Journal of MultiDisciplinary Evaluation
Volume 13, Issue 28, 2017

JMDE
Journal of MultiDisciplinary Evaluation

ISSN 1556-8180
http://www.jmde.com

A Mechanism-Centered Approach to Evaluating Complex Aid Interventions: The Case of Accompanying Measures to Budget Support

Johannes Schmitt
German Institute for Development Evaluation - DEval; University of Duisburg-Essen

Franziska Krisch
Deutsche Gesellschaft für Internationale Zusammenarbeit – GIZ

Background: Current methodological debates related to theory-based evaluations (TBE) focus on improving the explanatory strength of these approaches and how to integrate mechanisms as analytical concepts. Particularly in complex aid interventions, when multiple elements are expected to interact and thus create added value, exploring mechanisms as an analytical tool can be promising.

Purpose: This paper contributes to the discussion on the use of TBE for evaluating complex aid interventions by sharing experiences from a recent evaluation of accompanying measures to general budget support.

Setting: Nine countries of Sub-Saharan Africa which have received German budget support; Namely Burkina Faso, Ghana, Malawi, Mali, Mozambique, Rwanda, Tanzania, Uganda, and Zambia.

Intervention: Accompanying measures (mainly in the form of technical assistance and capacity development) are one element of the budget support package, which further encompasses financial contributions, policy dialogue, and conditionalities.

Research Design: We focus on interrelations between different elements of budget support and apply a mechanism-centered approach to program theory building.

Data Collection and Analysis: After defining accompanying measures and integrating them into the intervention logic of budget support used in recent multi-donor evaluations, key mechanisms were identified on an explorative mission to Mozambique, validated in an online survey, and further discussed in expert interviews and during field research in Tanzania.

Findings: For the specific example relating to two elements of budget support (policy dialogue and accompanying measures), some of the hypothesized mechanisms created added value and thus increased the effectiveness of budget support as a package. The applied approach helped generate a more comprehensive implementation theory and provided insights into potential benefits and challenges of combining different elements in one program. Beyond its use for future evaluations in the field of budget support, we argue that TBE of complex interventions can benefit from adopting such a mechanism-centered approach to create a better understanding of how different elements of the program interact. Moreover, the focus on mechanisms when analyzing program implementation enables evaluators to improve their empirical inquiry on the identified mechanisms and to draw valid conclusions on the program's contribution to the observed outcomes.

Keywords: *Theory-based evaluation; theory of change; development evaluation; budget support; technical assistance; capacity development.*

Introduction[1]

In light of development cooperation becoming more complex with the advent of harmonised and aligned aid modalities, approaches of theory-based evaluation (TBE) have gained momentum in recent years. Particularly in contexts of evaluating embedded and complex aid programs, where experimental designs based on counterfactual analyses are not feasible, evaluators have made increased use of TBE approaches. These approaches apply different forms of program theories in which the intervention logic and specific impact chains are outlined and which provide reference points to assess implementation and effectiveness of the interventions.

Current methodological debates related to TBE focus on improving the explanatory strength of these approaches and how to integrate mechanisms as analytic concept. Vividly discussed in TBE approaches such as Theory of Change or Realistic Evaluation, mechanisms are identified and analysed to assess how effective interventions contribute to observed outcomes. However, the potential of understanding mechanisms when evaluating the implementation side has yet to be realised. Particularly in complex aid interventions, when multiple elements are expected to interact and thus create an added value, exploring mechanisms as an analytical tool can be promising.

Described as one of the most complex aid modalities, general budget support combines multiple inputs by multiple donors in order to achieve results on multiple levels. Over the past decade, donors increasingly commissioned large-scale evaluations of budget support following a TBE approach thus stimulating the discussion among scholars and practitioners on how to improve the methodological approach for evaluating budget support. In this methodological debate, the role of causal mechanisms to substantiate the underlying program theory and to improve the empirical analysis of the complex aid instrument has been highlighted recently (ECA, 2014; Lawson, 2015; Schmitt and Beach, 2015).

This paper contributes to the discussion on the use of TBE for evaluating complex aid interventions by sharing experiences from a recent evaluation of accompanying measures to general budget support (Krisch et al., 2015).[2] Focusing on one element of budget support, this evaluation investigated if, how, and under which circumstances accompanying measures can contribute to achieving the objectives of general budget support. We focused on the implementation side of the program theory of budget support and specifically on interrelations between different elements of budget support.

Applying what we term a mechanism-centred approach to program theory building, we were able to identify benefits and challenges of combining different elements of budget support thus providing a more comprehensive implementation theory. We argue that TBE of complex interventions can benefit from adopting such a mechanism-centered approach to create a better understanding of how different elements of the program interact. Moreover, the focus on mechanisms when analysing program implementation enables evaluators to improve their empirical inquiry on the identified mechanisms and to draw valid conclusions on the program's contribution to the observed outcomes.

TBE, Program Theories and Mechanisms in Complex Evaluations

Today, theory-based [3] evaluations (TBE) are regularly used in a range of settings to create a better understanding of how and why a program works or fails (Fitzpatrick et al., 2012, p. 161). Different TBE approaches [4] share the idea of explaining how, why and for whom an intervention had an effect in a given context, thereby unpacking the 'programmatic black boxes' (Astbury & Leeuw, 2010). TBE all rely on some sort of theoretical account generally referred to as a program theory, defined as an 'explicit theory of how the

[1] The authors express their appreciation and sincere thanks to colleagues from the German Institute for Development Evaluation (DEval) who commented on an earlier version of this paper.

[2] The empirical results presented in this paper build on data collected during the evaluation conducted by the German Institute for Development Evaluation (DEval). The authors would like to thank all interview partners and participants of the online survey as well as all people involved in the evaluation.

[3] Among TBE scholars, different terms have been used like theory-driven, theory-based, or theory-oriented. As they generally describe similar concepts, they shall be used interchangeably. (see Fitzpatrick et al., 2012, p. 160).

[4] The most important TBE approaches include Realistic Evaluation (Pawson & Tilley, 1997) as well as several forms of Theory of Change Approaches such as Contribution Analysis (Mayne, 2012), General Elimination Methodology - GEM (Scriven, 2008), and Process Tracing (Beach & Pederson, 2013, George & Bennett, 2005).

intervention is understood to contribute to its intended and observed outcomes' (Funnel & Rogers, 2011, p. 34) thereby 'using the idea of 'mechanisms' that are at work' (Astbury & Leeuw, 2010, p. 365).

Especially in complex settings[5] which do not allow determining causation by comparison to a counterfactual, TBE represent an alternative for investigating how an intervention contributed to observed results (Astbury & Leeuw, 2010; Mayne, 2012, p. 4). Depending on the purpose of an evaluation, TBE can emphasise different levels of the program theory. 'Implementation theory' can be applied to better understand 'what is required to translate objectives into ongoing service delivery and program operation' (Weiss, 1997, p. 58). Focusing on the effectiveness of an intervention, 'program theory'[6] hypothesizes how mechanisms released by an intervention might trigger the anticipated outcomes. (Weiss, 1997, p. 46; see also Mayne, 2012).

Despite consensus regarding the use of program theories, there is a tendency among evaluators to break down the complexity of the evaluation topic for reasons of feasibility. In practical use of TBE in complex settings, model evaluations applying a thoughtful and complete program theory are rare (Weiss, 1997). Particularly in the context of complex aid interventions, evaluations tend to reduce complexity by using less explicit intervention logics instead of disentangling the program's mode of operation in an implementation theory. Such intervention logics describe the levels of inputs, outputs, outcomes and impact by listing all the components in each of the categories without necessarily describing the 'rational or reasoning behind program success.' (Fitzpatrick et al., 2012, p. 162). This results in a 'broad but weak' evaluation approach (Stern et al., 2012) in which the programmatic assumptions often remain implicit and consequently untested in the empirical inquiry (Astbury & Leeuw, 2010). In a similar vein, Nkwake (2012, p. 9) finds that the underlying assumptions of complex development programs are often taken for granted and points to the need to unearth assumptions in order to represent the complexity of the program. Focussing more on the role of causal mechanisms, Schmitt & Beach (2015) find that in TBE approaches such mechanisms are often hidden in the assumptions and therefore remain 'grey-boxed'.

The debate on how to use mechanisms in TBE has been ongoing and intensified in recent years. Mechanisms were introduced to the evaluation field by Chen (1990), who wrote about 'intervening mechanism evaluation' and referred to mechanisms as intervening variables (see also Chen, 2015, p. 312). Weiss (1997) shares this understanding of mechanisms and uses mediator and moderator variables as 'third variables that affect the relation between an independent and a dependent variable' (Weiss, 1997, p. 50).[7] With this understanding as intervening variable, mechanisms have become part of the evaluation lexicon today (Astbury & Leeuw, 2010, p. 366) and have long been used in TBE approaches, most notably in Realistic Evaluation (see Pawson & Tilley, 1997).

Although being widely discussed and applied in Theory of Change and Realistic Evaluation (see Blamey & Mackenzie, 2007), the use of mechanisms seems to be confined to the 'program theory' part of TBE approaches. To our knowledge, approaches or applications of mechanisms to evaluate 'implementation theory' do not exist. We argue that particularly in cases of complex settings of implementation, the use of mechanisms can be helpful to create a better understanding of how elements of a program interact and thereby create added value. Therefore, we diverge from the intervening understanding of mechanism and refer to them more as a system that transfers causal forces from cause to outcome.[8] In what can be termed a 'system' understanding of mechanisms, the analytical focus is on the theorized process itself. Opposed to the conception of mechanisms as intervening variables, (parts of) mechanisms in this understanding are not seen as separate entities that exist independently of the system. They can instead be directly linked to the program inputs and are integral parts of the causal process to produce the envisaged outcomes (Hedström & Ylikoski, 2010; Machamer et al., 2000).

[5] Based on Rogers' (2008) distinction between simple, complicated and complex, we define an intervention as complex if it follows a non-linear program logic and if its outcomes are likely to be disproportionate and emergent. Following this definition, budget support can be described as a complex aid instrument (Stern et al., 2012, p. 11; see also Betts, 2013, p. 255).

[6] To avoid confusion, we set the term 'program theory' in inverted commas wherever we refer to the part of the program theory that focuses on effectiveness.

[7] Moderator variables are characteristics that explain differences in outcomes. Mediator variables represent the generative mechanisms through which an intervention affects the outcome (Weiss, 1997, p. 50).

[8] For recent discussions of the mechanistic understanding, see Beach & Pedersen (2013); Gerring (2010).

The Case of Budget Support and Accompanying Measures

After the turn of the millennium, the international development community intensified the debate on how to increase the effectiveness of foreign aid and shifted portfolios from conventional project aid to new forms of so called program-based approaches (PBAs). Within the family of PBAs, budget support is conceived as the most far-reaching instrument to implement the new aid paradigm (Dijkstra et al., 2012; Faust et al., 2012).

In the case of budget support, multiple donors provide financial and non-financial inputs to support the partner government and to achieve various development outcomes (de Kemp et al., 2011). The core intention is to provide financial support to reform-oriented governments in implementing national strategies for development and poverty reduction. To that end, budget support donors provide non-earmarked financial means directly to the recipient country's national treasury. Besides this financing function, donors expect budget support to contribute to good governance in the recipient country, specifically with regard to public sector reforms, and to increase the efficiency of national public financial management systems. In order to achieve these governance objectives, donors complement their financial transfers with so-called non-financial elements of budget support comprising conditionality, policy dialogue between donors and the recipient government, as well as accompanying measures such as technical assistance and capacity development (TA/CD)[9].

The added value from simultaneously providing multiple budget support inputs has been described on a general level in policy documents and evaluation frameworks. Conditionality, policy dialogue, and accompanying measures are expected to directly strengthen political and administrative capacities in the recipient country and therefore increase the effectiveness of budget support funds. In turn, the financial contribution serves as a lever for the effectiveness of non-financial elements (de Kemp et al., 2011, p. 36ff; Nilsson, 2004). Particularly for the element of accompanying measures, donors envisage direct effects to overcome bottlenecks in the recipient countries' administrative systems. Admitting that the expected outcomes can hardly be realised by simply disbursing financial means to the partner's budget, donors emphasise the complementary nature of accompanying measures in the budget support package and describe expected synergies in their budget support strategies (BMZ, 2008; DANIDA, 2013; DFID, 2011; EC, 2012).

Despite their conceptual and operational importance, the added value of providing accompanying measures to budget support has not yet been analysed systematically. Since the start of the multi-donor aid instrument in the early 2000s, donors commissioned a series of evaluations of budget support. These evaluations followed a theory-based approach that builds on two components: i) an intervention logic of budget support represented in the Comprehensive Evaluation Framework (CEF) and ii) a set of analytical tools described in a '3-step approach' (OECD/DAC, 2012).[10] Due to its broad focus and openness to capture the extensive impact chain of the complex intervention, this methodological approach helped generate a substantial body of empirical evidence on the effectiveness of budget support. Yet, experience with the approach also shows that this strength is partly offset by limitations to produce in-depth information and reliable evidence on the contribution of budget support to the observed outcomes. The European Court of Auditors points to 'inherent limitations' referring to the 'difficulty in establishing unambiguous logical cause-effect links between results observed and the budget support programs' (ECA 2014, p. 20). The identified challenges have been concretised by experts of evaluating budget support who admit that causal mechanisms need to be taken into account more systematically in order to increase the explanatory account of future evaluations (EC & IEG 2015, p. 9; Lawson 2015). In a similar vein, Schmitt & Beach (2015) point out shortcomings related to the limited base for inferences about actual causal processes in existing evaluations and link these back to a lack of explicit theorisation of causal mechanisms in the existing intervention logic.

[9] Note that we use the terms accompanying measures and technical assistance and capacity development (TA/CD) interchangeably.

[10] The CEF represents the generic intervention logic of budget support. It structures the hypothesized sequence of expected effects of budget support programs across five analytical levels (inputs, direct outputs, induced outputs, outcomes and impacts). The CEF acknowledges the fact that budget support is embedded into the broader context of the recipient country's own development strategy and considers effects from various government programs and inputs by other aid activities. In addition to the CEF, the methodological approach provides guidance on how to structure the empirical inquiry in a 3-step approach (OECD/DAC 2012).

We argue that the challenges of present budget support evaluations to formulate credible attribution statements are mainly rooted in a weak implementation theory. As TBE generally assess interventions against their (existing or reconstructed) program theory, the focus of empirical analysis strongly relies on the quality of this program theory. For budget support, the applied intervention logic neither theorizes the causal mechanisms leading from individual inputs to results on the level of direct and induced outputs, nor explains how the combination of different elements of budget support is expected to create an added value. The theoretical account is particularly limited with respect to the role of accompanying measures and their interrelations with other elements of budget support.

The following section illustrates a mechanism-centered approach to program theory building. Drawing on a recent evaluation of accompanying measures to budget support, we will focus on the implementation part of the theory and show how program theory reconstruction can be used to identify and understand mechanisms of this complex intervention.

Applying a Mechanism-Centered Approach to Program Theory Building: The Example of Accompanying Measures to Budget Support

The following example draws on an evaluation of accompanying measures to general budget support in Sub-Saharan Africa. The evaluation asks if, how, and under which circumstances accompanying measures to budget support can contribute to achieving the objectives of general budget support. The evaluation aimed at gaining a better understanding of how accompanying measures function as an integral part of general budget support. In doing so, the team looked into the interdependencies between accompanying measures and other elements of the budget support package in order to identify the added value of strengthening partner countries' systems parallel to the financial contribution of budget support programs.

The evaluation adopted a theory-based approach building on the internationally-accepted intervention logic of budget support as developed under the lead of the EC and the OECD-DAC (OECD/DAC, 2012, p. 9). Against the backdrop of the identified challenges with the CEF described above, the evaluation team focused on the implementation side and applied a mechanism-centred approach to program theory building. Inspired by existing concepts of reconstructing program theories (Donaldson, 2007; Fitzpatrick et al. 2012), elements of a 'scientific' and a 'stakeholder approach' [11] (Chen, 1990) were combined across multiple analytical steps:

1. Defining accompanying measures to budget support

Based on explorative interviews with German and international stakeholders, we formulated a broad definition of the evaluation topic: 'Accompanying measures to budget support include all interventions of financial and technical cooperation which overlap in time with the allocation of general budget support and can contribute to the effectiveness of budget support due to interdependencies.' This conception expands the rather narrow definitions used in existing evaluations of budget support. While most evaluations only considered TA/CD as an element of budget support if it was directly linked to the budget support program,[12] in reality most donors implemented TA/CD parallel to their budget support and considered them as accompanying measures. Our comprehensive definition thus includes all forms of TA/CD that is provided simultaneously to budget support and pursues related objectives.

We categorised accompanying measures into the thematic areas of development policies, public financial management, and democratic control: The majority of accompanying measures to budget support aims at strengthening the public financial management system and addresses national institutions such as the auditor general's office or the tax authority. Furthermore, accompanying measures also aim at improving the formulation and implementation of the government's development policies and at enabling parliaments, civil society, and the media to better exercise democratic control in budget processes.

[11] According to Chen (1990), there are two ways of constructing a program theory which can also be combined: the stakeholder approach and the social science approach. While the first relies on the stakeholders' views, the latter builds on prior knowledge from social sciences and the evaluator's experience. Best practices usually combine both aspects (Fitzpatrick et al., 2012).

[12] That means on the same financing document and thus dependent on the budget support program.

Accompanying measures can take different forms. Addressing both government institutions and civil society, they encompass a range of interventions such as capacity development, short and long-term advisors, or software packages[13]. Multiple donors frequently provide them in the form of common basket funds, yet also implement them as bilateral program aid.

2. *Integrating accompanying measures into the intervention logic of budget support*

Based on the narrative in program documents of German accompanying measures and on explorative interviews with experienced staff from the German Federal Ministry for Economic Cooperation and Development (BMZ), German implementing agencies (KfW Development Bank and Deutsche Gesellschaft für Internationale Zusammenarbeit - GIZ) as well as international experts of evaluating budget support, the evaluation team developed a generic intervention logic for accompanying measures to budget support which helped to illustrate complexity and to map underlying impact assumptions. Further elaborating the accepted intervention logic of budget support (OECD/DAC, 2012), the team developed four hypotheses on the interrelations between different budget support inputs and identified a total of seven mechanisms. In this paper, we exemplify the four identified mechanisms for two hypotheses on interrelations between accompanying measures and policy dialogue. The first hypothesis states that accompanying measures are expected to improve the quality of policy dialogue. In return, the second hypothesis postulates that policy dialogue increases the effectiveness of accompanying measures.

3. *Identifying value-adding mechanisms of accompanying measures to budget support*

To open the so-called programmatic black box (Astbury & Leeuw, 2010: 364ff.) of how exactly accompanying measures contribute to the objectives of budget support, we formulated mechanisms for these hypotheses in a participative approach with stakeholders. Within the scope of an explorative mission to Mozambique, we asked representatives of both donors and partners to describe important mechanisms that explain the operation mode of different hypotheses. We identified four mechanisms for the two hypotheses on interrelations between accompanying measures and policy dialogue.

Hypothesis 1: Accompanying measures improve the quality of policy dialogue[14].

Identified Mechanism 1: Accompanying measures provide information about the quality of public services and the need for support of government institutions. In turn, this information is used in policy dialogue.

Identified Mechanism 2: By implementing accompanying measures, donor representatives expand their knowledge and skills and can act as qualified partners in policy dialogue.

Hypothesis 2: Policy dialogue increases the effectiveness of accompanying measures.

Identified Mechanism 3: Accompanying measures are planned and implemented on the basis of a needs assessment to eliminate the weaknesses of the system in a targeted way.

Identified Mechanism 4: Planning and implementation of accompanying measures is coordinated between donors.

Thus, the ideal operation mode of these hypotheses and mechanisms can be represented in a loop of mutual information, feedback and linkages between accompanying measures and policy dialogue (see figure 1).

[13] E.g. Integrated Financial Management Information Systems (IFMIS).

[14] Policy dialogue of budget support takes place at different levels: (i) technical policy dialogue about reform progress takes place on the operational level between donor representatives and respective partner representatives in ministries and other government authorities. ii) in high-level policy dialogue, ambassadors of donors and high-level representatives of the partner government discuss general strategic matters.

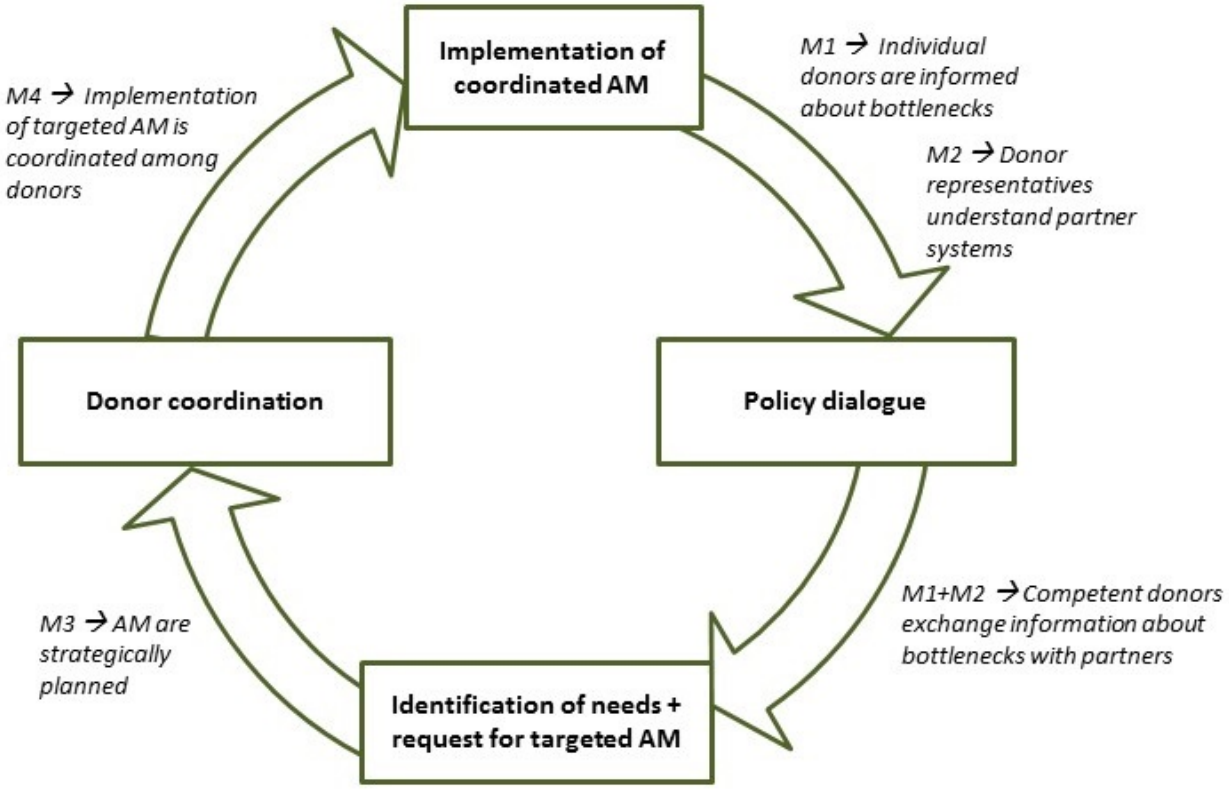

Figure 1. Interrelations between accompanying measures (AM) and policy dialogue: theory for the case of budget support (BS) in Mozambique.

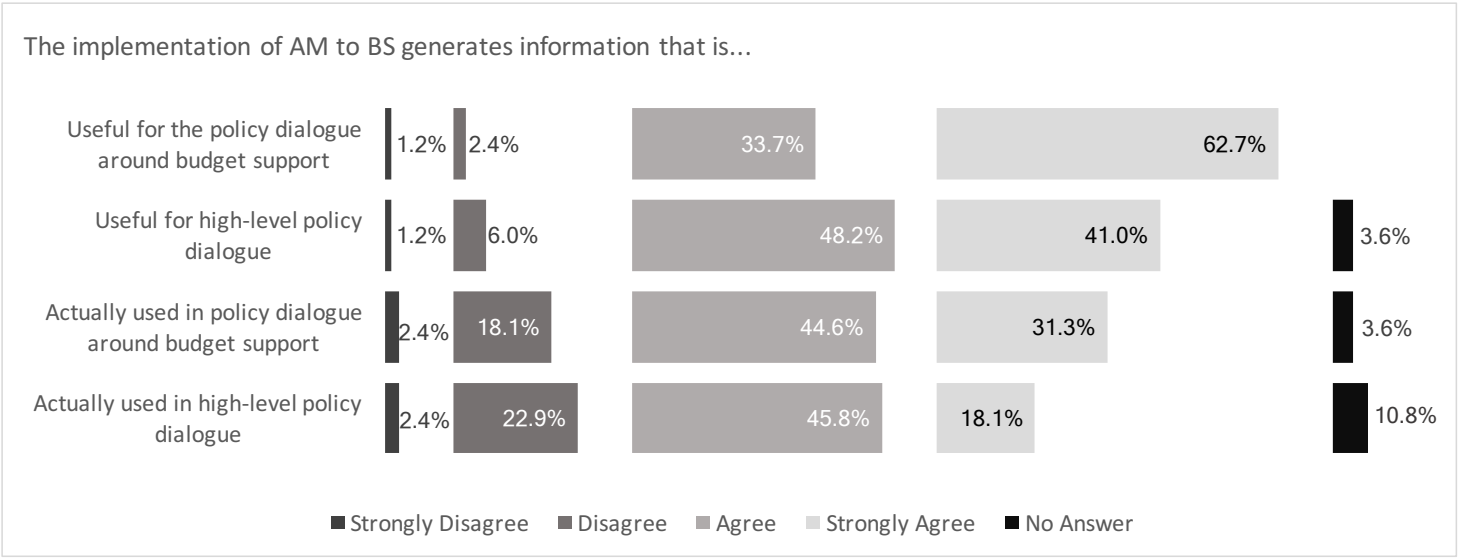

Figure 2: Results of online survey relating to M1 and M2

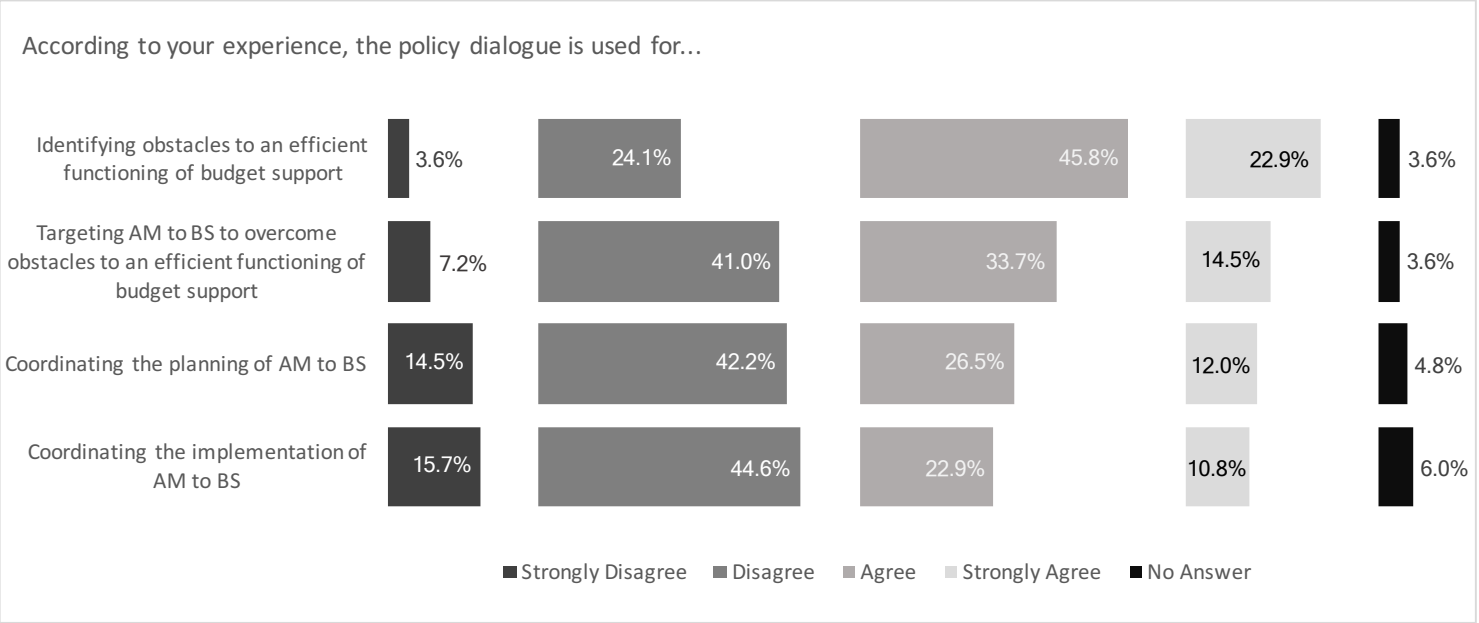

Figure 3. Results of online survey relating to M3 and M4.

Figure 1 represents the results of consultations with stakeholders in Mozambique and shows how they expect accompanying measures and policy dialogue to interact based on their experience in the specific case.

4. *Increasing validity for the hypotheses and mechanisms*

In a next step, the evaluation team proceeded to increase the validity of these hypotheses and mechanisms by conducting a standardised online survey[15]. We sent an invitation to participate in the online survey to representatives of donor countries implementing budget support in nine recipient countries of German budget support in sub-Saharan Africa; Namely Burkina Faso, Ghana, Malawi, Mali, Mozambique, Rwanda, Tanzania, Uganda, and Zambia. 83 out of 143 invited donor representatives (100 international heads of cooperation, 43 Germans including representatives of the implementing agencies GIZ and KfW) completed the questionnaire, an overall response rate of 58 percent (47 percent among international donors). The respondents rated statements describing the functioning of mechanism 1 on a 4-point scale.

The online survey results increase the confidence with respect to the first hypothesis that accompanying measures improve the quality of policy dialogue as well as the related mechanism 1. However, the agreement is strongest for the information function of implementing accompanying measures, especially in the technical policy dialogue surrounding budget support described in mechanism 1, and less pronounced for the actual utilisation of this information, particularly in the high-level political dialogue between ambassadors of the donor countries with representatives of the partner government (see figure 2).

Concerning hypothesis 2 that policy dialogue increases the effectiveness of accompanying measures and the related mechanisms 3 and 4, the results were mixed (see figure 3). 69 percent of the respondents agreed that the policy dialogue was used to identify obstacles to an efficient functioning of budget support, and 48 percent stated that accompanying measures are targeted to overcome these obstacles (mechanism 3). 61 percent of the respondents however disagreed that the policy dialogue was used to plan and implement such targeted accompanying measures in a coordinated manner (mechanism 4).

Overall, the survey results indicate that donor representatives hardly use policy dialogue to coordinate accompanying measures, neither to better address the identified weaknesses, nor to increase the efficiency in the implementation. The

[15] The survey was conducted from 8 to 30 September 2014 in English.

dotted-line arrows in the outer loop in figure 4 illustrate these shortcomings in fulfilling their planned commitment to coordination as associated with budget support.

5. *Deepening the understanding about key mechanisms*

The quantitative results from the online survey helped to move towards a more general implementation theory and strengthen the validity

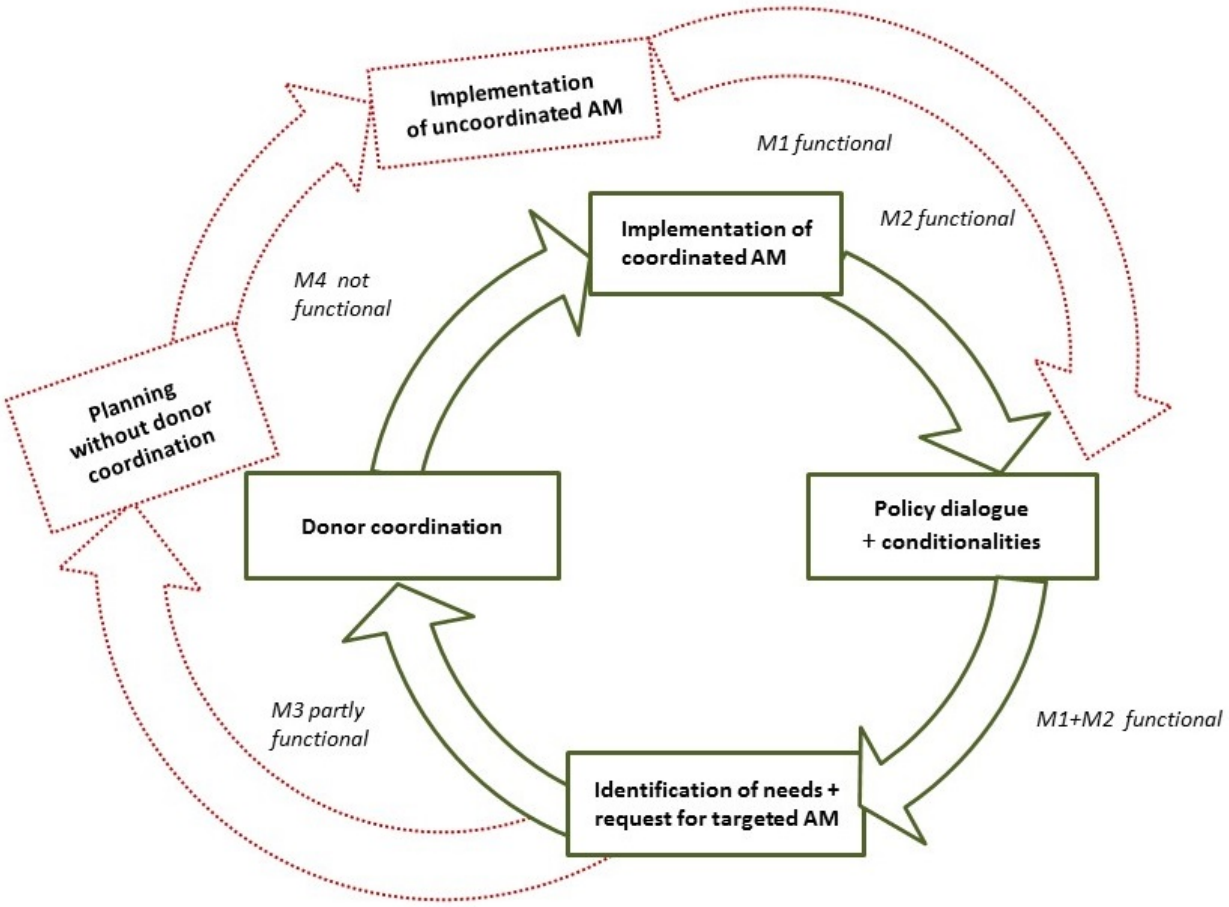

Figure 4. Interrelations between accompanying measures (AM) and policy dialogue: observed deviations from theory in Mozambique (own graph).

for M1 and M2 but posed new questions as to how and why problems of coordination occur. The results from the online survey were therefore further discussed during subsequent qualitative research to deepen the understanding about key mechanisms. In follow-up interviews, e.g. with technical experts in the field of public financial management as well as with representatives of government, public administration, civil society and media during a second field mission to Tanzania, the team investigated how exactly these mechanisms work or why they fail.

A common success factor for all four mechanisms was that they strongly depended on the quality of the technical policy dialogue and thus function mainly in the area of public financial management, in which the interests of partners and donors were found to align relatively well. When it comes to explaining the limited functioning of mechanism 4, the identified coordination problems can be traced to a range of shortcomings relating to ownership and harmonization, central principles for effective aid as stipulated in the Paris Declaration (High Level Forum, 2005).

When the partner government has assumed ownership for supporting a reform strategy by requesting accompanying measures from donors, and when the partner institution coordinates the implementation of accompanying measures, they

are used most intensively by the respective target groups. In some countries, the coordination of accompanying measures is exercised by the finance ministry, as well as by specific public institutions. The National Audit Office of Tanzania, for example, submitted a ten-year-plan for institution building which mentioned the need to improve the auditing infrastructure and the qualification of employees, and specified options for support from donors. Representatives of the National Audit Office reported that annual plans were drawn up based on this long-term strategy. They were then jointly financed from the National Audit Office's own budget, out of the basket fund for public financial management in the context of budget support, and by individual donors. If a specific donor expressed interest to cooperate, the pending needs of the National Audit Office and the thematic expertise of the respective donor were discussed and matched.

When partners do not assume ownership for reform implementation, and for coordination of accompanying measures, mechanism 4 functions only to a limited extent. In case of limited or no ownership, donors can still achieve a certain coordination of accompanying measures by harmonizing their TA/CD initiatives. However, donor representatives frequently even lack an overview of who is implementing which accompanying measures in which thematic field and institution. As a result, after receiving information about bottlenecks in partner systems in the policy dialogue, individual donors tend to design and implement their independent accompanying measures. Reasons for deviations from the ideal cycle of coordinated accompanying measures and good practice to overcome such obstacles emerged in follow-up interviews with different stakeholders. For example, as a first step, it is helpful to document the portfolio of accompanying measures of all donors. In Mozambique, the European Delegation took the initiative and set up an overview of all donor activities in the area of public financial management. In Zambia, donors signed a code of conduct on capacity development in 2008 to increase the effectiveness of technical assistance. Despite this commitment, the comprehensive PEMFA program – a common fund to support public financial management reforms – has been described as primarily donor-driven, and the overall effectiveness of donor support to public financial management systems was limited due to uncoordinated TA/CD side-lining the program (de Kemp et al., 2011, p. 69f).

Furthermore, interview partners explained the identified coordination problems with the political economy on the donor side. In the context of budget support, donors strive to convince their own parliaments that they take fiduciary risks seriously. Due to differing priorities of their headquarters, donor representatives in the partner countries cannot coordinate their accompanying measures effectively. Unless the partner government assumes strong ownership in the coordination of TA/CD initiatives, donors tend to follow their own preferences.[16]

Discussion

The objective of this paper is to contribute to the discussion on the use of TBE in the context of complex aid interventions. Focusing on the aid instrument of budget support and drawing on a recent evaluation of accompanying measures to general budget support, we present a mechanism-centred approach to program theory building focusing on the implementation side.

By applying this approach, we disclosed and assessed interrelations between different elements of budget support in retrospect. For the specific example relating to two elements of budget support (policy dialogue and accompanying measures), we reconstructed and further refined the implementation theory based on the identified mechanisms. While some of the mechanisms were found to create added value and thus increase the effectiveness of budget support as a package, others were not confirmed to be functional.

Beyond the relevance for the example of accompanying measures, we argue that the methodological approach of mechanism-centered program theory building can be useful to improve evaluations of budget support and related aid instruments and to strengthen TBE approaches in general. The applied approach provided insights into potential benefits and challenges of combining different elements in one program by generating a more comprehensive implementation theory.

In a narrow sense, the mechanism-centered approach to program theory building can be supportive to strengthening evaluations of budget support. Future evaluations could benefit from explicitly theorising the interrelations between different elements of a complex program through mechanisms. With a mechanism-centred reconstruction and the resulting substantiated

[16] The problem of uncoordinated technical assistance is not specific to budget support related accompanying measures but applies to TA/CD in the wider sense (Keijzer, 2013).

implementation theory, evaluations can generate a better understanding of how the program works. Moreover, knowing more about how the program generates the envisaged induced outputs, future evaluations can focus their empirical inquiry on the mechanisms and therefore come up with stronger conclusions on the contribution of budget support to the observed outcomes. Building on the identified mechanisms, the approach can also be useful to evaluate related programs that combine financial components with TA/CD and policy dialogue. The most relevant aid instruments in this aspect are sector budget support and sector-wide approaches in general, budget support in fragile states, and budget support for climate financing, as well as less related policy interventions such as aid for trade, or reconstruction after crises and conflicts.

Beyond its use for future evaluations in the field of budget support and related aid instruments, the presented approach provides the opportunity to strengthen TBE approaches in general. Particularly in settings of complex aid interventions or country program evaluations, the proposed approach might be useful to avoid applying simplistic intervention logics that are not able to illustrate complex program theories. We argue that the comparative advantages of complex programs can be evaluated only by making interdependencies between different elements of the program explicit and by uncovering the value-adding mechanisms at implementation level. Therefore, a 'system' understanding of mechanisms can help to illustrate complexity instead of treating them as assumptions which remain unquestioned. Based on a better understanding of interrelations between different elements of complex interventions, TBE can develop stronger conclusions on how to (re)design a complex development intervention. Moreover, knowing more about how a program really worked on the input and output levels enables the evaluator to draw more valid conclusions on the program's effectiveness to contribute to observed changes on outcome and impact level.

References

Astbury, B. & Leeuw, F.L. (2010). Unpacking black boxes: Mechanisms and theory building in evaluation. *American Journal of Evaluation* 31(3), 363 - 381.

Beach, D. & Pedersen, R.B. (2013). *Process-tracing methods: Foundations and guidelines,* Ann Arbor, Mich.: Univ. of Michigan Press.

Betts, J. (2013). Aid effectiveness and governance reforms: Applying realist principles to a complex synthesis across varied cases. *Evaluation* 19(3), 249–68.

Blamey, A. & MacKenzie, M. (2007). Theories of Change and Realistic Evaluation, Peas in a Pod or Apples and Oranges? *Evaluation* 13(4), 439-455.

BMZ. (2008). *Konzept zur Budgetfinanzierung im Rahmen der Programmorientierten Gemeinschaftsfinanzierung (PGF)*, Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, Bonn.

Chen, H-t. (1990). *Theory-driven evaluations,* Newbury Park, Calif.: Sage Publications.

Chen, H-t. (2015). *Practical program evaluation: Theory-driven evaluation and the integrated evaluation perspective,* Thousand Oaks, CA; London; New Delhi; Singapore: SAGE Publication.

DANIDA. (2013). *Guidelines for development contracts*. http://amg.um.dk/en/~/media/amg/Documents/Technical%20Guidelines/Guidelines%20for%20Development%20Contracts/Guidelines%20for%20development%20contracts%20final%20June%202013.pdf

de Kemp, A., Faust, J. & Leiderer, S. (2011). Between high expectations and reality: An evaluation of budget support in Zambia - Synthesis Report. Bonn; The Hague; Stockholm: BMZ; Ministry of Foreign Affairs; SIDA.

DFID. (2011). *Implementing DFID´s strengthened approach to budget support: Technical note*. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/214161/Strengthened-approach-budget-support-Tecnical-Note.pdf

Dijkstra, G., de Kemp, A. & Bergkamp, D. (2012). *Budget support: Conditional results - review of an aid instrument (200-2011).* IOB Evaluation. The Hague: Ministry of Foreign Affairs; Policy and Operations Evaluation Department (IOB).

Donaldson, S.I. (2007). *Program theory-driven evaluation science: Strategies and applications*. Mahwah, NJ: Erlbaum.

EC & IEG. (2015). *Joint evaluation of budget support to Uganda (2004-2013): Final report*. May 2015. http://ieg.worldbank.org/evaluations/joint-evaluation-budget-support-uganda-2004-2013

EC. (2012). *Budget support guidelines executive guide: A modern approach to budget support*. European Commission. http://ec.europa.eu/europeaid/sites/devco/files/methodology-budget-supportguidelines-executive-guide-201209_en_2.pdf

EU. (2014). *Mapping of support to PFM reforms in Mozambique*. Unpublished document.

European Court of Auditors. (2014). *EuropeAid's evaluation and results-oriented monitoring systems*. Special Report No. 18. European Union.

Faust, J., Leiderer, S. & Schmitt, J. (2012). Financing poverty alleviation vs. promoting democracy? Multi-donor budget support in Zambia, *Democratization*, 19, 438-464.

Fitzpatrick, J.L., Sanders, J.R. & Worthen, B.R. (2012). *Program evaluation: Alternative approaches and practical guidelines,* Upper Saddle River, N.J.: Pearson Education.

Funnell, S.C. & Rogers, P.J. (2011). *Purposeful program theory: Effective use of theories of change and logic models,* San Francisco, CA: Jossey-Bass.

George, A.L. & Bennett, A. (2005). *Case studies and theory development in the social sciences*. Cambridge, MS: MIT Press.

Gerring, J. (2010). Causal mechanisms: Yes, but…, *Comparative Political Studies*, 43, 1499-1526.

Hedström, P. & Ylikoski, P. (2010). Causal mechanisms in the social sciences. *Annual Review of Sociology,* 36, 49–67.

High Level Forum. (2005). *Paris declaration on aid effectiveness: Ownership, harmonisation, alignment, results and mutual accountability*.

Keijzer, N. (2013). *Who's the boss? Strengthening the effectiveness of capacity-development support*. Briefing Paper. Bonn: Deutsches Institut für Entwicklungspolitik (DIE).

Krisch, F., Schmitt, J. & Dörr, U. (2015). *Accompanying measures to general budget support in Sub-Saharan Africa*. German Institute for Development Evaluation (DEval). Bonn, https://www.deval.org/en/evaluation-reports.html

Lawson, A. (2015). *Synthesis of budget support evaluations (Vol.2): Review of methodological issues emerging from 7 country evaluations of budget support*.

Leeuw, F.L. (2003). Reconstructing program theories: Methods available and problems to be solved. *American Journal of Evaluation* 24, 5-20.

Machamer, P., Darden, L. & Craver, C.F. (2000). Thinking about mechanisms, *Philosophy of Science*, 67, 1–25.

Mayne, J. (2012). Contribution analysis: Coming of age? *Evaluation* 18, 270-280.

Nilsson, M. (2004). Effects of budget support: A discussion of early evidence. *UTV Working Paper* Stockholm: SIDA.

Nkwake, A.M. (2012). *Working with assumptions in international development program evaluation,* New York, NY: Springer.

OECD/DAC. (2012), *Evaluating budget support: Methodological approach*.

Pawson, R. & Tilley, N. (1997). *Realistic evaluation*. London; Thousand Oaks, Calif.; New Delhi; Singapore: Sage.

Rogers, P. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation* 14(1), 29-48.

Schmitt, J. & Beach, D. (2015). The contribution of process tracing to theory-based evaluations of complex aid instruments, *Evaluation*, 21(4), 429-447.

Scriven, M. (2008). A summative evaluation of RCT methodology: and an alternative approach to causal research. *Journal of Multidisciplinary Evaluation*, 5(9), 11–24.

Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., & Befani, B. (2012). *Broadening the range of designs and methods for impact evaluations*. Report of a study commissioned by the Department for International Development Working Paper (Vol. 38): DFID.

Weiss, C.H. (1997). Theory-based evaluation: Past, present, and future. *New Directions for Evaluation* 1997, 41-55.