

FRAGILE

GERMAN DEVELOPMENT COOPERATION IN FRAGILE CONTEXTS

Executive Summary

2019



DEval

GERMAN
INSTITUTE FOR
DEVELOPMENT
EVALUATION

Imprint

Published by

German Institute for Development
Evaluation (DEval)
Fritz-Schäffer-Straße 26
53113 Bonn, Germany
Phone: +49 (0)228 33 69 07-0
E-Mail: info@DEval.org
www.DEval.org

Authors

Dr Thomas Wencker
Ida Verspohl

Responsible

Dr Sven Harten

Design

MedienMélange:Kommunikation!, Hamburg
www.medienmelange.de

Editing

Martin Rickerd OBE MVO

Photo credits

Benjamin Nelan on Pixabay

Bibliographical reference

Wencker, T. and I. Verspohl (2019), German
Development Cooperation in Fragile Contexts,
German Institute for Development Evaluation
(DEval), Bonn.

Printing

Bonifatius, Paderborn



© German Institute for Development
Evaluation (DEval), 2019

ISBN 978-3-96126-091-1 (printed edition)
ISBN 978-3-96126-092-8 (PDF)

This is an excerpt from the publication "German Development Cooperation in Fragile Contexts".
Download the full report here: http://www.deval.org/files/content/Dateien/Evaluierung/Berichte/2020/DEval_Report_3_2019_German_Development_Cooperation_in_Fragile_Contexts.pdf

The German Institute for Development Evaluation (DEval) is mandated by the German Federal Ministry for Economic Cooperation and Development (BMZ) to independently analyse and assess German development interventions.

Evaluation reports contribute to the transparency of development results and provide policymakers with evidence and lessons learned, based on which they can shape and improve their development policies.

This report can be downloaded as a PDF file from the DEval website:

www.deval.org/en/evaluation-reports.html

Requests for printed copies of this report should be sent to:

info@DEval.org

BMZ's response to this evaluation is available at <http://www.bmz.de/en/ministry/evaluation/Evaluation/evaluierungsberichte-stellungnahmen/index.html>

EXECUTIVE SUMMARY

Fragility – a central challenge for development cooperation

State fragility is an important challenge for development cooperation. The 2030 Agenda for Sustainable Development reflects the challenge of development cooperation in fragile contexts: SDG 16 sets out the vision to “[p]romote peaceful and inclusive societies [...], provide access to justice for all and build effective, accountable and inclusive institutions” (UN, 2015, p.14). Beyond the immediate added value of functioning regulatory institutions, challenges to the state’s monopoly on the legitimate use of force, a limited ability to make and implement binding rules and to deliver basic services throughout a state’s territory, as well as lack of belief in the rightfulness of a state’s authority among the population, put at risk the efforts of development cooperation.

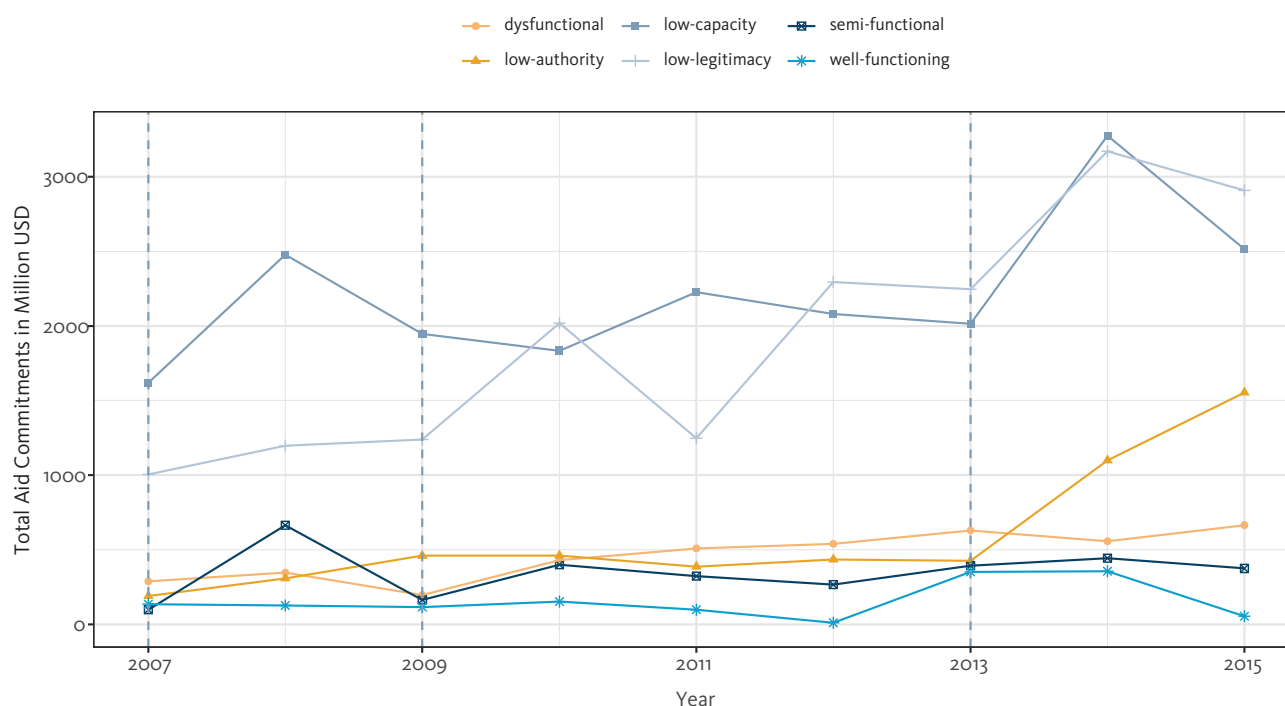
More than half of worldwide Official Development Assistance (ODA) goes to states affected by some or all of the mentioned aspects of fragility. German bilateral ODA commitments are no exception. Figure 1 shows committed grants and loans by type of fragility.¹ The largest share of German development cooperation goes to countries affected by some type of fragility – specifically low-capacity states, i.e. states with a low administrative capacity and limited supply of basic public services. It is for that reason that DEval is dedicating a Thematic Report on this topic which will compile the evidence of DEval evaluations and external contributions.

For development cooperation to be successful despite the challenging conditions in fragile contexts, the following four conditions should be satisfied:

1. A precise understanding of state fragility to align development policies to the local context
2. A political process that designs, possibly adapts, and implements strategies of development cooperation in fragile contexts effectively
3. Suitable procedures for learning what works in fragile contexts and what does not
4. Knowledge of the evidence with regard to which instruments of development cooperation work in fragile contexts.

The four self-contained chapters in this evaluation deal with the above-mentioned points. The following presents a brief summary.

¹ Grävingholt et al. (2018) distinguish six constellations of state fragility based on the state’s monopoly on the legitimate use of force (authority), the ability to provide basic services (capacity) and the recognition of state leadership (legitimacy). Dysfunctional, semi-functional and well-functioning states emerge from low, middle and high authority, capacity or legitimacy, respectively. States with low authority, capacity or legitimacy are characterized by low values in the respective dimension, while having average or high values in the other dimensions.

Figure 1 Bilateral ODA commitments (grants and loans) by type of fragility

Source: own figure based on data from OECD CRS and Grävingsholt et al. (2018).

Subject of this evaluation

This evaluation provides guidance for policymakers in the German Federal Ministry for Economic Cooperation and Development (BMZ) and other donors to adapt development cooperation to the context of partner countries with the aim to make it more effective. Each chapter examines one of the four above-mentioned conditions.

Following an introduction, Chapter 2 reviews the definition and measurement, i.e. the conceptualization of state fragility. Resulting indices and typologies constitute an integral part of decision-making processes in the BMZ and the implementing agencies. This evaluation provides a basis for assessing the suitability of available indices for decision-making. Chapter 2 critically compares approaches to defining and measuring state fragility. This is essential practical information since concepts influence how policymakers perceive current affairs. Resultant typologies classify partner countries, thereby influencing decisions about the suitability of certain development policy instruments. Consequently, the choice of concepts may influence aid allocation. With different approaches resulting in different classifications of states thus having tangible consequences, Chapter 2 speaks to a core issue of strategic steering by the BMZ.

The third chapter investigates policies by analysing strategies of development cooperation and allocations. Since 2005, the BMZ has released five major strategies guiding development cooperation in fragile states. This evaluation describes continuity and change in Germany's and the BMZ's development strategy towards fragile states and critically examines whether German ODA commitments are aligned to these strategies. The allocation analysis concisely points out which strategic decisions were actually put into practice and highlights where further action might be required.

Chapter 4 analyses the quality of evaluations and thus deals with one of the pivotal procedures for learning what works in fragile contexts. The implementing agencies, KfW Development Bank (KfW) and the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, have built up a large body of evidence on the success of development cooperation. This evidence spans all types of countries and context conditions – among them fragile settings in which a lack of security and poor infrastructure can offer

challenges in designing and conducting evaluations. Based on an earlier DEval evaluation on sustainability in German development cooperation, this evaluation analyses the quality of evaluations depending on the fragility of the context. The analysis identifies factors that impede evaluations in fragile contexts and offers ideas on how to increase the usefulness of evaluations.

The fifth chapter gives policymakers insights into which instruments of development cooperation are particularly suitable in fragile contexts. A synthesis of evaluation reports examines the correlation between context fragility and the evaluation's assessment of project success. The analysis differentiates dimensions of state fragility, sectors, and criteria to measure project success. This allows the identification, from the many possible influencing factors, of those that relate to successful development cooperation in fragile contexts.

Methods

The concept comparison and the strategy analysis are based on an evaluation of academic literature, official documents, publicly available socio-economic data, and interviews.

To delineate the structure and processes of the German bilateral development cooperation system, the evaluation team reviewed internal and public documents elaborating on the procedures for decision-making. The review placed particular emphasis on those steps in the process that relate to strategies on development cooperation in fragile contexts or fragility assessments. In order to tailor recommendations to the needs of the BMZ, GIZ, and KfW, they refer to the procedures as described in the BMZ's Joint Procedural Reform (GVR).

The analyses of evaluation quality and project success build on a synthesis of evaluation reports by the GIZ and KfW. This synthesis links characteristics of the evaluated project with data on the implementation context. To consider subnational variation of state fragility, this evaluation develops an approach to automatically retrieve geographical information from evaluation reports in order to locate projects. This data is then linked to conflict events to estimate the enforcement of the state's monopoly on the use of force at a subnational level. Moreover, the analysis of evaluation quality draws on evaluation guidelines from the GIZ and KfW and existing meta-evaluations.

The following sections summarize the four main chapters. Each section discusses the relevance, and presents the main results and recommendations.

Defining and measuring state fragility: Results and recommendations

Context-specific development cooperation requires an understanding of state fragility as a multidimensional phenomenon. Given an increased interest in the topic especially after 9/11, the number of approaches to measuring fragility has increased considerably. This resulted in more refined concepts of fragility that acknowledge its multidimensional nature. Rather than classifying states binarily or assessing fragility on a one-dimensional scale, more recent approaches measure fragility based on multiple constitutive attributes and derive typologies.

The increasing availability of sophisticated approaches to measuring state fragility presents policymakers with the challenge of assessing the strengths and weaknesses of different approaches. This is relevant, since the use of a particular concept has concrete consequences in terms of the:

- number of states classified as fragile
- ability to distinguish different types of fragility
- accuracy and timeliness of measurement, and
- number of years and countries on which information is available.

The conceptual comparison gives an overview and provides guidance. It compares four concepts: the OECD's States of Fragility (SFR), the Constellations of State Fragility (CSF) developed at the German Development Institute, the Crisis Early Warning (CEW) provided by the German Institute of Global and

Areas Studies (GIGA) and used by the BMZ,² and the fragility component of the Catalogue of Criteria (CoC) for Assessing Development Orientation overseen by the Governance division of the BMZ.

The comparison reveals significant variation between concept designs, even for similar definitions.

These conceptual differences have implications for allocation and programming decisions. The concepts differ, for example, in their empirical scope. The BMZs CEW defines fragility focusing very narrowly on violent conflict. By contrast, the OECD's SFR follows a very broad understanding which includes, e.g., environmental risks such as exposure to natural hazards. Narrow concepts leave out aspects of statehood that may influence the success of development cooperation. Broad concepts, on the other hand, conceive state fragility as an overarching term describing a broad range of political, economic, societal, and environmental deficiencies and thereby become imprecise and lose analytical value and policy relevance. Consequently, focusing on the core state functions of authority, capacity, and legitimacy appears most useful.

Since no single concept on its own fulfils all requirements for supporting strategic policy-making, a combination of existing approaches seems to be expedient.

Concepts with a large scope, suitable extension, and founded on qualitative and quantitative public data and an explicit concept design are readily available but usually do not provide timely data. The latter are, however, available through the existing reporting system currently used in the BMZ. As it would be inefficient to develop yet another approach to defining and measuring fragility, the challenge lies in combining existing approaches. This requires a standardized process of combining and weighting information through mixed- and multi-method approaches. For instance, a combination of different sources could increase the precision of measurement, give insights into uncertainties, and improve the imputation of missing data.

Fragility indices are well suited to the strategic management of German bilateral development cooperation as they enable comparisons between states on a single, ideally multidimensional scale.

By contrast, a lack of guidance on what information should be considered relevant, combined with time constraints, can lead decision-makers to over-rely on volatile subjective reasoning prone to psychological biases, increasing the danger of incorrect judgements. The mandatory use of indices as currently practised in the development of country strategies thus promises to improve decision-making. At the same time, portfolio management, programming, and implementing development cooperation also requires an idiographic, in-depth understanding. In particular, highly aggregated measures that reduce dozens of indicators to a binary classification entail the danger of oversimplification.

One approach to reduce the danger of oversimplification is to present results of indices more comprehensively.

More specifically, a visual representation of concepts that includes indicators and the applied weighting mechanism would convey the wealth of information provided by multidimensional concepts and make transparent the influence of individual indicators. Another way to tackle the risks of oversimplification is to make clear the uncertainty of measurement. For instance, a country classified as non-fragile might change to being classified as fragile one year later because it barely passed a certain threshold. Since the underlying measurement is uncertain, such a change in category might be entirely due to chance. Moreover, quantitative indicators can convey a certainty that seldom exists. One way to prevent policymakers from drawing wrong conclusions from data is to report distributions rather than pseudo-exact point estimates.

Beyond possible improvements in presentation, this evaluation also discusses three ways to improve the understanding of fragility.

First, taking into account how fragility spreads across borders would allow possible spillover effects to be assessed. Second, spatial disaggregation of measures of fragility promises to improve our understanding because the level of fragility often varies within states, specifically in developing ones. We present one approach to disaggregate state authority subnationally in Chapter 5. Third, applying

² The BMZ replaced the CEW with a new system called the Escalation Potential Measurement (ESKA – Eskalationspotentialmessung) at the end of 2018.

the notion of fragility and its constitutive dimensions to the regime, administration, and government of a state could further improve the added value of fragility indices in context-sensitive programming.

In summary, we make the following recommendation:

Recommendation 1

Based on a systematic conceptual comparison of four approaches to defining and measuring fragility, this evaluation recommends that the BMZ continue to use fragility indices for strategic portfolio management. Furthermore, the BMZ should continue to constantly evaluate and possibly improve its definitions and measurement of fragility on the basis of clearly defined criteria. This should include an assessment of the extent to which existing approaches could be combined.

As part of the current revision of approaches to measuring fragility, the BMZ should ensure that the results of indices are presented to political decision-makers more comprehensively in order to make underlying concept designs transparent and to do justice to the multidimensional nature of fragility. To achieve this, the BMZ should adopt a visual representation of concepts that includes indicators and the applied weighting mechanism. Second, the BMZ should include the uncertainty of classification and measurement in presentations of fragility scores.

Moreover, with a view to possible future revisions, the BMZ should further disaggregate and contextualize its measurement of state fragility in order to further increase the value of indices for the context-sensitive management of development cooperation. Further disaggregation could be achieved by reporting information for lower-level geographical units or more regularly and by differentiating fragility between different actors within partner countries. Contextualization could be achieved by taking account of possible spillover effects of fragility across borders.

Strategies and allocation: Results and recommendations

For development cooperation to be successful, it needs to be aligned to fragile contexts. Consequently, it is no surprise that the increasing attention given to fragile contexts has resulted in a growing number of strategies on context-sensitive implementation of development cooperation.

This evaluation examines to what extent strategy papers shape German ODA commitments. Strategies provide guiding norms and principles for operations, define policy goals, set action areas, propose appropriate instruments, and explicate assumptions underlying development cooperation in fragile contexts. Beyond influencing policies and politics, strategy papers signal to the public and partner countries the German government's preferences concerning fragility, state-building, and crisis prevention. However, some practitioners and researchers argue that official policies as embodied in strategies have no effect on practice. Furness (2014), for instance, argues that, for the EU, development cooperation in fragile states reveals a "policy-operations" gap. Practitioners sometimes follow a similar reasoning when they claim that strategies are especially hard to implement in fragile contexts. The volatility and complexity of fragile contexts clearly impedes the detailed planning of interventions.

This evaluation uses three steps to investigate whether allocation patterns reflect the content of strategies. First, we describe changes and continuities in the BMZ's strategy towards fragile states. Second, we extract testable implications from strategy documents that evolved after critical junctures. And third, we test empirically to what extent ODA commitments match these testable implications.

The analysis identifies important strategic shifts during two critical junctures: the Al-Qaeda attacks in 2001 and the "Arab Spring" of 2011–12. With each critical juncture, new discourses emerged on how development cooperation can be adapted to the challenges arising from fragility. The 2004 Interministerial Action Plan "Civilian crisis prevention, conflict transformation and peace consolidation" can be considered Germany's first joint strategy targeting fragile states. Subsequently, the BMZ released two major strategies to operationalize its 161 recommendations. This resulted in two approaches to development cooperation supporting crisis prevention in fragile states: peacebuilding and institution-building. In 2012, the Federal Government released the "Interdepartmental Guidelines for a Coherent Policy of the German Federal

Government in Fragile States”. It was released at a time of disappointing results from liberal state-building efforts in Afghanistan, Iraq, and the DRC. Moreover, the five peace- and state-building goals (PSG5) concluded in the 2011 “New Deal for Engagement in Fragile States” served as a reference. The BMZ operationalized the guidelines in the strategy “Development for Peace and Security”. In 2017, the Federal Government released the interministerial guidelines “Preventing Crises, Managing Conflicts, Building Peace”. Efforts to operationalize the guidelines in the BMZ are continuing, with particular attention paid to the 2030 Agenda for Sustainable Development.

This evaluation finds that two approaches have emerged in German development cooperation towards fragile states: peacebuilding and institution-building. Although we have seen major innovations within each of these approaches and, more recently, attempts to integrate them more closely, they have both maintained a considerable degree of specialization. This specialization must be considered against the background that the foci of peace- and institution-building are distinct but complementary. They are distinct because, on the one hand, peacebuilding is about more than building strong and inclusive institutions. For instance, it also includes disarmament, demobilization, and reintegration. On the other hand, institutions affect environmental, economic, and social development and thus go beyond peacebuilding. At the same time, peace- and institution-building are themselves complementary. Institutions form the structural conditions that shape the incentives for societal, economic, and political actors and thereby influence the decision to resort to violence or play by the rules as embodied in institutions. Consequently, peacebuilding processes often centre on the question of shaping institutions. But peace and conflict also affect institutions: institutional changes due to violent conflict can render obsolete an assessment of governance and development orientation.

In the light of the fact that strategies have been implemented well, the institutional specialization within the BMZ seems fit for purpose. At the same time, the scientifically demonstrated mutual influence of institutions and political conflict illustrates that development cooperation in fragile states must always take into account peace- and institution-building together. Consequently, a close collaboration among the responsible organizational units can ensure complementarity. The joint reflection of development orientation, the level of governance, and the CEW are a good step in this direction. The current revision of the CEW and the measurement of the level of governance provides an opportunity to further deepen the cooperation, develop a uniform understanding of fragility, and eliminate possible redundancies by possibly integrating existing approaches (see Recommendation 1). On the one hand, this may include more clearly defining and demarcating the objectives of both approaches; on the other, joint data collection on shared indicators could increase efficiency and thereby improve assessments.

We thus make the following recommendation:

Recommendation 2

Based on a systematic conceptual comparison of approaches to defining and measuring fragility, an analysis of strategy documents, a review of the structure and processes of the German bilateral development cooperation system, and interviews, this evaluation recommends that the BMZ maintain its practice of dealing separately with questions of peace and security, and of governance. At the same time, the BMZ should ensure coordination and complementarity between the responsible organizational units wherever possible. The BMZ should systematically compare concept designs across both divisions with the aim of clearly demarcating approaches in order to avoid duplication and redundancy. Where concept designs complement one another, a formal mechanism for sharing indicators and data collection efforts should be established.

The investigation of patterns of commitment shows that bilateral aid allocations broadly follow the strategies. After 2005, more aid was committed to conflict transformation and peacebuilding, to countries with a German military presence, and to post-conflict countries. For the 2007 and 2009 strategies, more aid was committed to the direct delivery of basic services in countries with low governance and deteriorating development orientation, to public institutions in countries with low governance and improving development orientation, and to the support of government at local level in countries with low governance and constant development orientation. After the release of the 2013 strategy, transitional aid commitments

to non-partner countries in conflict or transition rose, and conflict-prevention programmes in fragile states saw an increase in total and relative commitments. We also find that the more general strategy of exiting from budgetary support also applies to fragile states.

Although these results show that allocation patterns are mostly in line with the policy recommendations tested, we also find some strategies to be less well implemented. First, we do not find evidence of an increase in commitments to basket funds, an instrument promising greater donor coordination and harmonization as well as increased ownership and reduced administrative burden faced by partners. Second, we find weak evidence of decreasing support to non-state actors, contrary to strategic recommendations. Both findings touch upon the question of whether state or non-state actors are suitable partners for development cooperation in fragile contexts. The low level of commitments to non-state actors may influence the efficient use of resources and the political regime of the respective country.

In summary, one of the main challenges lies in the strategic orientation of German development cooperation with regard to the question of what channels of delivery – among them the partner governments and non-state actors – offer the greatest benefits of development cooperation in fragile contexts. The strategies under analysis in this chapter already go a long way in tailoring instruments to the local context by making recommendations that are conditional on the fragility, level of governance, and development orientation of partner countries. However, in some cases implementation is falling behind.

We thus make the following recommendation:

Recommendation 3

Based on an analysis of strategy documents and commitments, and a screening of the literature, this evaluation recommends that the BMZ ensure the implementation of its strategies specifically with regard to the support of non-state actors in fragile contexts. The BMZ should also develop a concrete action plan to assess and possibly adapt guiding principles on cooperation with non-state actors in fragile contexts based on available evidence.

Moreover, the BMZ should continue the good practice of constantly adapting strategies to the ever-improving understanding of successful development cooperation in fragile contexts. The existing practice of tailoring strategic recommendations to contextual conditions is likely to do justice to the increasing analytical precision of scientific research and the validity of scientific explanations that depend on particular background conditions.

Evaluation in fragile contexts: Results and recommendations

Evaluating development interventions helps us to understand the effects of development cooperation and learn lessons for future activities. Fragile contexts can influence the quality of evaluation due, for example, to security concerns, poor infrastructure, or lack of mutual trust. Since evaluation reports are crucial for understanding what works in fragile contexts, a negative impact of fragility on evaluation quality would impede better knowledge about aid effectiveness.

A quantitative analysis shows no correlation between fragility and evaluation quality, but finds an indirect relation through reduced mobility in low-capacity states. On the one hand, we therefore conclude that evaluators generally succeed in maintaining evaluation quality – operationalized through the quality criteria developed by Noltze et al. (2018c) – at the usual level, even under difficult circumstances. On the other hand, a more detailed analysis shows that low state capacity is associated with reduced mobility and thus indirectly with a lower quality of evaluation.

In a complementary qualitative analysis, experts identified reduced mobility and poor project M&E capacities as the main challenges. The experts mentioned travel restrictions and poor infrastructure as causes for the limited mobility. Project M&E is related to various factors such as resource questions – in terms of budget as well as availability and competence of staff – leading to lacking or poor-quality indicators and other project data available for external evaluators.

In summary, the mobility of evaluators and M&E at the project level are central concerns for increasing evaluation quality in fragile contexts. Consequently, a critical review of M&E systems could be useful to identify potential to improve data quality and availability. Approaches to improve M&E systems could include reserving sufficient resources for M&E tasks at project/programme level, incentives such as M&E-capacity-related annual objectives for leadership personnel, and sensitization training specifically for leadership personnel. Moreover, the evaluation capacity of local partners could be improved by creating a sustainable pool of evaluation experts in partner countries.

Additional approaches to increase M&E capacity are budgeting and project planning. Separate budgets for project implementation on the one hand and M&E on the other should limit competition for resources. This has already been recently implemented with regard to evaluations. For instance, the GIZ has a separate budget for financing evaluations. Under the recent GIZ evaluation reform, project evaluations will now be more centralized, being commissioned by the evaluation department of the GIZ headquarters. With regard to project monitoring, module proposals (*Modulvorschläge*) for technical development cooperation do not separately report monitoring costs. However, separate reporting of costs could improve the BMZ's steering capacity in respect of the quality of monitoring and evaluation.

Mobility challenges can be addressed through further strengthening the capacity of conducting remote monitoring and evaluation. Suitable approaches to data collection include remote sensing, digital data collection using hand-held devices, and remote surveys. To make use of these new types of digital data in decision-making processes, it is moreover important to strengthen the capacity to apply and process data.

We thus make the following recommendation:

Recommendation 4

Based on qualitative interviews, observational evidence from an evaluation synthesis, and a screening of the literature, this evaluation recommends that the BMZ, GIZ, and KfW address mobility restrictions on evaluators in fragile contexts and improve M&E at the project level.

First, we recommend that the BMZ require the GIZ and the KfW to further separate out budgets for project implementation on the one hand, and project monitoring and evaluation on the other, to further strengthen the prevention of fungibility and competition for resources.

Second, this evaluation recommends that the GIZ and the KfW address mobility challenges in fragile contexts through the increased use of innovative data-collection methods. This includes strengthening capacity in conducting remote evaluation and monitoring using, e.g., geospatial methods, remote surveys, and digital data collection.

Development interventions in fragile contexts: Results and recommendation

This evaluation sheds new light on the nexus between fragility and project success by synthesizing the results of 471 digitized evaluation reports on GIZ and KfW projects in 85 countries between 2006 and 2016. As a rather broad phenomenon, state fragility is linked to many factors that have been discussed in terms of the success of development cooperation.

Two innovative features of this evaluation promise to improve our understanding of the role of state fragility. First, a multidimensional yet state-based concept of fragility measures the authority, capacity, and legitimacy of partner countries. Second, this evaluation presents a subnational measure of fragility based on violent events in the vicinity of georeferenced locations. Since the level of fragility often varies within states, the use of national fragility scores might be insufficient. Consequently, we retrieve information on project locations from evaluation reports through natural language processing and assign coordinates to each extracted location. This assessment of challenges to the state's monopoly on the legitimate use of force provides a more realistic picture of the context in which development cooperation takes place.

The synthesis shows no evidence that fragility of context correlates with project success as measured by evaluation ratings. We find no evidence of project success differing among most types of national fragility, or of varying project success conditional on conflict events near project sites. The causal

mechanism behind this perhaps surprising result is not clear. It might be that projects are well adapted to fragility in terms of programming and design. The fact that this evaluation finds strategies on fragility to be well implemented in German development cooperation seems to point in this direction. It could also be that the goals of interventions in fragile contexts are set lower, which makes them easier to achieve. However, interventions in highly fragile contexts usually have a dual objective in that they additionally aim at stabilization. A second possible explanation is that any negative effect of fragility is offset by greater benefits of projects due to the possibly low initial level of development. It could also be that evaluators take into account the fragility of the context in their ratings, which would lead to spurious results.

A closer look at the dimensions of fragility shows that projects in countries with higher levels of capacity are, on average, more successful. The results indicate that the extent to which development interventions achieve their objectives inducing sustainable positive change depends on the capacity of partner countries. By contrast, whether interventions do so in cost- and time-efficient ways and whether they are well adapted to the target group seems to depend less on capacity. This finding supports the results of an earlier evaluation by Noltze et al. (2018b), which finds the role of partner countries to be particularly important in rating sustainability.

Correlations between state capacity and project ratings are particularly noticeable in the health sector. This resonates with existing findings that interventions aimed at improving health require efficient and independent bureaucracies as well as an efficient infrastructure of service provision across the whole country (Briebe, 2018). However, we should be cautious with these results as health outcomes are also included in the measurement of state capacity.

The relation between fragility and project success turns out to be context-, sector-, and outcome-specific. Tailoring development cooperation to fragile contexts requires a deeper understanding of the complex mechanisms that determine project success in different sectors across different contexts. Although further research and better data are needed, strengthening state capacity seems to be crucial in order to create a context for successful development cooperation. Consequently, it seems advisable to continue, and possibly to extend, the good practice of assessing state capacity in the management of bilateral development cooperation. Differentiating fragility profiles and a detailed Peace and Conflict Assessment are important steps in this direction. The capacity-related risk assessment could also draw on the successor to the Governance Criteria Catalogue as well as monitoring and evaluations of similar interventions in similar contexts.

Moreover, measures of state capacity could be further improved by disaggregating capacity among different actors and by taking into account subnational variations in state capacity (also see Recommendation 1).

Furthermore, a higher standardization of evaluations should improve the comparability of evaluation reports. From an evaluator's point of view, the validity of indicators with regard to project goals as well as comparability of indicators across different projects are of primary importance. Current efforts by the BMZ, GIZ, and KfW in the joint working group "AG Evaluierung" are aimed in this direction and hence should improve the quality of observational data analysis across different contexts.

We thus make the following recommendation:

Recommendation 5

Based on weak observational evidence from an evaluation synthesis, a screening of research and evaluations, and a review of the structure and processes of the German bilateral development cooperation system, this evaluation recommends that the BMZ systematically assess the risks emanating from a lack of capacity of partner countries for the success of development projects even more than has hitherto been the case. This should include a detailed assessment in module proposals of implementation risks in low-capacity states.

Moreover, the analysis points to the added value of improving the quality of, and standardizing, monitoring and evaluation. Consequently, the joint BMZ-GIZ-KfW working group “AG Evaluierung” should strive to increase comparability across evaluations with a view to enhancing the validity of evaluation syntheses and thereby facilitate joint learning.