 1) A plan of action for improving the state labour inspectorate, agreed with state, private sector and civil society actors, is in place.	MOI 2) 75% of newly recruited labour inspectors from July 2018 have, within the first 6 months of their work, undergone the introductory curriculum developed within the framework of the project.	MOI 3) 60% of 50 surveyed textile or garment companies confirm that the products of the SAA-CIWCE (an agency of the LHRD of the province of Punjab, SAA-CIWCE) for occupational health and safety (e.g. campaigns, informational material, guidelines, training courses) are used.	MOI 4) 5 textile and garment producers have concluded contracts with private sector and/or public service providers (producer associations, chambers of commerce, consulting firms, universities, etc.) for advisory services based on the project's dialogue methodology, to ensure compliance with social standards while also increasing productivity.
Achievement	100%	109%	113%	100%

Synergies: As already mentioned in the assessment of internal coherence, the lack of strategic coherence in the design of the development cooperation programme has resulted in the absence of further synergies in the field of sustainable economic development in Pakistan (Int_1). Nevertheless, based on the responsibilities of the regional project FABRIC, knowledge management within the textile sector took place through conferences in Pakistan, Cambodia and Myanmar (see 4.3). Furthermore, cooperation with GIZ FABRIC and GIZ SP-SHP supported the achievement of results (environmental standards and monitoring occupational accidents).

Cofinancing: The project had no cofinancing, and partner contributions (office space) were not realised due to security factors and limited partner resources. The evaluation team did not find further evidence on efforts towards cofinancing opportunities through the project.

Human resource management: In terms of project management, clear roles and responsibilities concerning the cluster managers were in place and managed well through the management team. Overarching costs in particular could therefore be limited. The high level of involvement of national project staff was praised as a very efficient style of project management. It should also be mentioned that capacity building among the project team was facilitated (e.g. M&E, individual coaching for team leaders, digital moderation skills). As such, human resource management was highlighted for its systematic approach (Int_8).

Consideration and realisation of possibilities for scaling-up: With regard to the DfS, scaling up to other regions in Pakistan and expanding the scope within Punjab was not an initial strategy of the project but slowly developed as considerations of further multipliers were made. The project has already put some effort into identifying multipliers and starting collaboration (see MOI 4) during its term (see section 4.5). However, the initiative will be continued in the follow-on project (Int_1).

Overall, the evaluation team concluded that the project's use of resources was mostly appropriate with regard to achieving its objectives. Outcomes could have been maximised if scaling-up had been considered earlier in the project design.

Efficiency dimension 2 – Allocation efficiency – **scores 27 out of 30 points**.

Methodology for assessing efficiency

Table 26: Methodology for assessing OECD/DAC criterion: efficiency

Efficiency: assessment dimensions	Basis for assessment	Evaluation design and empirical methods	Data quality and limitations
Production efficiency (Input/outputs)	Transformation of inputs to outputs based on: • GIZ efficiency tool • cost commitment report of the project • comparison of planned budget figures with actual figures • results matrix • RBM system	**Evaluation design:** • The analysis follows the analytical questions from the evaluation matrix (see Annex) • Follow-the-money approach **Empirical methods:** Interview with project management and project team, document analysis	• Data quality: good • No limitations
Allocation efficiency (Input/outcome)	Transformation of inputs to outcome based on: • GIZ efficiency tool • cost commitment report of the project • comparison of planned budget figures with actual figures • results matrix • RBM system	**Evaluation design:** • The analysis follows the analytical questions from the evaluation matrix (see Annex) • Follow-the-money approach **Empirical methods:** Interviews, document analysis	• Data quality: good • Limitations: qualitative findings considered are plausible assumptions and anecdotal evidence

4.7 Sustainability

This section analyses and assesses the sustainability of the project Improvement of labour and social standards in the Pakistani textile industry, Pakistan (PN 2016.2029.3).

Summarising assessment and rating of sustainability

Table 27: Rating of OECD/DAC criterion: sustainability

Criterion	Assessment dimension	Score and rating
Sustainability	Capacities of the beneficiaries and stakeholders	15 out of 20 points
	Contribution to supporting sustainable capacities	22 out of 30 points
	Durability of results over time	35 out of 50 points
Sustainability score and rating		Score: **72 out of 100 points** Rating: Level 3: moderately successful

Stakeholders overall confirmed that capacities are long lasting because they have been institutionalised into the factories. Sustainable effectiveness has been achieved through the methodological approach. Nevertheless, refresher workshops to systematically follow up on factories were deemed necessary to maintain quality and to be able to increase learning further. Institutional multipliers (four local consulting firms, one governmental service provider and one NGO) were identified to further implement the DfS without the project. Nevertheless, they will need further support as their business model and sales strategy is not yet working

sustainably. In addition, a CoP was formed to facilitate informal learning and has the potential for further institutionalisation. Further increased organisational capacities of the public sector have already been introduced and systematically integrated. External factors remain a big challenge for the durability of results concerning the labour inspection system. Furthermore, wider effects in the textile and garment sector will rely on additional investments in TVET. Overall, poorly coordinated policies and governmental structures will continue to hinder further effects.

In total, the sustainability of the project is rated Level 3: moderately successful, with 72 out of 100 points.

Analysis and assessment of sustainability

Sustainability dimension 1: Capacities of the beneficiaries and stakeholders

The first dimension assessed the extent to which the beneficiaries and stakeholders of the project have the institutional and human resources as well as the willingness to sustain the results of the project over time. Factories that participated in the earlier stages of the project reported that capacities were still there: 'We focused on the system so that changes do not stop after project phase out (…) many things, which we learned are happening as these are now part and parcel of the system' (FGD_4). Nevertheless, factories also demanded further refresher courses for them and factory workers. A very active CoP, where knowledge and best practices are shared made it possible to sustain the partnership with factories where cooperation has officially ended. Nevertheless, this remains rather informal, and further refresher workshops to systematically follow up on factories does not so far exist (FGD_1). Although it could be a valuable entry point for multipliers of the project to provide refresher workshops, it remains unclear how far factories would be willing to pay for them (FGD_3). Due to their increased capacities to identify where support is needed, factories are more likely to invest in external consultants. However, it will depend on the pricing models and the quality of services (Int_1).

With regard to the public and private sectors, the project focused on concrete products such as the DfS. Ensuring the sustainable effectiveness of the project is therefore achieved not through additional capacity development measures but through the methodological approach of the DfS on the one hand and through concrete improvements, such as the monitoring system of PESSI and the organisation of campaigns, on the other. In the follow-on project, multipliers will be trained in more depth and deployed in the field. In the future, the project should thus become unnecessary for implementing DfS. Moreover, through the regional project FABRIC, the DfS an be scaled to other countries such as Bangladesh (where it was originally developed).

Overall, the evaluation team concluded that five points should be deducted because follow-on activities are still not present.

Sustainability dimension 1 – Capacities of the beneficiaries and stakeholders – scores **15 out of 20 points.**

Sustainability dimension 2: Contribution to supporting sustainable capacities

As the analysis of sustainability goes very much hand in hand with the assessment of impact and effectiveness of the project, the evaluation team used a similar methodological basis. The evaluation team collected evidence on the project's contribution to institutionalisation and systematisation.

Regarding the institutionalisation of the DfS, the project conducted a comprehensive assessment of 33 firms and organisations to find qualified institutional multipliers and signed Memoranda of Understanding in 2019 and 2020 with six multipliers. The institutional multipliers are four local consulting firms (Institute of Quality, NEC Consultants Pvt. Limited, Aftec Pvt. Limited, Environmental Services Pakistan), one governmental service

provider (Pakistan Institute of Management) and one NGO (Pakistan Society for Training and Development). It should be kept in mind that the project did not have a target for multipliers to work independently until the end of 2020 (Int_1). Institutional multipliers will therefore be further integrated into the follow-on project.

Stakeholders identified several hindering factors and challenges that will need to be tackled in the follow-on project to increase the sustainability of DfS and to improve the business model of the institutional multipliers:

- Overall, the institutional multipliers confirmed that capacity building had taken place, but pointed out that they are not yet ready to disseminate the DfS independently. Factories confirmed that the quality of the multipliers is perceived to be lower than that of the project's support (FGD_3). Moreover, as the multipliers were new and the COVID-19 pandemic restricted implementation, they had not yet been able to secure any contracts in 2020 (Int_3). Contracts relating to MOI 4 were concluded not with the institutional multipliers but with the service providers from the project.
- As the project implemented the approach, it was free of charge for companies. Hence, the buy-in from companies was considered easier, enabling the methodology to be further piloted and developed (FGD_10). Nevertheless, the lack of financial contribution is considered a challenge for the multipliers, as existing partners will not be willing to pay. Therefore, multipliers recommended the inclusion of an increased financial contribution from the factories to facilitate the transition (FGD_7). Moreover, in future, the costs for consultancy would have to be included in the calculation of how the DfS leads to increased productivity and earnings. Sales strategies could not be further developed at the end of the project period and must be taken up again in the follow-on project. They will be needed as part of sustainable business models.
- Multipliers still lack the reputation possessed by the project from previous assignments (FGD_3).
- Multipliers need further training and especially practical experience. Given that the project team themselves had a learning curve, the same will also apply to the multipliers (Abacus, 2020: 10).

Another approach to institutionalisation was to build a CoP in 2019. The CoP mainly consists of an alumni network of (former) partner factories. As the project recognised that informal networks play a major role (see section 4.5), the CoP should facilitate and institutionalise mutual learning between factories (GIZ, 2020d: 10). Project partners confirmed that through the CoP, sustainable capacities are being strengthened as mutual learning takes place (Int_17). Part of the CoP is a learning platform/portal of all alumni factories of all textile projects that GIZ has implemented since 2014 (GIZ, 2020d: 13). However, at the end of the project, the CoP portal was perceived by the CoP members not to be working yet (FGD_5). The motivation of CoP members remains high and further collaboration is required from them to play a more crucial role: 'We need clear objectives for CoP for three years, broken down on yearly basis, and then we can make action plans to implement accordingly' (FGD_5). Further institutionalisation of informal networks can be considered to improve sustainability. Furthermore, the project is already considering combining the multipliers and CoP as a related consortium in the follow-on project.

During the inception mission an additional question was raised regarding what would have happened if crucial project staff (e.g. the project manager) had left the project. Based on the perception of key project partners and related consultants, the project's management was praised multiple times for its transparency of informed decisions, its ability to share responsibilities and its systematic capacity building within the project team (Int_8, Int_13, Int_14, Int_17). It was also praised for its technical expertise, from which other projects had been able to learn. Thus, if a crucial member of the project team had left, the level of expertise would be diminished, but the project's management had already invested in more capacity building than other projects do. Therefore, the evaluation team concluded that success was not bound to an individual but that the project managed responsibilities as a team effort.

Overall, the evaluation team concluded that the project was able to identify promising strategies for the follow-on project to support sustainable capacities. However, within the current project phase, stakeholders are not

yet ready to take over. Nevertheless, with the CoP, a movement has been created that continues to share learning experiences independently of the project. Moreover, organisational capacities of the public sector were increased and have been introduced and systematically integrated (e.g. monitoring system, curriculum, Vision Zero).

Sustainability dimension 2 – Contribution to supporting sustainable capacities – scores **22 out of 30 points**.

Sustainability dimension 3: Durability of results over time

The core question regarding the durability of the project's positive results was based on the analysis of contextual challenges that were raised by key interview partners.

With regard to the labour inspection system, key stakeholders pointed out that despite the success of the project, labour law and the labour inspection system remain very low on the priority of LHRD. As labour inspectors' pay remains low and they lack support in terms of equipment, office space and means of transportation, long-term behavioural change is hindered to a great extent. Moreover, resources are only part of the solution, the main issues being organisational set-up, management practices, and lack of accountability and liability. The project's efforts with the accident monitoring system have increased the visibility of poor labour inspection. Corruption is perceived by international stakeholders to be endemic among labour inspectors. A systematic change in the behaviour of labour inspectors might only happen if the environmental conditions improve and if the institutions involved in prevention and in compensation and rehabilitation further increase their collaboration. Hence, LHRD was perceived as not providing inspectors with the means to offer professional advisory services. Although ownership and motivation improved, external stakeholders remain pessimistic, believing that the system in which labour inspectors are paid for not entering factories rather than for acting as advisors will continue (Int_14). Inspectors themselves reported receiving phone calls and threats from other inspectors warning them not to do their jobs according to the new advisory approach (FGD_9). Moreover, the risk of a suspension of labour inspections remains. This means that the investments made in the project and the successes achieved so far are at stake. Nevertheless, the evaluation team concluded that the project had started a very holistic approach to changing the labour inspection system. The question remains to what extent LHRD will further realise the importance of the labour inspection system for the GSP+.

With regard to the textile and garment factories, the overall capacities of the management and factory workers remains limited. Although the project realised some noteworthy achievements, wider effects on the textile industry will need investments in TVET (FGD_2). With regard to the international textile market, Pakistan might not be able to compete with countries such as Bangladesh and China. Therefore, GSP+ status might not become a strong lever for systematic change at governmental level towards compliance with labour, social and environmental standards. Nevertheless, Pakistan encompasses all products of the entire value chain of the textile industry and therefore has potential for further growth.

Overall, political decisions are determined by a multitude of forces that on the one hand act for the government and on the other use their influence on the government to achieve their own goals. This leads to poorly coordinated policies and actions of the various departments within the provinces or between provinces, while at the federal level, policy documents can be almost worthless as soon as they are published. Nevertheless, the project remained realistic in supporting the private sector and downstream institutions. After all, it might take decades, and not just a project period, to achieve realistic impact at the governmental level (FGD_10).

Sustainability dimension 3 – Durability of results over time – scores **35 out of 50 points**.

Methodology for assessing sustainability

Table 28: Methodology for assessing OECD/DAC criterion: sustainability

Sustainability: assessment dimensions	Basis for assessment	Evaluation design and empirical methods	Data quality and limitations
Capacities of the beneficiaries and stakeholders	As capacity building is a crucial part of the indicators, the dimension will be assessed against the project's indicators (A2, B2, C1, C2, C3) and their maintained application (MOI 3, MOI 4).	**Evaluation design:** The analysis follows the analytical questions from the evaluation matrix (see Annex) **Empirical methods:** Interviews and focus group discussions with factory management, participants from LHRD and labour inspectors	• Data quality: strong • No limitations
Contribution to supporting sustainable capacities	See above	**Evaluation design:** The analysis follows the analytical questions from the evaluation matrix (see Annex) **Empirical methods:** Interviews, focus groups	• Data quality: good • Limitations: anecdotal evidence
Durability of results over time	See above	**Evaluation design:** Interviews, validation workshop with the project team **Empirical methods:** Interviews, online survey	• Data quality: good • Limitations: anecdotal evidence

4.8 Key results and overall rating

The success of the project lies in its participatory, bottom-up approach with motivated private sector actors as well as with public sector downstream institutions, with ownership for shared responsibility. The success factor is the very close cooperation of different actors at the same level (e.g. factory workers and factory management). Project partners from different institutions were successfully supported to increase their own efficiency on the one hand and to create synergies through cooperation on the other. This was made possible through the very targeted needs assessment facilitated by the project and the establishment of ownership through jointly developed action plans and self-defined KPIs. Particularly noteworthy is the methodology of the DfS, with associated change management teams for implementation. Although success at the private sector level in increasing productivity and improving labour, social and environmental standards is happening on a small scale, further impacts resulting from public actors' improved capacities are to be expected in the next project phase.

Overall, the framework conditions and the commitment of higher government agencies are the greatest obstacles to broader achievements. The financial and logistical resources as well as the basic professional training of factory workers up to management level, factory owners, labour inspectors and other public sector employees all require further systematic support. Moreover, the sustainability of results is not yet ensured. Although approaches to improve sustainability can be observed in multipliers and the CoP in the private sector, along with improvements in digitalisation and process optimisation in the public sector, their long-term market potential has yet to be demonstrated.

The project is rated successful, with an overall mean score of 86 out of 100 points (see Table 30).

Photo 4: Group work during change management team workshop (Source: GIZ Pakistan.).

Table 29: Rating and score scales

100-point scale (score)	**6-level scale (rating)**
92–100	Level 1: highly successful
81–91	Level 2: successful
67–80	Level 3: moderately successful
50–66	Level 4: moderately unsuccessful
30–49	Level 5: unsuccessful
0–29	Level 6: highly unsuccessful
Overall rating: The criteria of effectiveness, impact and sustainability are knock-out criteria: If one of the criteria is rated at level 4 or lower, the overall rating cannot go beyond level 4 although the mean score may be higher.	

Table 30: Overall rating of OECD/DAC criteria and assessment dimensions

Evaluation criteria	Dimension	Max	Score	Total (max.100)	Rating
Relevance	Alignment with policies and priorities	30	30	**94**	Level 1: highly successful
	Alignment with the needs and capacities of the beneficiaries and stakeholders	30	27		
	Appropriateness of the design	20	17		
	Adaptability – response to change	20	20		
Coherence	Internal coherence	50	35	**75**	Level 3: moderately successful
	External coherence	50	40		
Effectiveness	Achievement of the (intended) objectives	30	30	**90**	Level 2: successful
	Contribution to achievement of objectives	30	24		
	Quality of implementation	20	18		
	Unintended results	20	19		
Impact	Higher-level (intended) development changes/results	30	30	**90**	Level 2: successful
	Contribution to higher-level (intended) development results/changes	40	30		
	Contribution to higher-level (unintended) development results/changes	30	30		
Efficiency	Production efficiency	70	65	**92**	Level 1: highly successful
	Allocation efficiency	30	27		
Sustainability	Capacities of the beneficiaries and stakeholders	20	15	**72**	Level 3: moderately successful
	Contribution to supporting sustainable capacities	30	22		
	Durability of results over time	50	35		
Mean score and overall rating		100	**86**		**Level 2: successful**

5 Conclusions and recommendations

5.1 Key findings and factors of success/failure

To facilitate learning from this evaluation, this section corroborates key factors of success and central weaknesses of the project. During the evaluation mission it became evident that key results can be centred around five (success) dimensions. Efforts and positive achievements in these dimensions (which sometimes overlap) appear to have the potential to leverage current achievements, mitigate current or future risks, or be transferred to other similar projects.

Success factors

- **DfS** and related **change management teams**: The methodological approach guaranteed bottom-up participation, mutual collaboration and a specific action plan.
- Needs assessment and **building on the awareness of challenges**: The joint analysis of potentials and needs have led to an awareness of the necessity and ownership of change processes.
- **Informal exchange**: Change processes work mainly through informal exchange between textile and garment factories.
- Establishing **mutual trust** between private and public and between private and private stakeholders: Common goals and win–win situations were identified, creating shared responsibility and ownership.
- **Business-driven approach**: Linking compliance with labour, social and environmental standards to increased productivity has led to the debunking of prevailing opinions about the increased costs of complying with standards.

Weaknesses

- **Framework conditions**: Other framework conditions, such as labour inspectors' pay and transportation, and workers' rights regarding the implementation of law, have not improved.
- Limited **practical experience**: Multipliers and labour inspectors do not have the complete set of skills and competences to successfully take up their new role in the textile and garment industry.
- **Business models** and marketing strategies for institutional multipliers: These need to be revised and improved.
- **Limited influence on higher-level policy-making:** High turnover among high-level bureaucrats and policy-makers in the province hindered further results at impact level (e.g. towards the GSP+).
- **Lack of strategic coherence**: This resulted in the absence of further synergies in the field of sustainable economic development in Pakistan.

Findings regarding 2030 Agenda

Universality, shared responsibility and accountability

The project focused on the contribution to several SDGs (1, 3, 5, 6, 8 and 12). There was a focus on SDG 8 – to 'promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all' – with specific reference to Goal 8.8 on protecting labour rights and promoting safe and secure working environments for all workers. The evaluation team identified contributions to several achievements for improved productivity and labour standards. The working conditions in 15 partner factories (approximately 20,000 people) improved in terms ergonomic chairs, personal protection equipment, safe drinking water, hygienic and subsidised lunches, regular evacuation drills, and the provision of first aid boxes and medical

attendants. Furthermore, a sample of six export-oriented factories showed a 6–10% growth in profits compared with GDP growth of 4.46%.

Other international partners are not very active in Punjab, though they are in other regions in Pakistan. Therefore, no significant coordination effort at the operational level was needed. The project established good working relations with ILO and the Embassy of the Netherlands (SDG 17). Furthermore, the project built its activities to a large extent on existing downstream departments' responsibilities (PESSI, IRI, SAA-CIWCE), and thus improved its partners' capacities to implement policies such as the Punjab Labour Policy (SDG 16).

Interplay of economic, environmental and social development

Social, economic and environmental results were strongly related within the project structure. As textile and garment factories in Pakistan have a poor efficiency into a final product, the project aimed to increase efficiency, for example to reduce solid waste by saving raw material and introducing chemicals management systems, reducing both environmental and accident risks and costs, thus increasing the competitiveness of the textile industry. Measures to increase productivity therefore simultaneously contributed to improved social, economic and environmental results, and vice versa. Further unintended positive results were found in the changing company culture. Informal exchange of best practices between partner factories became much more frequent. Furthermore, the private sector started to push the public sector for further support and increasingly demand government accountability (SDG 8).

Inclusiveness/leave no one behind

Gender is a cross-cutting social, economic and cultural issue. The project always tried to include women, closely monitored the potential for gender mainstreaming, and tried to increase the gender balance accordingly. The needs of female labour inspectors and workers were taken into account in new strategies and documents (revision of curriculum), and female participation in change management teams was ensured. However, the project realised that its influence at the gender level was limited due to the low number of female employees and labour inspectors (GIZ, 2018c).

Human rights, especially work-related human rights, are at the centre and are the main objective of a project working on labour standards. The project contributed to improved work-related human rights through improved compliance with social and labour standards by textile and garment factories. Through change management teams in textile factories, employees were sensitised to their rights, and functional relationships between employees and employers were promoted.

Other contributions to the strengthening of the supervisory and regulatory role of state institutions was limited and was outside the project's sphere of influence due to the political culture in Pakistan. However, the project improved the prerequisites for compliance to a great extent. Cooperation with the private sector in developing procedures and methods for meeting its responsibility for human rights was geared towards international standards and guidelines on the private sector's responsibility for minimum standards and human rights (e.g. UN Guiding Principles on Business and Human Rights, UN Global Compact). The project was based on international ILO conventions and Articles 6 and 7 of the International Covenant on Economic, Social and Cultural Rights.

In summary, there are many thematic overlaps with results in the SDG areas of no poverty (SDG 1), good health and wellbeing (SDG 3) and clean water and sanitation (SDG 6). Gender equality (SDG 5) and responsible consumption and production (SDG 12) are cross-cutting themes of the project. However, all results are related to decent work and economic growth (SDG 8).

5.2 Recommendations

Recommendations for similar project interventions and the design of new projects (directed to GIZ sectoral unit):

- The project was designed somewhat in isolation, as other projects in Pakistan did not focus on the textile industry or on textile workers. In future, there should be a focus on strategic coherence between projects (especially with TVET and SP-SHP).
- Based on various developments that are taking place in connection with the textile industry (Supply Chain Transparency Act), approaches to the integration of international brands and buyers may already be needed at the planning level. Action in this regard is recommended to GIZ for future projects.
- The DfS should also be considered for further projects and should be further marketed within GIZ. Other international partners have already made use of the project's experiences; therefore, the products have potential for cofinancing and further extension to other countries/regions. Coordination with the sectoral units is recommended.

Recommendation on the general project implementation and the follow-on project (directed to the project team):

- Practical involvement of institutional multipliers should be further increased in the implementation of the DfS. As planned by the project, they should gradually take on more and more responsibilities. With increased responsibility, the financial contribution of companies should also increase, so that they pay the full price at the end of the cooperation. The costs of the work of the institutional multipliers should be incorporated into a revised business model. Furthermore, the linking of multipliers and the CoP is already part of the offer for the follow-on project and is recommended by the evaluation team. In addition, consideration should be given to whether the multipliers should be more strongly involved in the cooperation with IRI in the future so that they can take over refresher workshops for labour inspectors.
- The BMZ programme develoPPP should be considered in order to test initial approaches to cooperation with international brands and buyers. An example from a partnership in Bangladesh shows promising results regarding capacity building for factory management on gender issues.
- In order to strengthen the institutionalisation of human rights, a joint strategy with other partners, such as ILO and EU, should be considered. Together, synergies and areas of responsibility can be shared in a targeted manner.
- As international partners are already building on the experience of the project, consideration should be given to whether this potential can be further exploited. For example, ILO in Sindh could also apply the DfS in a very concrete way by introducing the curriculum for the training of labour inspectors or the monitoring system for work-related accidents.

List of references

GIZ standard project documents

GIZ (2018): *LSP CD-Strategie_EN_180423.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2019): *LSP Stakeholder Landscape and RM 08052019.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2020a): *LSP Results Model.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2020b): *Indicator progress update sheets and several documents for verification.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2020c): *Updated LSP Plan of Operation 23092020.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

Other GIZ project documents

GIZ (no year): *Post assessment sheet of Labour Inspectors and Officers.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2016): *Programme proposal, Part B, TC module Improvement of Labour Standards in the Textile Industry of Pakistan, Project no. 2016.2029.3.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2017a): *2016.2029.3 PV Arbeits- und Sozialstandards Pakistan 02.11.2017.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2017b): *Programmvorschlag, Teil A, Förderung nachhaltiger Wirtschaftsentwicklung in Pakistan.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2018b): *2016.2029.3 PV Arbeits- und Sozialstandards Pakistan 25.06.2018.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2018c): Gender Analysis. Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2018d): *Vertiefte Umweltprüfung LSP.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2018e): *Integrated Context Analysis.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2018f): *Fortschrittsbericht zu einem TZ-Modul I.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2019a): *Programme proposal, Part B, TC module.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2019b): *Fortschrittsbericht zu einem TZ-Modul II*. Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2019c): *Dresden Declaration.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2020d): *Fortschrittsbericht zu einem TZ-Modul III.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2020e): *PÖK Pakistan 2020.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2020f): *Description of Results 01032019_RK.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2020g): *Angebot für die Durchführung eines TZ-Folgemoduls.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2020h): *Joint Plan of Action-Implementation of Dresden Declaration for Promoting Culture of Prevention.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2021): *Assessment of GIZ LSP Outcome 3 & Output C1, C2 and C3 Indicators.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2021a): *Fortschrittsbericht zu einem TZ-Modul IV*. Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

Other GIZ documents

GIZ (2007): *Peace and Conflict Assessment (PCA). Ein methodischer Rahmen zur konflikt- und friedensbezogenen Ausrichtung von EZ-Maßnahmen,* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2014a): *Guidelines on Designing and Using a Results-based Monitoring System,* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2017): *Guide for Central Project Evaluations,* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GIZ (2018a): *GIZ's Evaluation System – Central Project Evaluations for BMZ Business,* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH [online] www.giz.de/en/downloads/GIZ_EVAL_EN_ZPE_BMZ%20business.pdf [14.08.2019].

GIZ (2018g): *Central Project Evaluations: Implementation Guidelines*, Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, unpublished document.

GIZ (2020): *Report Writing Guidelines for GIZ Central Project Evaluations,* Eschborn/Bonn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

GTZ (2004): *Protokoll der Berliner Vereinbarung*, Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

Further references

Abacus /GIZ (2020): *DFS Research Report: Improving Labour Standards in Pakistan's Textile Industry.* Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) (2006): *Evaluierungskriterien für die deutsche bilaterale Entwicklungszusammenarbeit: Eine Orientierung für Evaluierungen des BMZ und der Durchführungsorganisationen,* BMZ [online] https://www.bmz.de/de/zentrales_downloadarchiv/erfolg_und_kontrolle/evaluierungskriterien.pdf [26.04.2021].

Board of Investment (2020): *Sector brief* [online] https://invest.gov.pk/textile#:~:text=Sector%20Brief&text=It%20is%204th%20largest%20producer,listed%20on%20the%20stock%20exchange. [10.12.2020].

Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung BMZ (2015): *Das Bündnis für nachhaltige Textilien.* Paderborn: Bonefatius.

BMZ (2015a): *Gute Arbeit im Textilsektor.* Berlin: mediabogen.

BMZ (2018): *Gemeinsame Fortschritte im Textilsektor.* Berlin: Hauer+Dörfler.

BMZ (2016) *Länderstrategie zur bilateralen Entwicklungszusammenarbeit mit der Islamischen Republik Pakistan 2016 – 2020.*

The Collaborative for Development Action Inc (CDA) (2004): *The Do No Harm Handbook*. *The Framework for Analyzing the Impact of Assistance on Conflict*, Cambridge, MA: CDA, pp. 11–12.

European Commission (2020): *The EU Special Incentive Arrangement for Sustainable Development and Good Governance ('GSP+') assessment of Pakistan covering the period 2018 – 2019* [online] https://ec.europa.eu/transparency/regdoc/rep/10102/2020/EN/SWD-2020-22-F1-EN-MAIN-PART-1.PDF [19.01.2021].

Government of Punjab (2002): *Our Secretaries*. Lahore: Labour and Human Resource Department [online] https://labour.punjab.gov.pk/oursecretaries [7.12.2020].

Government of Punjab (2020): *Textile Industries*; *Cottage and Small Scale Industries*, Lahore: Punjab Board of Investment and Trade [online] http://pbit.gop.pk/textile_ind and http://www.pbit.gop.pk/smeda [7.12.2020].

Government of Punjab (2021): https://punjab.gov.pk/provincialdepartments_labour [19.01.2021].

Government of Pakistan (2020): *Pakistan Economic Survey 2019-20*, Islamabad: Finance Division, pp. 47
Provincial Assembly of Punjab, 2020: *Members Directory,* Lahore: Provincial Assembly of Punjab [online] https://www.pap.gov.pk/members/listing/en [7.12.2020].

Kirkpatrick, James D.; Kirkpatrick, Wendy (2016) *Kirkpatrick's Four Levels of Training Evaluation.* Alexandria, USA: ATD Press.

Mayne, John (2012): *Contribution analysis: Coming of age?* Evaluation; 18 (3): 270–280.

Ministry of Planning, Development & Reform (2013): *Pakistan 2025. One Nation – One Vision.* [online] https://www.pc.gov.pk/uploads/vision2025/Pakistan-Vision-2025.pdf [26.04.2021].

Ministry of Planning, Development & Reform (2019): *Pakistan's Implementation of the 2030 Agenda for Sustainable Development Voluntary, National Review* [online] https://sustainabledevelopment.un.org/content/documents/233812019_06_15_VNR_2019_Pakistan_latest_version.pdf [20.01.2021].

Organisation for Economic Co-operation and Development (OECD) (no date): *Evaluation Criteria*, OECD [online] www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm [10.12.2020].

Provincial Assembly of Punjab (2020): *Cabinet and Functionaries*, Lahore: Provincial Assembly of [online] https://www.pap.gov.pk/about_assembly/cabinet/en [7.12.2020].

United Nations Evaluation Group (UNEG) (2014): *Integrating Human Rights and Gender Equality in Evaluations,* New York: UNEG [online] www.unevaluation.org/document/detail/1616 [30.07.2020].

Annex: Evaluation matrix

OECD-DAC Criterion Relevance - Is the intervention doing the right things? (max. 100 points)
The 'relevance' criterion focuses on the intervention's design. It refers to the extent to which the objectives and design of a development intervention are consistent with the (global, country and institution-specific) requirements, needs, priorities and policies of beneficiaries and stakeholders (individuals, groups, organisations and development partners). It also identifies the ability of the intervention's design to adapt to a change in circumstances. 'Relevance' is assessed in relation to 1) the **time of the intervention design**[1] and 2) from **today's perspective**[2].

Assessment dimensions	Filter - Project Type	Evaluation questions	Clarifications	Basis for Assessment / Evaluation indicators	Evaluation Design and empirical methods	Data sources	Data Quality and limitations	Data Quality Assessment (weak, moderate, good, strong)
Alignment with policies and priorities	Standard	To what extent are the intervention's objectives aligned with the (global, regional and country specific) policies and priorities of the BMZ and of the beneficiaries and stakeholders and other (development) partners? To what extent do they take account of the relevant political and institutional environment?	• Orientation at BMZ country strategies and BMZ sector concepts • Strategic reference framework for the project (e.g. national strategies including the national implementation strategy for Agenda 2030, regional and international strategies, sectoral and cross-sectoral change strategies, in bilateral projects especially partner strategies, internal analytical framework e.g. safeguards and gender[4] • Orientation of the project design at the (national) objectives of Agenda 2030 • Project contribution to certain Sustainable Development Goals (SDGs) • Explanation of a hierarchy of the different policies, priorities (especially in case of contradictions)	Comparison of objectives and goals between project and frameworks	Document review and criteria-led analysis	BMZ country strategy for Pakistan (2016-2020) BMZ sectoral strategies: • Bündnis für nachhaltige Textilien • Gemeinsame Fortschritte im Textilsektor • Zukunftspapier: Gute Arbeit weltweit Pakistan Vision 2025 (also as national strategy for Agenda 2030) Documents on EU GSP+ with Pakistan	• No foreseen limitations	strong
	and Fragility	To what extent was the (conflict) context of the project adequately analysed and considered for the project concept?	• Key documents: (Integrated) Peace and Conflict Assessment (I)PCA, Safeguard Conflict and Context Sensitivity documents	Context documents exists and their recommendations are considered	Document review and criteria-led analysis	PCA. Gender analysis, environmental assessment document	• No foreseen limitations	moderate
	and SV/GV	To what extent does the project complement bilateral or regional projects?	• Please use CPE factsheet on SV / GV / IZR	Perception of key partners	Document analysis, Interviews and focus group discussions	Direct target group: • Staff of the LHRD and subsidiary institutions (IRI, SAA-	• Representation of indirect stakeholders need to be identified • Data triangulation	good

		To what extent does it complement other global projects?				CIWCE) • Implementing partners from the District Labour Offices Management of factories in the textile sector Indirect target group • Employees in the textile sector with a focus on the province of Punjab	between primary and secondary data	
	and SV/GV	To what extent is the project geared towards solving a global challenge that cannot only be effectively addressed bilaterally/ regionally?	• Please use CPE factsheet on SV / GV / IZR	Comparison of proposal and progress reports with PCA column I and II and other documents regarding conflict assessment	Document analysis, Interviews and focus group discussions	PCA. Gender analysis, environmental assessment document		moderate
Alignment with the needs and capacities of the beneficiaries and stakeholders	Standard	To what extent are the intervention's objectives aligned with the development needs and capacities of the beneficiaries and stakeholders involved (individuals, groups and organisations)?	• Also: consideration of stakeholders such as civil society and private sector in the design of the measure	Comparison of proposal and progress reports with PCA column I and II and other documents regarding conflict assessment	Document analysis, Interviews	PCA. Gender analysis, environmental assessment document		moderate
	and Fragility	How were deescalating factors/ connectors[5] as well as escalating factors/ dividers[6] in the project context identified and considered for the project concept (please list the factors)?[7]	• e.g. see column I and II of the (Integrated) Peace and Conflict Assessment	Disadvantaged groups are considered in key project documents	Document analysis, Interviews and focus group discussions	Direct target group: • Staff of the LHRD and subsidiary institutions (IRI, SAA-CIWCE) • Implementing partners from the District Labour Offices Management of factories in the textile sector Indirect target group • Employees in the textile sector with a focus on the province of Punjab	• Representation of indirect stakeholders need to be identified • Data triangulation between primary and secondary data	good
	and Fragility	To what extent were potential (security) risks for (GIZ) staff, partners, target groups/final beneficiaries identified and considered?		The proposal was realistic and results and indicators are achievable Documents were updated with exchange offers.	Analysis of updated proposals, change offers and related indicators	Proposals and change offers 2017, 2018, 2019 Capacity Development Strategy Change offers	· No foreseen limitations	good

		Standard	To what extent are the intervention's objectives geared to the needs and capacities of particularly disadvantaged and vulnerable beneficiaries and stakeholders (individuals, groups and organisations)? With respect to groups, a differentiation can be made by age, income, gender, ethnicity, etc. ?	• Reaching particularly disadvantaged groups (in terms of Leave No One Behind, LNOB) • Consideration of potential for human rights and gender aspects • Consideration of identified risks	The results model represents the project logic in an adequate way. Documents were updated with exchange offers.	Document review; Interviews	Results model (including results hypotheses) Capacity Development Strategy Change offers	· No foreseen limitations	good
	Appropriateness of the design[3]	Standard	To what extent is the intervention's design appropriate and realistic (in terms of technical, organisational and financial aspects)?	• Realistic project goal from today's perspective and in view of the available resources (time, finances, partner capacities) • Consideration of potential changes in the framework conditions • Dealing with the complexity of framework conditions and strategic reference frameworks and with possible overloading • Strategic focusing	Synergies and trade-offs are considered within the project design	Analysis of updated proposals, change offers and related indicators	Proposals and change offers 2017, 2018, 2019 Capacity Development Strategy Change offers	· No foreseen limitations	good
		Standard	To what extent is the intervention's design sufficiently precise and plausible (in terms of the verifiability und traceability of the system of objectives and the underlying assumptions)?	Assessment of the (current) results model and results hypotheses (Theory of Change, ToC) of the actual project logic: • Adequacy of activities, instruments and outputs in relation to the project objective to be achieved • Plausibility of the underlying results hypotheses • Clear definition and plausibility of the selected system boundary (sphere of responsibility) • Appropriate consideration of potential influences of other donors/	change offers represent sectoral knowledge	Analysis of updated proposals, change offers and related indicators; Interviews	Three change offers from 2017, 2018 and 2019 BMZ, FMB	· No foreseen limitations	good

			organisations outside the project's sphere of responsibility • completeness and plausibility of assumptions and risks for the project results • How well is cofinancing (if any) integrated into the overall concept of the project and what added value could be generated for the ToC/project design?					
	Standard	To what extent is the intervention's design based on a holistic approach to sustainable development (interaction of the social, environmental and economic dimensions of sustainability)?	• Presentation of the interactions (synergies/trade-offs) of the intervention with other sectors in the project design - also with regard to the sustainability dimensions in terms of Agenda 2030 (economic, ecological and social development)					
Adaptability – response to change	Standard	To what extent has the intervention responded to changes in the environment over time (risks and potentials)?	• Reaction to changes during project including change offers (e.g. local, national, international, sectoral changes, including state-of-the-art sectoral know-how)	change offers represent sectoral knowledge	Analysis of updated proposals, change offers and related indicators; Interviews	Three change offers from 2017, 2018 and 2019 BMZ, FMB	· No limitations	strong

(1) The 'time of the intervention design' is the point in time when the offer/most recent modification offer was approved.

(2) In relation to the current standards, knowledge and framework conditions.

(3) The design of an intervention is usually assessed by evaluating its intervention logic. The intervention logic depicts the system of objectives used by an intervention. It maps out the systematic relationships between the individual results levels. At the time an intervention is designed, the intervention logic, in the form of a logical model, is described in the offer for the intervention both as a narrative and generally also on the basis of a results framework. The model is reviewed at the start of an evaluation and adjusted to reflect current knowledge. Comprehensive (re)constructed intervention logics are also known as 'theories of change'. In GIZ the 'project design' encompasses project objective (outcome) and the respective theory of change (ToC) with outputs, activities, TC instruments and especially the results hypotheses as well as the implementation strategy (e.g. methodological approach, Capacity Development (CD) strategy). In GIZ the ToC is described by the GIZ results model as graphic illustration and the narrative results hypotheses.

(4) In the GIZ Safeguards and Gender system risks are assessed before project start regarding following aspects: gender, conflict, human rights, environment and climate. For the topics gender and human rights not only risks but also potentials are assessed. Before introducing the new safeguard system in 2016 GIZ used to examine these aspects in separate checks.

(5) Deescalating factors/ connectors: e.g. peace-promoting actors and institutions, structural changes, peace-promoting norms and behaviour. For more details on 'connectors' see: GIZ (2007): 'Peace and Conflict Assessment (PCA). Ein methodischer Rahmen zur konflikt- und friedensbezogenen Ausrichtung von EZ-Maßnahmen', p. 55/135.

(6) Escalating factors/ dividers: e.g. destructive institutions, structures, norms and behaviour. For more details on 'dividers' see: GIZ (2007): 'Peace and Conflict Assessment (PCA). Ein methodischer Rahmen zur konflikt- und friedensbezogenen Ausrichtung von EZ-Maßnahmen', p. 135.

(7) All projects in fragile contexts, projects with FS1 or FS2 markers and all transitional aid projects have to weaken escalating factors/dividers and have to mitigate risks in the context of conflict, fragility and violence. Projects with FS1 or FS2 markers should also consider how to strengthen deescalating factors/ connectors and how to address peace needs in its project objective/sub-objective.

OECD-DAC Criterion Coherence - How well does the intervention fit? (max. 100 points)

Assessment dimensions	Filter - Project Type	Evaluation questions	Clarifications	Basis for Assessment / Evaluation indicators	Evaluation Design and empirical methods	Data sources	Data Quality and limitations	Data Quality Assessment
Internal coherence	Standard	Within German development cooperation, to what extent is the intervention designed and implemented (in a sector, country, region or globally) in a complementary manner, based on the division of tasks?	• Also analysis of whether the project takes the necessary steps to fully realise synergies within German development cooperation	*Synergies with further projects within the textile cluster*	Interviews	Förderung von Nachhaltigkeit in der Textil- und Bekleidungsindustrie in Asien (FABRIC, PN 2019.2141.0 Other GIZ projects within the textile cluster: • Supporting Technical and Vocational Education and Training (TVET) Reform (PN 2016.2042.2) • Support to Social Security including Health Insurance in Pakistan (PN 2015.2186.3). • The regional programme Social and Labour Standards in the Textile and Garment Sector in Asia (PN 2014.2279.9) Further sector projects, e.g. • SV Förderung von Multi-Akteurs-Projekten für nachhaltige Textil-Lieferketten • SV Nachhaltigkeit in Textil-Lieferketten Bündnis für nachhaltige Textilien / Grüner Knopf	Data collection with each project representative might not have a strong evidence, therefore representatives from the GIZ sectoral unit might provide more reliable information	strong
	Standard	To what extent are the instruments of German development cooperation (Technical and Financial Cooperation) meaningfully interlinked within the intervention (in terms of both design and implementation)? Are synergies leveraged?	• if applicable, also take into account projects of different German ressorts/ministries	Alignment with further projects within the textile cluster	Interviews	Förderung von Nachhaltigkeit in der Textil- und Bekleidungsindustrie in Asien (FABRIC, PN 2019.2141.0 Other GIZ projects within the textile cluster: • Supporting Technical and Vocational Education and Training (TVET) Reform (PN 2016.2042.2) • Support to Social Security including Health Insurance in Pakistan (PN 2015.2186.3). • The regional programme Social and Labour	Data collection with each project representative might not have a strong evidence, therefore representatives from the GIZ sectoral unit might provide more reliable information	strong

						Standards in the Textile and Garment Sector in Asia (PN 2014.2279.9) Further sector projects, e.g. • SV Förderung von Multi-Akteurs-Projekten für nachhaltige Textil-Lieferketten • SV Nachhaltigkeit in Textil-Lieferketten Bündnis für nachhaltige Textilien / Grüner Knopf		
	Standard	To what extent is the intervention consistent with international and national norms and standards to which German development cooperation is committed (e.g. human rights)?		Comparison with standards German development cooperation committed with	document analysis, interviews	international ILO convention and articles 6 and 7 of the International Covenant on Economic, Social and Cultural Rights	Data collection with each project representative might not have a strong evidence, therefore representatives from the GIZ sectoral unit might provide more reliable information	moderate
External coherence	Standard	To what extent does the intervention complement and support the partner's own efforts (principle of subsidiarity)?		Ownerships and Motivation of project partners is high	Interviews	Direct target group: private sector factories, public sector partners	no foreseen limitations	strong
	Standard	To what extent has the intervention's design and implementation been coordinated with other donors' activities?	• Also: To what extent could synergies be achieved through cofinancing (where available) with other bilateral and multilateral donors and organisations and how did cofinancing contribute to improved donor coordination?	Synergies identified and used	Interviews	Government of Netherlands • Buyers' Forum EU Sozial- und Umweltstandards in der Textil- und Bekleidungsindustrie • Project: ILES Government of Japan • Project for Skills Development and Market Diversification of Garment Industry of Pakistan (PSDMD)	no foreseen limitations	good
	Standard	To what extent has the intervention's design been designed to use existing systems and structures (of partners/other donors/international organisations) for implementing its activities? To what extent are these systems and structures used?	• Also analysis of whether the project is taking the necessary steps to fully realise synergies with interventions of other donors at the impact level	Cooperation with further systems and other partners outside GIZ and German DC	Interviews	Government of Netherlands Buyers' Forum EU Sozial- und Umweltstandards in der Textil- und Bekleidungsindustrie Project: ILES Government of Japan Project for Skills Development and Market Diversification of Garment Industry of Pakistan (PSDMD)	no foreseen limitations	good

	Standard	To what extent are common systems (together with partners/other donors/international organisations) used for M&E, learning and accountability?		Data and knowledge exchange with other partners	Interviews	ILO	no foreseen limitations	good

OECD-DAC Criterion Effectiveness

Assessment dimensions	Filter - Project Type	Evaluation questions	Clarifications	Basis for Assessment / Evaluation indicators	Evaluation Design and empirical methods	Data sources	Data Quality and limitations	Data Quality Assessment
Achievement of the (intended) objectives[1]	Standard	To what extent has the intervention achieved, or is the intervention expected to achieve, the (intended) objectives as originally planned (or as modified to cater for changes in the environment)?	• Assessment based on the project objective indicators (agreed with BMZ) • Check whether more specific or additional indicators are needed to adequately reflect the project objective	Achievement of project objective indicators	analysis of project's monitoring data, Interview	• Indicator Progress Update Sheets and RBM system • Perception of key partners, perception of project team members • SMART* criteria have been met.	Low number of survey participants	moderate
	and Fragility	For projects with FS1 or FS2 markers: To what extent was the project able to strengthen deescalating factors/ connectors?[2, 4]						
Contribution to achievement of objectives	Standard	*To what extent have the intervention's outputs been delivered as originally planned (or as modified to cater for changes in the environment)?*		Achievement of output indicators	Document analysis, analysis of project's monitoring data	• Indicator Progress Update Sheets and RBM system • Perception of key partners, perception of project team members • SMART* criteria have been met.	No limitations	strong
	Standard	To what extent have the delivered outputs and increased capacities been used and equal access (e.g. in terms of physical, non-discriminatory and affordable access) guaranteed?		Training effectiveness is proven (Kirkpatrick); Evidence for hypotheses established/rejected As capacity building is a crucial part of the indicators. The dimension will be assed against the project's indicators (A2, B2, C1, C2, C3) and their maintained application (MO3, MO4)	Contribution analysis	Interviews with all key stakeholders, validation workshop with project team	No limitations	strong

				Kirkpatrick Level 3 (behaviour)				
	Standard	To what extent has the intervention contributed to the achievement of objectives?	• Assessment based on the activities, TC instruments and outputs of the project (contribution analysis as focus of this assessment dimension and minimum standard, see annotated reports) • What would have happened without the project? (usually qualitative reflection)	Results Hypothesis output outcome	Contribution analysis	Interviews with all key stakeholders, validation workshop with project team		
	Standard	To what extent has the intervention contributed to the achievement of objectives at the level of the intended beneficiaries?		Results hypothesis outcome - impact	Contribution analysis	Interviews with all key stakeholders, validation workshop with project team	No limitations	strong
	Standard	To what extent has the intervention contributed to the achievement of objectives at the level of particularly disadvantaged or vulnerable groups of beneficiaries and stakeholders? (These may be broken down by age, income, gender, ethnicity, etc.)?		Assessment of do-no-harm and unintended impact	Contribution analysis	Interviews with all key stakeholders, validation workshop with project team	No limitations	good
	Standard	*Which internal factors (technical, organisational or financial) were decisive for achievement/non-achievement of the intervention's intended objectives?*	• Internal factors = within the project's sphere of responsibility / system boundary. The project is implemented jointly by GIZ and the official partner(s).	Assessment of external factors	Contribution analysis	Interviews with all key stakeholders, validation workshop with project team	No limitations	strong
	Standard	*Which external factors were decisive for achievement/non-achievement of the intervention's intended objectives (taking into account the anticipated risks)?*	• External factors = outside the project's sphere of responsibility / system boundary. The project is implemented jointly by GIZ and the official partner(s).	Assessment of external factors	Contribution analysis	Interviews with all key stakeholders, validation workshop with project team	No limitations	strong
Quality of implementation	Standard	What assessment can be made of the quality of steering and implementation of the intervention in terms of the achievement of objectives? What assessment can be made of the quality of steering and implementation of, and participation in, the intervention by the partner/executing agency?	Capacity WORKS considerations: - **Results-oriented monitoring (RoM / WoM)** is established and used, e.g. for evidence-based decisions, risk management. Data are disaggregated by gender and marginalised groups. unintended positive and negative results are monitored. Conflict-sensitive monitoring and explicit risk-safety monitoring are particularly important for projects in fragile contexts. - A bindingly communicated **strategy** agreed with the partners is pursued - Involvement and **cooperation** of all relevant actors (including partners, civil society, private sector) - **Steering**: decisions influencing the project's results are made in time and evidence-informed. Decision	Capacity works is established and used	document analysis, interviews	Capacity works considerations: • RBM System • Capacity development strategy • Plan of Operations • Involvement of all relevant stakeholder	No limitations	good

			processes are transparent. - **Processes**: Relevant change processes are anchored in the cooperation system; project-internal processes are established and regularly reflected and optimised. - **Learning and innovation**: There is a learning and innovation-friendly work culture that promotes the exchange of experience; learning processes are established; context-specific adjustments are possible					
Unintended results	Standard	To what extent can unintended positive/negative direct results (social, economic, environmental and among vulnerable beneficiary groups) be observed/anticipated?	• The focus is on the outcome level, but for the analysis the unintended effects can also be included on the output level	Most Significant Change - Motivation and ownership of project partners	Focus group discussion, interviews, document analysis	Interviews with all key stakeholders, validation workshop with project team	anecdotal evidence expected	good
	and Fragility	To what extent was the project able to ensure that escalating factors/ dividers[3] have not been strengthened (indirectly) by the project[4]? Has the project unintentionally (indirectly) supported violent or 'dividing' actors?		Assessment of monitoring system on conflict, fragility and violence	Focus group discussion, interviews, document analysis	Interviews with all key stakeholders, validation workshop with project team		good
	Standard	What potential benefits/risks arise from the positive/negative unintended results? What assessment can be made of them?	• also check whether the risks were already mentioned and monitored in the design phase	Conflict sensitivity and human rights (2020): Impact on human rights regulation in partner factories; do-no-harm analysis, mitigation measures	document analysis, interviews	Activity reports, progress reports	anecdotal evidence expected	moderate
	and Fragility	To what extent have risks and unintended negative results in the context of conflict, fragility and violence[5] been monitored (context/conflict-sensitive monitoring) in a systematic way?		Conflict sensitivity and human rights (2020): Impact on human rights regulation in partner factories; do-no-harm analysis, mitigation measures	document analysis, interviews	Activity reports, progress reports	anecdotal evidence expected	moderate
	Standard	How has the intervention responded to the potential benefits/risks of the positive/negative unintended results?	• Check if positive results at the outcome level have been monitored and set in value	Additional results are identified; Update of results model within the change offers	document analysis, interviews	Activity reports, progress reports	anecdotal evidence expected	moderate

OECD-DAC Criterion Impact

Assessment dimensions	Filter - Project Type	Evaluation questions	Clarifications	Basis for Assessment / Evaluation indicators	Evaluation Design and empirical methods	Data sources	Data Quality and limitations	Data Quality Assessment
Higher-level (intended) development changes[1]	Standard	To what extent can the higher-level development changes (social, economic and environmental dimensions and the interactions between them) to which the intervention will/is designed to contribute be identified/foreseen)? (Specify time frame where possible.)	• Consider module proposal for suggested impact and program objective indicators (program proposal), if it is not an individual measure • Potential basis for assessment: program objective indicators, identifiers, connection to the national strategy for implementing 2030 Agenda , connection to SDGs	Degree of contribution to Programme indicators; Overarching development results the project is contributing to	Interviews with all key stakeholders, validation workshop with project team	Programme objective indicators Further impact on: • SDG 8 • Vision 2025 • GSP+	Programme is limited to Punjab Anecdotal evidence expected as robust statistical data is expected to be missing	Good
	Standard	To what extent can the higher-level development changes (social, economic, environmental dimensions and the interactions between them) be identified/foreseen at the level of the intended beneficiaries? (Specify time frame where possible.)		Degree of contribution at target group level; Perception of partners on impact for final beneficiaries	Interviews with all key stakeholders, validation workshop with project team	Further impact on: • SDG 8 • Vision 2025 • GSP+ • TD 2 • PD/GG 1 • AO 1	see above	Good
	Standard	To what extent can higher-level development changes to which the intervention will/is designed to contribute be identified/foreseen at the level of particularly disadvantaged/vulnerable groups of beneficiaries and stakeholders? (These may be broken down by age, income, gender, ethnicity, etc.) (Specify time frame where possible.)		Perception of partners on impact for vulnerable target groups	Interviews	• TD 2 • PD/GG 1 • AO 1	see above	Good
Contribution to higher-level (intended) development changes	Standard	To what extent has the intervention actually contributed to the identified and/or foreseeable higher-level development changes (social, economic, environmental dimensions and their interactions, taking into account political stability) that it was designed to bring about?	• Contribution analysis (evaluation design) as minimum standard and focus of this assessment dimension, further approaches are possible and welcome, see also annotated reports • Evaluation of the project's contribution to impacts based on an analysis of the results hypotheses from outcome to impact level	Achievement of outcome indicators	Contribution analysis	Interviews with all key stakeholders, validation workshop with project team,	Too early for development results to be established	good

Standard	To what extent has the intervention achieved its intended (original and, where applicable, revised) development objectives?	• This question can already be assessed in Dimension 1 Question 1, the contribution to impact is assessed in Dimension 2, Question 1	Training effectiveness is proven (Kirkpatrick); Evidence for hypotheses established/rejected As capacity building is a crucial part of the indicators. The dimension will be assed against the project's indicators (A2, B2, C1, C2, C3) and their maintained application (MO3, MO4) Kirkpatrick Level 3 (behaviour)	Contribution analysis	Interviews with all key stakeholders, validation workshop with project team, online survey	see above	good
Standard	To what extent has the intervention achieved its (original and, where applicable, revised) development objectives at the level of the intended beneficiaries?		Training effectiveness is proven (Kirkpatrick); Evidence for hypotheses established/rejected (Behaviour and results)	Contribution analysis	Interviews with all key stakeholders, validation workshop with project team, project's monitoring data, online survey	see above	good
Standard	To what extent has the intervention contributed to higher-level development changes/changes in the lives of particularly disadvantaged or vulnerable groups of beneficiaries and stakeholders that it was designed to bring about? (These may be broken down by age, income, gender, ethnicity, etc.).		Do-no-harm analysis	Contribution analysis	Interviews with all key stakeholders, validation workshop with project team	see above	Good
Standard	*Which internal factors (technical, organisational or financial) were decisive for achievement/non-achievement of the intervention's intended development objectives?*	• Internal factors = within the project's sphere of responsibility / system boundary. The project is implemented jointly by GIZ and the official partner(s)	Assessment of internal factors	Contribution analysis	Interviews with all key stakeholders, validation workshop with project team,	see above	Good
Standard	*Which external factors were decisive for the achievement/non-achievement of the intervention's intended development objectives?*	• External factors = outside the project's sphere of responsibility / system boundary. The project is implemented jointly by GIZ and the official partner(s). • Take into account the activities of other actors or other policies, framework conditions, other policy areas, strategies or interests (German ministries, bilateral and multilateral development partners)	Influence of framework conditions	Contribution analysis	Interviews with all key stakeholders, validation workshop with project team,	see above	Good

	Standard	To what extent has the intervention achieved structural or institutional changes (e.g. for organisations, systems and regulations)?		Training effectiveness is proven (Kirkpatrick); Evidence for hypotheses established/rejected (Behaviour and results)	Contribution analysis	Interviews with all key stakeholders, validation workshop with project team,	see above	good
	Standard	To what extent did the intervention serve as a model and/or achieve broad-based impact?	• Scaling-up is a consciously designed process to anchor changes in organisations and cooperation systems (e.g. concepts, approaches, methods) to generate broad impact • There is vertical scaling-up, horizontal scaling-up, functional scaling-up or a combination of these[2] • also analyse possible potential and reasons for not exploiting it	Scaling-up assessment - number of factories	Contribution analysis	Interviews with all key stakeholders, validation workshop with project team,	see above	Good
	Standard	*How would the situation have developed without the intervention?*	• usually qualitative reflection, quantitative approaches welcome	Counterfactual situation	Contribution analysis	Interviews with all key stakeholders, validation workshop with project team,	see above	good
Contribution to higher-level (unintended) development changes	Standard	To what extent can higher-level, unintended development changes (social, economic and environmental dimensions and their interactions, taking into account political stability) be identified/foreseen? (Specify time frame where possible.)		Evidence for widespread impact established	Most Significant Change - Focus group discussion, interviews	FGD	Anecdotal evidence expected	Good
	and Fragility	To what extent did the project have (unintended) negative or escalating effects on the conflict or the context of fragility (e.g. conflict dynamics, violence, legitimacy of state and non-state actors/institutions)? To what extent did the project have positive or deescalating effects on the conflict or the context of fragility (e.g. conflict dynamics, violence, legitimacy of state and non-state actors/institutions)?		Mitigation measures mentioned	Document analysis	proposal documents		Good

	Standard	To what extent has the intervention brought about foreseeable/identifiable unintended (positive and/or negative) higher-level development results?	• Analyse whether the risks were already known in the design phase • Check how the assessment of risks in connection with (unintended) negative or (not formally agreed) positive results at the impact level in the monitoring system has been carried out (e.g. use of 'compass') • measures taken to avoid or counteract the risks/ negative effects/ trade-offs[3] • Determine relevant framework conditions for negative results and the project's reaction to them • Examine to what extent potential (not formally agreed) positive results and synergies between the ecological, economic and social development dimensions have been monitored and exploited	Degree of assessment ing monitoring tools and within proposal documents	Document analysis of monitoring documents	monitoring system web-tool, indicator progress sheets, progress reports	Anecdotal evidence expected	good
	Standard	To what extent has the intervention contributed to foreseeable/identifiable unintended (positive and/or negative) higher-level development results at the level of particularly disadvantaged or vulnerable groups of beneficiaries and stakeholders? (These may be broken down by age, income, gender, ethnicity, etc.)		Do-no-harm analysis,	Focus group discussion, interviews	FGD	Anecdotal evidence expected	good

OECD-DAC Criterion Efficiency - How well are resources being used? (max. 100 points)

Assessment dimensions	Filter - Project Type	Evaluation questions	Clarifications	Basis for Assessment / Evaluation indicators	Evaluation Design and empirical methods	Data sources	Data Quality and limitations	Data Quality Assessment
Production efficiency	Standard	*How are the intervention's inputs (financial, human and material resources) distributed (e.g. by instruments, sectors, sub-interventions, taking into account the cost contributions of partners/executing agencies/other beneficiaries and stakeholders etc.)?*	• Description of the data: Costs per output, type of costs, agreed and provided partner contributions • Description of the deviations between original planned costs and actual costs (with comprehensible justification, changes are certainly desirable for increased efficiency)	*Transformation of inputs to outputs based on:* • *GIZ efficiency tool*	• Follow-the-money approach; Interviews, document analysis	• 'Kostenträger-Obligo" report of the project, • the comparison of planned budget figures with actual figures, • the results matrix • Progress reports Project	No limitations	good

						management and team		
	Standard	To what extent have the intervention's inputs (financial, human and material resources) been used economically in relation to the outputs delivered (products, investment goods and services)? If possible, refer to data from other evaluations in a region or sector, for instance.	• Use of 'Efficiency tool' including instructions and use of the follow-the-money approach as evaluation design (may be combined with other high-quality approaches) • Output level: Analysis of approaches and activities as well as TC instruments (personnel instruments, financing, materials and equipment)[1] compared to possible alternatives with a focus on the minimum principle (use of comparative data if available) • The project is oriented on internal or external benchmarks in order to achieve its effects economically • Regular reflection of the resources used by the project with focus on economically use of resources and cost risks • The overarching costs of the project are in an appropriate proportion to the costs of the outputs	*Transformation of inputs to outputs based on:* • *GIZ efficiency tool*	Follow-the-money approach; Interviews, document analysis	• 'Kostenträger-Obligo" report of the project, • the comparison of planned budget figures with actual figures, • the results matrix • Progress reports Project management and team	No limitations	good
	Standard	To what extent could the intervention's outputs (products, investment goods and services) have been increased through the alternative use of inputs (financial, human and material resources)? If possible, refer to data from other evaluations of a region or sector, for instance. (If applicable, this question adds a complementary perspective*) * This case is always applicable in the technical cooperation (TC), please answer the question bindingly	• Use of 'Efficiency tool' including instructions and use of the follow-the-money approach as evaluation design (may be combined with other high-quality approaches) • Output level: Analysis of approaches and activities as well as TC instruments (personnel instruments, financing, materials and equipment)[1] compared to possible alternatives with focus on output maximisation (use of comparative data if available) • Analysis of alternative options for allocating resources and shifts between outputs for output maximisation • saved resources can and should be used to maximise outputs • Reflection of the resources during the design phase and regularly during the implementation of the project with focus on output maximisation (with comprehensible justification, changes are certainly desirable for increased efficiency) • 'maximising outputs' means with the same resources, under the same conditions and with the same or better quality	*Transformation of inputs to outputs based on:* • *GIZ efficiency tool*	Follow-the-money approach; Interviews, document analysis	• 'Kostenträger-Obligo" report of the project, • the comparison of planned budget figures with actual figures, • the results matrix • Progress reports Project management and team	No limitations	good
	Standard	Were the outputs (products, investment goods and services) produced on time and within the planned time frame?		Assessment of indicator achievement	Follow-the-money approach; Interviews, document analysis	progress reports, indicator progress update sheets	No limitations	good
Allocation efficiency	Standard	*By what other means and at what cost could the results achieved (higher-level project objective) have been attained?*		Transformation of inputs to outcome based on: • GIZ efficiency tool	Follow-the-money approach; Interviews, document analysis	Further interviews with key stakeholders	Qualitative findings are considered plausible assumptions and anecdotal evidence	good

	Standard	To what extent – compared with alternative designs for the intervention – could the results have been attained more cost-effectively?	• Outcome level: Analysis of approaches and activities as well as TC instruments in comparison to possible alternatives with focus on minimum principle (use of comparative data if available) • Regular reflection in the project of the input-outcome relation and alternatives as well as cost risks • The partner contributions are proportionate to the costs for the outcome of the project	Transformation of inputs to outcome based on: • GIZ efficiency tool	Follow-the-money approach; Interviews, document analysis	Further interviews with key stakeholders		
	Standard	To what extent – compared with alternative designs for the intervention – could the positive results have been increased using the existing resources? (If applicable, this question adds a complementary perspective*) * This case is always applicable in the technical cooperation (TC), please answer the question bindingly	• Outcome level: Analysis of applied approaches and activities as well as TC instruments compared to possible alternatives with focus on maximising the outcome (real comparison if available) • The project manages its resources between the outputs in such a way that the maximum effects in terms of the module objective are achieved • Regular reflection in the project of the input-outcome relation and alternatives • Reflection and realisation of possibilities for scaling up • If additional funds (e.g. cofinancing) have been raised: Effects on input-outcome ratio (e.g. via economies of scale) and the ratio of administrative costs to total costs • Losses in efficiency due to insufficient coordination and complementarity within German DC are sufficiently avoided	Transformation of inputs to outcome based on: • GIZ efficiency tool	Follow-the-money approach; Interviews, document analysis	Further interviews with key stakeholders	Qualitative findings are considered plausible assumptions and anecdotal evidence	good

OECD-DAC Criterion Sustainability - Will the benefits last?

Assessment dimensions	Filter - Project Type	Evaluation questions	Clarifications	Basis for Assessment / Evaluation indicators	Evaluation Design and empirical methods	Data sources	Data Quality and limitations	Data Quality Assessment
Capacities of the beneficiaries and stakeholders	Standard	To what extent do the beneficiaries and stakeholders (individuals, groups and organisations, partners and executing agencies) have the institutional, human and financial resources as well as the willingness (ownership) required to sustain the positive results of the intervention over time (once assistance has drawn to a close)?	• Transitional Development Assistance (TDA) projects primarily address final beneficiaries, whose resilience to crises and recurring shocks is to be strengthened. The focus for TDA projects is thus often on the resilience of final beneficiaries and/or at least the continuity of the measure (see explanation in dimension 3) (clarification in the inception phase of the evaluation).	*Training effectiveness (Kirkpatrick); behaviour and results* *As capacity building is a crucial part of the indicators. The dimension will be assed against the project's indicators (A2, B2, C1, C2, C3) and their maintained application (MO3, MO4)*	Endline-assessment, Interviews	Key project partners	No limitations	strong
	Standard	To what extent do the beneficiaries and stakeholders (individuals, groups and organisations, partners and executing agencies) have the resilience to overcome future risks that could jeopardise the intervention's results?		Training effectiveness (Kirkpatrick); behaviour and results As capacity building is a crucial part of the indicators. The dimension will be assed against the project's indicators (A2, B2, C1, C2, C3) and their maintained application (MO3, MO4)	**Endline-assessment, Interviews**	Key project partners	No limitations	strong

Contribution to supporting sustainable capacities	Standard	To what extent has the intervention contributed to the beneficiaries and stakeholders (individuals, groups and organisations, partners and executing agencies) having the institutional, human and financial resources as well as the willingness (ownership) required to sustain the intervention's positive results over time and to limit the impact of any negative results?	• Analysis of the preparation and documentation of learning experiences • Description of the anchoring of contents, approaches, methods and concepts in the partner system • Reference to exit strategy of the project • If there is a follow-on project, check to what extent the results of the evaluated project are taken up; the anchoring of the effects in the partner's organisation should be pursued independently of a follow-on project, since sustainability should be achieved even without donor funds • Transitional Development Assistance (TDA) projects primarily address final beneficiaries, whose resilience to crises and recurring shocks is to be strengthened. The focus for TDA projects is thus often on the resilience of final beneficiaries and/or at least the continuity of the measure (see explanation in dimension 3) (clarification in the inception phase of the evaluation).	Perception on project's exit strategy	**Endline-assessment, Interviews**	Key project partners	Anecdotal evidence	good
	Standard	To what extent has the intervention contributed to strengthening the resilience of the beneficiaries and stakeholders (individuals, groups and organisations, partners and executing agencies)?		Training effectiveness (Kirkpatrick); behaviour and results As capacity building is a crucial part of the indicators. The dimension will be assed against the project's indicators (A2, B2, C1, C2, C3) and their maintained application (MO3, MO4)	**Endline-assessment, Interviews**	Key project partners	Anecdotal evidence	good
	Standard	To what extent has the intervention contributed to strengthening the resilience of particularly disadvantaged groups? (These may be broken down by age, income, gender, ethnicity, etc.)		Explorative question	**FGD**	Female beneficiaries and factory workers	Anecdotal evidence	good
Durability of results over time	Standard	*How stable is the context in which the intervention operates?*		Perception of sustainability of conflict factors	Interviews	Interviews with key partners, validation with project team	Anecdotal evidence	good
	Standard	*To what extent is the durability of the intervention's positive results influenced by the context?*	• Consideration of risks and potentials for the long-term stability of the results and description of the reaction of the project to these	Perception of sustainability of conflict factors	Interviews	Interviews with key partners, validation with project team	Anecdotal evidence	good
	Standard	To what extent can the positive (and any negative) results of the intervention be deemed durable?	• Consideration of the extent to which continued use of the results by partners and beneficiaries can be foreseen • Reference to conditions and their influence on the durability, longevity and resilience of the effects (outcome and impact) • In the case of projects in the field of Transitional Development Assistance (TDA), at least the continuity of the measure must be examined: To what extent will services or results be continued in future projects (of GIZ or other donors/organisations) or their sustainability ensured? (Clarification in the inception phase)	Perception of sustainability of conflict factors	Interviews	Interviews with key partners, validation with project team	Anecdotal evidence	good

Photo credits and sources

Disclaimer:

Maps:

Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH

Registered offices
Bonn and Eschborn

Friedrich-Ebert-Allee 32 + 36
53113 Bonn, Germany
T: +49 228 44 60-0
F: +49 228 44 60-17 66

Dag-Hammarskjöld-Weg 1–5
65760 Eschborn, Germany
T: +49 6196 79-0
F: +49 6196 79-11 15

E: info@giz.de
I: www.giz.de

giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH