



Effective Joint Sector Reviews as (Mutual) Accountability Platforms



Abstract

This paper examines the ways in which joint sector reviews (JSRs) contribute to the development of mutual accountability in an aid environment and may be most effective as instruments of a domestic policy cycle. Using a wide variety of sources, including an extended literature review, interviews with key education stakeholders in five countries and data collected from a desk-based analysis of 39 JSRs, we address two questions: (1) How are JSRs perceived as promoting (mutual) accountability? and (2) How is mutual accountability likely to be affected by JSR effectiveness? Our research suggests a less unified understanding of mutual accountability across key stakeholders than is touted in international development rhetoric. We therefore offer a more nuanced model of mutual accountability and the ways in which this intersects with JSRs that encompasses an expanded understanding of responsibilities and channels of responsibility. To achieve this, we identified five interconnected dimensions of effective JSRs that served as framework for analyzing 39 JSRs. We noted key intersections between these dimensions and the extent to which the JSR mechanism, when effective, is able to nurture mutual accountability. We also observed a number of weaknesses in how JSRs are currently operationalized and point to areas where improvement could unleash the potential of JSRs to lock together policy planning and implementation through policy dialogue. To the authors' knowledge, this is the first comprehensive and systematic review of JSRs in education to date.

This paper was written by Raphaëlle Martínez Lattanzio, Margaret Irving and Vania Salgado.

For more information, please contact Raphaëlle Martínez Lattanzio at rmartinez3@globalpartnership.org.

All rights reserved.

This paper is a product of the staff of the Secretariat of the Global Partnership for Education. The designations employed and the presentation of material throughout this publication do not imply the expression of any opinion on the part of the Global Partnership for Education, or the World Bank Group concerning the legal status of any country, territory, city, area or its authorities, frontiers or boundaries.

Published by

Global Partnership for Education
1850 K Street, NW, Suite 625, Washington, DC, 20006 USA
information@globalpartnership.org
www.globalpartnership.org

Acknowledgments

The authors would like to thank Karen Mundy, Chief Technical Officer at the Global Partnership for Education (GPE), for her ongoing support throughout this research undertaking. Her encouragement and intellectual guidance were central drivers of the success of this project. Particular thanks also goes to Hugues Moussy from the UNESCO International Institute for Educational Planning and William Smith from the Global Education Monitoring Report for their thoughtful comments and review of this paper. The coordination of the many moving parts of this project would not have been possible without the able assistance of Marc-Antoine Percier, who also provided key research contributions. Tongai Makoni and Sai Sudha Kanikicharla also provided invaluable research assistance. Much of the data collected for our analysis reflects the kind cooperation of a large number of interview respondents: Peter Harrold, Veronica Walford and Abby Riddell, who contributed useful insights into the historical emergence of joint sector reviews (JSRs) in the development landscape, as well as the 32 anonymized respondents from Bangladesh, Burkina Faso, Guinea, The Gambia and Zambia, who informed our review of stakeholder perspectives. We, of course, acknowledge that any errors or omissions are our own and should not reflect on those mentioned above.

The quality of the final product was greatly enhanced by guidance provided by Stuart Robinson and Noelle Roth on effective use of the NVivo research software as well as editing from Renata Harper. We are also grateful for the contributions of a number of GPE colleagues and consultants: Shilpa Bista, Adrien Boucher, Naoko Hosaka, Suha Kaouk, Karen Nzakimuena, Janne Perrier, Alexandra Solano, Krystyna Sonnenberg, Jeannette Vogelaar and Arianne Wessal, whose inputs included but were not limited to assistance with key logistics; conducting interviews and recording observations in the field; data management; collecting, quality assuring and analyzing survey and questionnaire data; and providing useful insights on methodologies, theoretical framing and interpretation of findings at various stages of this project. Finally, we wish to thank GPE country leads and education specialists for their patience in assessing and completing JSR questionnaires and for the rich perspectives into JSRs they offered during focus groups.

Table of Contents

Abstract.....	ii
Acronyms.....	vi
Introduction	1
SECTION 1: Research Strategy.....	4
SECTION 2: What Do We Know about JSRs? A Review of the Literature	6
JSRs in Historical Perspective: How Did the Mechanism Emerge?	7
Defining JSRs Today	9
Mutual Accountability as a Construct and Its Role in JSRs	13
SECTION 3: JSRs as a Mutual Accountability Platform	14
Education Stakeholders' View of Accountability in the Context of JSRs.....	14
Perspective of Development Partners	15
Perspective of Government Officials.....	17
Perspective of CSOs.....	18
SECTION 4: How Mutual Accountability Is Likely to Be Affected by JSR Effectiveness: An Empirical Assessment.....	20
Dimension 1: Inclusive and Participatory	20
Definition	20
Analysis	21
Dimension 2: Aligned to Shared Policy Frameworks	23
Definition	23
Analysis.....	23
Dimension 3: Based on Evidence.....	24
Definition	24
Analysis.....	24
Dimension 4: Monitoring Tool	26
Definition	26
Analysis.....	27

Dimension 5: Instrument for Change Embedded Effectively into a Policy Cycle	28
Definition	28
Analysis.....	29
Discussion and Conclusion	30
Bibliography.....	33
ANNEX 1: Methodological Annex.....	38
JSR Variable Analysis	38
Literature Review.....	39
Interviews with Key Country Stakeholders	45
Survey on JSR Function Preference	45
Additional Data	46
ANNEX 2: JSR Questionnaire Used for Variable Analysis	48

List of Figures

1 Dimension 1: Participatory and inclusive.....	21
2 Dimension 2: Aligned to shared policy frameworks	23
3 Dimension 3: Based on evidence	25
4 Dimension 4: Monitoring tool	27
5 Dimension 5: Instrument for change embedded effectively into a policy cycle.....	29
7 Literature type	42
8 Methodology	43
9 Instrument used	43
10 Distribution of articles by language	44

List of Tables

1 Timeline with key literature portraying JSR evolution.....	8
2 Stakeholder expectations on JSRs as a platform for accountability by stakeholder (% of N)	15
3 JSRs included in sample.....	40
4 Question-level Coding	41
5 Theme and sector identified in articles.....	41
6 Country cases most frequently covered in the articles	44
7 Interview list	46
8 JSR Function Preference	47

Acronyms

Busan Partnership	Busan Partnership for Effective Development Cooperation
CSO	Civil Society Organization
CL	country lead
EFA	Education for All
FTI	Education for All Fast Track Initiative
EMIS	Education Management Information System
ESP	education sector plan
GPE	Global Partnership for Education
IHP	International Health Partnership
IMF	International Monetary Fund
JSR	joint sector review
M&E	monitoring and evaluation
OECD	Organisation for Economic Co-operation and Development
Paris Declaration	Paris Declaration on Aid Effectiveness
SDG	Sustainable Development Goal
SIP	Sector Investment Program
SMART	specific, measurable, achievable, relevant and timely
SWAp	sector-wide approach
TEP	transitional education plan
Toc	Theory of Change
UNDP	United Nations Development Programme

Introduction

Conducting joint sector reviews (JSRs) is now a widespread development practice, with many aid recipient countries regularly organizing annual reviews in different sectors. However, little is known about the origin of this mechanism, the purposes it serves, the ways it contributes to sector results and how it reinforces aid effectiveness principles.

More broadly, fundamental shifts in the international aid paradigm over the past two decades have led to a re-envisioning of the relationship between key stakeholders in development. The aid effectiveness agenda epitomized in the Paris Declaration on Aid Effectiveness (2005) and the Accra Agenda for Action (2008), underpinned by the core principles of country ownership, donor alignment and harmonization, results management AP, and improved mutual accountability, has driven the formulation of new approaches to building more meaningful partnerships in the context of development. Among these approaches, JSRs have emerged as a central forum for making progress toward key sector goals and developing more robust country-level systems, particularly within the health and education sectors.

There is no agreed-upon definition or standardized set of practices that constitutes a JSR. The most comprehensive characterization is arguably that provided by Holvoet and Inberg (2009, 205): a “type of joint periodic assessment of performance in a specific sector with the aim of satisfying donors’ and recipients’ accountability and learning needs,” understanding performance in a broad way that includes inputs, activities, outputs, outcomes and impacts as well as systemic and institutional issues. One government official interviewed for this paper reinforced this understanding, noting:

The joint sector review is a key moment to assess the implementation of our policies and programs, to draw lessons and to anticipate the next steps for the following year. [...] it is also an opportunity for having substantive exchange with our partners, either social partners or technical and financial partners. It is also an essential moment to measure the impact of our actions and redefine new actions. [...] It contributes [...] to the involvement of all.

A survey of the limited existing literature demonstrates a relatively homogenous theoretical picture of the key aspects of a well-performing JSR; yet experience in the field suggests a far wider set of expectations and values attributed to JSRs, which diverge widely across education stakeholders. This calls for a more systematic review of the evidence through an action-oriented research agenda designed to learn more about approaches and expectations at the country level. It is our contention that developing a stronger evidence base will contribute to creating consensus concerning what core functions a JSR should serve. This should, in turn, help reinvigorate stakeholder views of the mechanism, and contribute to making a more effective use of JSRs in driving improved outcomes in the education sector.

JSR effectiveness is clearly embedded in the Global Partnership for Education (GPE’s) strategies. In October 2015, GPE published a draft Theory of Change (ToC), which lays out the core philosophy underpinning our approach to achieving improved education outcomes. The ToC forms the basis for GPE’s new Strategic Plan 2020 and its 2016–2020 results framework, which set out GPE’s longer-term goals and desired high-level impact and outcomes, namely strengthened systems delivering more equitable educational opportunities, and stronger learning outcomes. From GPE’s perspective, JSRs

lock planning and implementation together through inclusive and evidence-based *policy dialogue*. This reinforces the need for a broad array of stakeholders from across the sector to be involved.

In this paper, we take mutual accountability as a core construct, reflecting both its centrality to the primary aims of Sustainable Development Goal (SDG) 4 and its necessity as a key aid effectiveness principle for a sector aid partnership representing multiple stakeholders with divergent agendas to work effectively. Mutual accountability is broadly understood to reflect a compact under which two or more parties hold one another accountable for agreed commitments they have voluntarily made to one another. Acknowledging GPE's institutional focus on the importance of systems strengthening in conceptualizing the role of JSRs, we explore two interlinked questions, assuming that mutual accountability is a key objective achieved by JSRs where these are effective. We therefore first explore the ways education stakeholders conceptualize and understand the JSR's role in achieving accountability. We then examine how JSRs in GPE partner countries operate, by applying an analytical framework that defines JSR effectiveness in five key dimensions as well as showing how each of these dimensions intersect with the objective of mutual accountability.

We frame our research by asking two key questions:

- 1. How are JSRs perceived as promoting (mutual) accountability?** As above, we consider the mutual accountability principle to be of central importance in an aid environment, given the division of sector roles and responsibilities across a diverse range of stakeholders. We therefore pay particular attention to the competing views of these stakeholders in considering this question. We also recognize that selection of the mutual accountability lens limits us to primarily examining the relationship as though it acts in only one direction (i.e., the role of JSRs in fostering accountability), but of course acknowledge that in practice accountability is also a key input into the production of an effective JSR. Furthermore, our underlying assumption is that accountability, mutual or otherwise, is not an end in itself; it can support other equally important principles, such as ownership and transparency, as well as serving as a means to improved education outcomes.
- 2. How is mutual accountability likely to be affected by JSR effectiveness?** While JSRs are arguably a product of the aid effectiveness agenda and have only been researched in the development context, we contend that the capacity of a JSR to serve as an accountability platform is directly related to its effectiveness. Here, we apply an analytical framework that defines JSR effectiveness in five key dimensions: (i) inclusive and participatory, (ii) aligned to shared policy frameworks, (iii) based on evidence, (iv) a monitoring tool, and (v) an instrument for change embedded effectively into a policy cycle. This analysis provides valuable insights into how JSRs in a large sample of GPE partner countries currently operate and allows us to better understand how to position the JSR as a tool for flexible and adaptive planning. Our aim here is to stimulate further research into better understanding the pathways to creating successful JSRs. The dimensions included in our framework are intended as reference points for focusing dialogue among the community of policymakers, education practitioners and researchers.

The discussion of the evidence collected is framed in four sections. In **Section 1**, we discuss the multiple sources of data used to inform our study as well as the methodology employed to collect and analyze this data, showing how we intertwined inductive and normative research approaches in a complementary way. Our research strategy entails an inductive approach to understanding mutual accountability in JSRs that applies across stakeholders, and then superimposes this on a normative

conception developed by GPE of how JSRs should ideally be conducted. Applying a normative analytical framework to apprehend JSR effectiveness allowed us to better understand how JSR effectiveness intersects with the goal of mutual accountability and provided an opportunity to check the robustness of this framework.

In **Section 2**, we present a review of the literature, contextualizing the emergence of JSRs within the evolution of the architecture of international aid and providing a definition of JSRs (including their key characteristics) as we know them today. We also explain how both the literature and aid effectiveness principles inform the five key dimensions of the analytical framework for JSR effectiveness. Finally, we explore how other researchers have defined the concept of (mutual) accountability as well as how mutual accountability theoretically intersects with the JSR mechanism.

In **Section 3**, we discuss the conceptualization of mutual accountability within the JSR mechanism, based on the perspectives of key education stakeholders: governments, development partners and civil society organizations (CSOs). Here we make the argument that for JSRs to serve as an accountability platform, it is critical to understand the variety of expectations education stakeholders hold with respect to accountability in the context of a JSR. We find considerable deviation in views both across and, in some instances, within stakeholder groups. Development partners tend to reflect the international development rhetoric in a relatively homogenous way; government officials implicitly prefer national ownership of sector planning and implementation over holding other partners to account for key commitments; and CSOs (the most heterogeneous group) often view themselves as holding no particular accountability piece in the compact. These findings underpin our development of a more nuanced model of mutual accountability, based on an expanded understanding of sectoral responsibilities and differential distribution across partners.

Finally, in **Section 4**, we apply the analytical framework composed of five key dimensions describing JSR effectiveness to a sample of 39 JSRs from GPE developing country partners, so as to better understand how JSRs currently operate. We aim to highlight the ways in which the key dimensions of this framework intersect with the model of mutual accountability outlined in the previous section. We also discuss a number of weaknesses in the way in which JSRs are operationalized; these weaknesses serve as a threat to the effectiveness of the instrument, thereby reducing mutual accountability. At the same time, identifying these weaknesses allows us to consider ways in which the JSR mechanism may be strengthened.

SECTION 1: Research Strategy

This paper constitutes the first findings of a broader research project undertaken by GPE to explore the dynamics of what might define JSR effectiveness, with the longer-term objective of constructing a robust measure of this outcome. We adopted a mixed-methods approach, using a variety of primary and secondary data in complementary ways to enrich our discussion. This data was analyzed using a two-pronged approach, contained within distinct sections. In the first, Section 3, we apply an inductive lens in reviewing inputs from key education stakeholders in conceptualizing the construct of mutual accountability. Here, we attempted to compare and contrast the sometimes divergent perspectives of multiple education stakeholders against the broad understanding of mutual accountability embedded in recent international development compacts. Our objective is to delineate a more nuanced model of mutual accountability in JSRs that better reflects the reported interests and concerns of key actors across the sector. It is important to acknowledge that while we attempted to remain neutral on the issues, our views are inevitably colored by a GPE lens.¹ This approach is also exploratory, as data on mutual accountability and how this is supported by JSRs were extracted from instruments that were not specifically designed for examining this question, but rather to generate a broader understanding of the functions served by JSRs.

Then, in Section 4, we shift away from an inductive approach, applying a normative analytical framework that was developed as part of the GPE 2016-2020 results framework for monitoring the quality of JSRs in GPE partner countries. This methodology outlines five dimensions of JSR effectiveness. Here, our primary purpose was to provide a detailed descriptive snapshot of what could be observed about JSRs and how they operate today across a large sample of countries. However, given the extensive volume of data collected to support our broader research agenda, we also used the opportunity to refine this JSR effectiveness typology, making adjustments where these would not compromise the quality or interpretation of the data, which was collected for the GPE results framework prior to inception of this paper. Our descriptive findings are complemented by observations that arise from these other data sources.

We consider the use of these two approaches, namely the heuristic and the normative lenses, to offer two benefits. First, removing ourselves in our first findings chapter (Section 3) from the normative framework developed by GPE (Section 4) affords us an opportunity to critically review the construction of the JSR effectiveness methodology and to consider changes that could push us toward the development of a more robust framework. During the initial process of developing the GPE results framework JSR indicator, GPE drew on a wide range of resources, including literature and engagement with GPE stakeholders. Second, it allows us to be pragmatic in our use of available data. Furthermore, *not* exploiting this data source would considerably weaken the overall narrative we are able to build around JSRs and mutual accountability.

The data collected to inform our research are described below (further details on the methodology are presented in Annex 1).

¹ As outlined in the introduction, the ToC positions JSRs as the preferred monitoring mechanism in the education policy cycle as well as the central vehicle for achieving mutual accountability in the context of developing stronger education systems.

1. A comprehensive and systematic review of available literature (referred to as *literature review*)

The literature review was conducted to consolidate existing knowledge on JSRs, in order to both better understand how researchers perceive this mechanism as supporting accountability among stakeholders and what form effective JSRs take. The literature review included 49 articles,² mainly published by international organizations and based on a documentary review and single- or multiple-case studies. Of these, 10 articles (20 percent) primarily discuss JSRs, four of which are in the education sector. Approximately a third of the articles (17, or 35 percent) focus on sector-wide approaches (SWAps, discussed below), 10 of them (20 percent) on aid effectiveness, and the rest on monitoring and evaluation (12 percent), sector budget support (8 percent) and donor-government reviews (4 percent). Almost a third of the articles refer to the education sector (31 percent) and slightly more than a third refer to the health sector (39 percent); only 4 percent refer to the agricultural sector, and the rest refer to no sector, or to multiple sectors. A total of 57 different country cases were studied in the articles; on average, five countries were discussed per article. However, the same country cases were discussed in a number of articles: Cambodia, Bangladesh, Ghana, Mozambique, Uganda and Vietnam were the most frequently discussed across both the eight articles focusing primarily on JSRs and the full sample of 49 articles.

2. Interviews conducted with education stakeholders (referred to as *stakeholder interviews*)

Thirty-two semi-structured interviews with open-ended questions were conducted with key country-level stakeholders (government officials, development partners, CSO representatives and two GPE country leads³) across five sites: Bangladesh, Burkina Faso, Guinea, The Gambia and Zambia. Interviews took place between March and June 2016. The broad interview protocol investigated the experience of the interviewees and their roles in their organizations; their involvement in JSR processes; their understanding of a JSR and its functions or purposes; their opinions on an “ideal” JSR; and their perspective on the successes and challenges of the particular JSR in which the interviewee participated and the interviewer observed. The objective was to collect information on the perspective of JSR participants that could assist in identifying themes of relevance for understanding effective JSRs, in order to complement the findings from our normative framework (presented in Section 4). Although the interview protocol was not structured specifically to examine the question of mutual accountability, interviewees naturally discussed the issue, allowing us to leverage this material to inductively explore the links between JSRs and mutual accountability. Interview data was imported into NVivo, and coded and analyzed using this software.

3. An online survey of key education stakeholders (referred to as *stakeholder surveys*)

In September 2016, GPE administered a brief online survey to multiple stakeholders across GPE partner countries, which included a question interrogating the relative value that they place on different functions or purposes of JSRs. This was used to inform our analysis of stakeholder perspectives on the role of mutual accountability in JSRs. The survey was filled out by 70 representatives across 28 countries and included the views of governments, development partners (both donors and technical agencies), CSOs (encompassing NGOs, teachers unions and parents associations) and representatives from the private sector. The question as well as the related survey results are included in Annex 1.

² Articles used are noted in the bibliography with an asterisk (*).

³ Country leads are GPE staff who form part of the Country Support Team. They are responsible for supporting GPE processes across partner countries that form their portfolio. They also advise countries on sector policy processes, including sector planning and monitoring.

4. Data collected across 39 JSRs, examining five dimensions of effective JSRs, as part of GPE's expanded monitoring and evaluation framework (referred to as *variable analysis*)

In an attempt to better monitor outcomes, and to provide a stronger evidence base for policy dialogue, GPE's Secretariat developed a series of indicators under the umbrella of the 2016–2020 results framework, designed to align directly with the core goals and objectives outlined in the GPE 2020 strategic plan. As part of this process, an assessment questionnaire (detailed in Annex 1) was developed to evaluate JSR quality, as represented by five key dimensions identified as central to effectiveness. This approach entails a desk-based review of principal JSR documents to extract observations about a number of variables that we argue serve as proxies for core drivers of robust and effective JSRs. It is important to reinforce that while the results provide critical inputs into this research paper, this questionnaire was not developed primarily as a research tool, but rather prior to the elaboration of this research as a tool for GPE corporate monitoring.

This exercise enabled the construction of a unique database for JSRs conducted between July 2014 and December 2015, and for systematic comparison of JSR variables across states. This offers unique insights compared to the exclusively case-based approach in the existing literature. In Section 4 of this paper we draw heavily on this data to highlight findings across 39 JSRs.⁴ Correlations across the dimensions of this framework were statistically tested and we found little or no evidence of clear relationships. While one might assume that strength in one dimension could reinforce good performance in others, we take this as evidence of independence across the dimensions that were investigated, suggesting that the analytical framework is robust in its construction.

In addition, our insights are colored by a number of other sources, including direct observations of JSR meetings, interviews with four members of the GPE management team, and two focus groups with GPE country leads who were asked to reflect on their experiences with JSRs both at GPE and prior to joining the Secretariat. While these sources are not directly cited here, they helped us to generate a more holistic approach to answering our core research questions.

SECTION 2: What Do We Know about JSRs? A Review of the Literature

Despite their emergence almost 20 years ago, and their use today across a wide range of countries as a key monitoring instrument, there has been little systematic investigation of the role of JSRs in the strengthening of national systems. In order to build a robust research framework and develop a clearer conceptualization of the form and contribution of JSRs within the national policy cycle, we undertook an extensive review of the literature in three sectors: education, health and agriculture. To our knowledge, this is the first time that such a comprehensive review of JSRs has been conducted. We identified 49 articles,⁵ related directly or indirectly to JSRs, from which information about 57 different country cases was collected.

⁴ Note that this constitutes a different (wider) sample of JSRs than was included in the GPE 2016–2020 results framework indicator, which only included the 2015 calendar year.

⁵ Idem, note 1.

An analysis of these articles demonstrates that, to date, knowledge production on JSRs has been biased toward the publications of international organizations, based on JSRs conducted in health and education, with in-depth discussion of only a limited sample of country cases (primarily Cambodia, Bangladesh, Ghana, Mozambique, Uganda and Vietnam). The approach to analysis in these articles has largely been via documentary review, with the vast majority of publications in English,⁶ which may explain the limited selection of country cases studied. With these shortcomings in mind, we first provide historical context for the emergence of JSRs, before presenting the key themes that emerged from existing research and helped to inform the framework of effective JSRs discussed in Section 4.

JSRs in Historical Perspective: How Did the Mechanism Emerge?

Aid Paradigms

The first education JSRs recorded in the literature date back to 1999 in Mozambique (World Bank, 2001) and Uganda (Penny et al., 2008). In the same year, the first health care JSRs also took place, in Bangladesh (Martínez et al., 2013; Foster et al., 2000) and Ghana (Foster et al., 2000). They emerged at a time when the macroeconomic dialogue between donors and governments shifted away from the failure of structural adjustment programs to what came to be known as the “aid effectiveness agenda,” first formalized in 2003 at the First High-Level Forum on Aid Effectiveness in Rome and further developed at the Paris Declaration on Aid Effectiveness (Paris Declaration) in 2005. Propelled by the principles of ownership, alignment, harmonization, managing for results, and mutual accountability, frameworks for wider donor coordination were established. Common procedures were also adopted, particularly for joint missions, monitoring, progress reporting, procedures, and norms for technical assistance (OECD, 2005, 2008). Aid effectiveness emerged in parallel with the above-mentioned transformations in the broader macroeconomic dialogue between donors and governments, in line with a shifting focus toward expenditure management and the provision of essential social services.

In education, the aid effectiveness agenda shared principles with the Education for All (EFA) consensus that emerged in the late 1990s among developing countries and their development partners. The central tenet of the EFA consensus, as expressed in the World Education Forum in Dakar (2000), echoed the international commitment endorsed by members of the OECD Development Assistance Committee in the reference document *Shaping the 21st Century* (OECD, 1996). This declaration aimed to increase bilateral aid, harmonize donor activities and focus on development priorities, including universal education (Mundy, 2007). The EFA consensus was also based on the promise of increased donor coordination and support for country ownership, as delineated in the joint declaration, A Better World for All (IMF et. al, 2000), signed by key multilateral organizations (Mundy, 2007). It is in this context that the Education for All Fast Track Initiative (EFA-FTI), GPE’s predecessor, was launched in 2002 as an alliance between donors and developing countries to ensure accelerated education progress in the lowest-income economies toward EFA goals. Echoing global priorities of the time, EFA-FTI placed considerable emphasis on the development of education components in Poverty Reduction Strategy Papers (PRSPs) and/or the preparation of sound national education sector plans (ESPs) as central instruments for driving sector results via aid effectiveness principles.

Progress toward development results and alignment with the aid effectiveness principles at country level was appraised through multi-stakeholder “inclusive mutual assessments of progress” (OECD

⁶ Searches were conducted in English, French and Spanish.

TABLE 1 TIMELINE WITH KEY LITERATURE PORTRAYING JSR EVOLUTION

Timeline	Literature Review	JSR Evolution
1995	Discussion paper delineating sector investment programs (SIPs) at the World Bank (Harrold and Associates, 1995)	SIP supervision missions
1997	Guide delineating sector-wide approaches (SWAps) in health sector (Cassels, 1997)	SWAp joint appraisals
1999	First identified JSRs in education (World Bank, 2001; Penny et al., 2008) and health (Martínez et al., 2013; Foster et al., 2000)	First JSRs in education in Mozambique and Uganda, and in health in Bangladesh and Ghana
2005	Paris Declaration on Aid Effectiveness (OECD, 2005)	Mutual assessment reviews at the national level and JSRs at the sector level
2008	Accra Agenda for Action (OECD, 2008)	
2011	Busan Partnership for Effective Development Cooperation (Busan Partnership, 2011)	

and UNDP, 2014). These assessments are arguably comparable to today's JSRs as discussed later in this section, given that they were meant to track stakeholders' progress against agreed commitments, albeit focused at the level of overall development objectives and not at the sector level. These assessments also introduced the principle that such reviews should be *mutual* and *inclusive* (OECD, 2005, 2008, 20). The Busan Partnership for Effective Development Cooperation (Busan Partnership) in 2011 promoted a number of additional features, including the need to assess performance toward development targets and indicators in line with countries' policy priorities, the importance of regular assessment, the inclusion of multiple stakeholders and the wide dissemination of reviews (Busan Partnership, 2011, 5; OECD and UNDP, 2014).

Aid Instruments

Early forms of JSRs can be found in the monitoring mechanisms conceived as part of sector investment programs (SIPs), which were promoted by the World Bank in light of the unsatisfactory results achieved in individual aid projects and public expenditure allocations (Harrold and Associates, 1995) as well as government frustration with donor fragmentation and managerial overload (Cassels, 1997). Embedded in the same international aid paradigm shift—but articulated prior to formalization of the aid effectiveness agenda—SIPs were driven by recognition of the need for increased efficacy in aid disbursements at the sector level, as well as a shift from donor-driven to government-owned aid processes. This required substituting measures of success from individual donor project performance with measures related to government sector strategy performance (P. Harrold, 2016, personal communication, September 19). SIPs were distinguished by their comprehensive sector coverage, promotion of close coordination among government and donor partners, and a focus on local ownership and management of all sector processes (Harrold and Associates, 1995). SIPs did not exist as a separate lending instrument; they were envisioned as a new sector-based approach within

existing investment activities (P. Harrold, 2016, personal communication, September 19). Within the SIP approach, donor-driven “supervision missions” were seen as the core monitoring mechanisms, although these were called “annual review meetings” in practice (P. Harrold, 2016, personal communication, September 19).

The same shifting perspectives that gave rise to SIPs, as conceptualized by a single donor (the World Bank), prompted the parallel evolution of sector-wide approaches (SWAps), developed in collaboration with a wider range of development partners alongside the World Bank, including Denmark, Ireland, Norway, Sweden and the United Kingdom (V. Walford, 2016, personal communication, September 20; A. Riddell, 2016, personal communication, September 21). While SIPs and SWAps evolved as a consequence of the same underlying philosophy and shared key characteristics, they were developed in parallel rather than as linked mechanisms. Both were conceptualized as alternatives to financing development in Africa (Buchert, 2002), especially in countries heavily dependent on external donors (Harrold and Associates, 1995). A commonly accepted definition of SWAps demonstrates clear similarities in the philosophy underlying SIPs and the broader agenda of aid effectiveness: SWAps are referred to as a type of international development cooperation modality in which significant funding supports a single sector policy and expenditure program under government leadership, driving the adoption by all funding parties of a common approach across the sector and moving toward the use of government processes to disburse and account for financing (Foster et al., 2000; Riddell, 2007a). SWAps reinforced the idea that aid should be responsive to *national policies and priorities* (Boak and Ndaruhutse, 2011).

Under SWAps, “joint appraisals” organized between government and donors became the main monitoring and evaluation (M&E) instrument to assess progress across a sector (Cassels, 1997; World Bank, 2001). Joint appraisals were considered the culmination of a SWAp and were designed to assess policy implementation including strategic objectives, performance indicators, annual work plan and medium-term expenditure framework (Riddell, 2007a; Bartholomew, Takala and Ahmed, 2010).

Defining JSRs Today

With this understanding of the historical evolution of JSRs, we turn next to review how the literature can help identify the common characteristics and objectives of JSRs today, in an attempt to consolidate a working definition. There is no widely accepted definition of JSRs; however, Holvoet and Inberg (2009) identify a number of common characteristics of JSRs, many of which reflect key elements of the joint monitoring platforms envisaged under SIPs and SWAps, namely:

1. They are organized once or twice a year, for a duration of one to two weeks.
2. They are led by the sector ministry and engage various state and non-state stakeholders, and are divided into working groups to discuss thematic issues.
3. Inputs include a variety of primary and secondary data sources, mostly sector performance reports and financial reports prepared by the sector ministry and the Ministry of Finance, as well as commissioned studies on specific topics and information from donors and CSO projects.
4. Field visits may feed into the main discussions.
5. Conclusions and recommendations are gathered into an annual review report that is shared with all stakeholders. This constitutes the main output of the process.

Packer (2006) adds that JSR discussions are framed by an agreed-upon set of objectives, targets and performance indicators. In 2012, GPE conducted an unpublished study based on 67 JSR-related documents, produced between 2002 and 2011 in 21 countries (GPE, 2012a). This report concluded that there was considerable conformity across the basic characteristics of the education JSRs in the sample. Most JSRs are conducted on an annual basis, with field visits of three to five days followed by three to five days of a conference or workshop in the capital city. The latter typically includes both plenary sessions and working groups, during which presentations, official statements and formally agreed records of conclusions are documented. However, some variation exists across countries with respect to the operationalization of these characteristics (GPE, 2012a) and the terminology used to describe the review.⁷ In research conducted in the health sector, Martínez, Dodd and Walford (2013) identify two main objectives of sector reviews: (1) they allow stakeholders to take stock of sector progress at key points in the planning cycle, and (2) they facilitate agreement on strengths and weaknesses in implementation, enabling an informed discussion on how to continue toward sector progress.

This leads us to an examination of what constitutes an “effective” JSR, or one that is likely to achieve these objectives. Our definition of the key elements of effectiveness draws on a normative tool developed by GPE to inform its corporate results framework, and corresponds to the five dimensions around which our analysis in Section 4 is based. These comprise three key JSR characteristics, namely that a JSR should be (1) participatory and inclusive, (2) aligned to a policy framework and (3) evidence-based; and should serve two important functions, namely to act as (4) a monitoring tool and (5) an instrument for change embedded effectively into a policy cycle. We organize the research collected during our literature review along these lines, noting that the evidence collected during this element of our research project extends beyond what was initially collected to develop the results framework indicator, creating a far richer evidence base. This expanded review broadly supports the robustness of the original framework.

(1) Participatory and inclusive

JSRs provide an important platform for a variety of stakeholders to comment on sector progress and influence key priorities, strategies and future activities related to the ESP (Packer, 2006). In addition to the traditional categories of actors formally included in the negotiation of the aid partnership—governments and development partners (OECD, 2005, 2008)—the Busan Partnership (2011) provided a broader and more inclusive platform where CSOs and the private sector were framed as equal contributors to development.⁸ In particular, the inclusion of the CSOs in policy fora was emphasized in the EFA consensus, in which CSOs were expected to play two key roles: (1) legitimizing national plans and (2) holding governments accountable for commitments (ODI, 2006; Lister and Nyamugaisira, 2003; Brock et al., 2002; Gould and Ojanen, 2003; McGee, Levene and Hughes, 2002; Smillie and Lecomte, 2003, as cited in Mundy, 2008). However, the literature discusses the inherent challenge of the participation of *nationally representative* CSOs, which necessitates aggregating a diversity of local citizen voices at an upper governance level for the compact to work (Mundy, 2008).

Across the literature, it is also argued that the participation of different groups serves diverse and complementary functions. *Government* is in charge of reviewing the ESP, its strategic objectives, performance indicators and related operational tools, in collaboration with its partners (Riddell, 2007a); *development partners* prioritize the need for the accountability and reinforce the learning functions

⁷ In this paper, we use the generic terminology *JSR* for all multi-stakeholder annual reviews.

⁸ Historically, CSOs and the private sector were already acknowledged in the Paris Declaration on Aid Effectiveness and the Accra Agenda for Action, but only referenced in the compact through the actions of governments and donors (OECD, 2005, 2008). The Busan Partnership was the first attempt to formally include them among key education stakeholders, and CSOs were invited to the negotiation table [Busan Partnership, 2011].

of JSRs (Holvoet and Inberg, 2009); and CSOs and local representatives, as key agents in the delivery of services, provide a unique perspective on what is working, or not, from the community's perspective (Hattori, 2009). Cooperation across stakeholders therefore facilitates more robust policy formulation and implementation. However, attendance by a variety of actors at a JSR does not necessarily translate into meaningful engagement by all parties. Token participation by some stakeholders, in particular CSOs, is often reported as problematic. For example, Asuming-Brempong and Frimpong (2014) argue that "poor participation" in JSRs is often not a function of low attendance levels, but rather attributable to the *ineffective contribution* of participants.

(2) Aligned to shared policy frameworks

What constitutes a relevant policy framework on which a JSR should be based is extensively debated in the literature, with variation in practices both across countries and across sectors. In some contexts, JSRs are encompassed into broader monitoring processes associated with macro-level development or budgetary frameworks such as PRSPs or medium-term budget frameworks (Packer, 2006; Gerster and Chickwekwe, 2007; de Kemp et al., 2011). In the education sector, the compact that specifies the common goals and reciprocal commitments of stakeholders tends to be found in sector plans and strategies. However, regardless of whether the JSR is conceived of as part of a larger review or strictly as a sector-monitoring exercise, it should be based on an agreed-upon policy framework that formulates steps for action (Domingo et al., 2009). This provides the JSR with a clearly defined, shared agenda, along which stakeholders can coordinate on the basis of "shared principles, common goals and differential commitments" (Busan Partnership, 2011, 1). This in turn allows (1) cross-checking to determine whether stakeholders entrusted with specific roles and responsibilities carried out their assigned functions or tasks (Steer and Wathne, 2009), and (2) an examination of sector performance against a predefined and agreed-upon set of objectives and targets (Packer, 2006; Domingo et al., 2009; Penrose and Sacklokham, 2009; Martínez et al., 2013; Asuming-Brempong and Frimpong, 2014). Lack of consensus on the framework against which a JSR is based often results in parallel monitoring mechanisms targeting different audiences, which may contribute to divergent and competing lines of accountability (Domingo et al., 2009) as well as increased transaction costs.

(3) Based on evidence

A key principle of the aid effectiveness agenda is the importance of "managing for results," or the idea that *evidence-based decision making* is central to effective planning and policy implementation (OECD, 2005, 2008). It is argued that evidence should be viewed as "a measure of technical credibility" in monitoring, incorporating information that is clearly defined, quality assured, unbiased and transparent. Furthermore, evidence is the true "currency of accountability" (Droop et al., 2008, 18), given that sound policy dialogue can only be supported in the presence of a robust evidence base on progress toward key goals. To allow stakeholders in a JSR to fully analyze and discuss evidence, it is also necessary that governments ensure the timely provision of relevant data to develop an annual implementation report, the core document around which the assessment conducted during a JSR is typically framed. The production of quality annual implementation reports is a central input in ensuring the analytical value of a JSR (Holvoet and Inberg, 2009). Quality of annual implementation reports is often contingent on the robustness of the Education Management Information System (EMIS), which should produce timely and reliable quantitative data on education for supporting sector monitoring and decision making (Penny et al., 2008). The corollary here is that a weak EMIS may compromise JSR quality (Boak and Ndaruhtse, 2011). An area that has not been fully explored in the literature is the *relative weight* JSR participants place on different forms of evidence.

(4) Monitoring tool

Monitoring is the most commonly cited function of JSRs in the literature, where this is understood to be a review of sector performance over the past year, particularly with respect to plan implementation, financial management, previous JSR recommendations, and core government and donor-funded activities (Martínez et al., 2013; Packer, 2006; Asuming-Brempong and Frimpong, 2014; World Bank, 2001; Penrose and Sacklokhamb, 2009). Holvoet and Rombouts (2008) propose that, rather than representing a simple exercise in stocktaking and accounting outputs, the most effective JSRs reflect a more analytical approach that asks why targets were not met and questions the value of various targets and their related interventions, as well as how processes might be improved based on lessons learned. Good monitoring suggests the use of quality and on-time evidence to understand what is (and what is not) working and how best to address key bottlenecks. It was in this spirit that JSRs were conceived. As part of a sector-wide approach, the JSR should serve the essential function of providing a platform for dealing with problems in plan implementation within the sector (Cassels, 1997). It is important to note that aligning a JSR to a clear policy framework, per dimension 2, is a necessary but not sufficient condition for coherent monitoring, in that the framework sets the appropriate parameters for monitoring but does not guarantee its performance. Of course, aligning to a policy framework has benefits beyond effective monitoring, including improved coordination across stakeholders and greater consensus around JSR objectives.

(5) Instrument for change embedded effectively into a policy cycle

Many authors argue that a JSR may be regarded as effective only when it ultimately serves as a management and support tool for sector planning, programming, budget preparation and execution (Packer, 2006; Martínez et al., 2013; Asuming-Brempong and Frimpong, 2014). This presupposes that the four dimensions described above are operational in the JSR, culminating in a forward-looking consensus among stakeholders on how to work together to improve processes and generate better sector results (Steer and Wathne, 2009). Notably, existing research across multiple sectors highlights that the major challenge to this function is a failure to properly translate JSR recommendations into policy, operational planning and budgeting actions (Packer, 2006; Holvoet and Inberg, 2009; Hattori, 2009; GPE, 2012a; Martínez et al., 2013; Asuming-Brempong and Frimpong, 2014). GPE's earlier unpublished study (2012a) finds that recommendations are rarely prioritized and often scattered throughout the JSR output documentation. In the health sector, one study reported that JSRs very rarely formulate "SMART" recommendations specifying the "who, what, how and when" (IHP, 2013a).

However, the overall challenge does not lie simply with the robustness or number of the recommendations. It also reflects the absence of proper mechanisms for follow-up. This is said to contribute to a palpable fatigue, as stakeholders become less likely to see JSRs as an instrument for change (Packer, 2006; Holvoet and Inberg, 2009; GPE, 2012a; Martínez et al., 2013; Asuming-Brempong and Frimpong, 2014). Embedding JSRs into the policy cycle at the right moment is of critical importance in addressing this challenge. Harding and Gerster (2004) reported that sector review timing was too often arbitrarily set for the convenience of development partners, instead of being aligned with sector ministries' planning and budgeting cycles. Hattori (2009) and Asuming-Brempong and Frimpong (2014) suggested JSR recommendations should be operationalized as soon as possible, and the required financial and human resources allocated in an effective and timely way. The authors also

⁹ Indicating the acronym for specific, measurable, achievable, relevant and timely [SMART] recommendations (IHP, 2013a).

argue that linking the JSR with the annual planning process (or with the annual budget programming cycle if the latter does not occur) appears to be crucial for influencing negotiations with the Ministry of Finance on budget allocations. Packer (2006) discusses the potential value of embedding JSRs into a macro policy process such as the poverty reduction strategies, even though he could not report any evidence that this was actually happening by that time.

Mutual Accountability as a Construct and Its Role in JSRs

Accountability is a complex and multifaceted concept that may have different meanings for different stakeholders. However, as the central focus of this paper, it is necessary to provide a framework for what is meant when we discuss mutual accountability. We therefore adopt the working definition provided in a background paper prepared for the Third High-Level Forum on Aid Effectiveness in Accra (Droop et al., 2008, 10):¹⁰

Accountability is about holding those with responsibilities to account with the purpose of securing more responsive, efficient and effective behaviour. It means ensuring that actions are answered for, that performance is evaluated, and that consequent steps are taken. In essence, *mutual* accountability is the process by which two (or multiple) parties hold one another accountable for the commitments they have voluntarily made to one another. But it is also more than that. It is a process through which commitment to, and ownership of, shared agendas is created and reinforced by: building trust and understanding; shifting incentives towards results in achievement of shared objectives; embedding common values; deepening responsibilities and strengthening partnership; and openness to external scrutiny for assessing results in relation to goals.

Along with JSRs, the way the development community has operationalized the concept of accountability has evolved over time. The joint appraisal mechanism under SWAs had as its near-exclusive concern the provision of support to government in ensuring transparency and accountability for funds through adequate safeguards (Foster, 2000). This focus on financial accountability was reinforced by the aid effectiveness agenda, which called for both strengthened public financial management capacities from governments, and greater financial responsibility and aid predictability from donors (Hattori, 2009).

However, both the Paris Declaration and the Accra Agenda for Action (OECD, 2005, 2008) sought to broaden the perimeter of accountability and actively engaged with the concept of *mutual* accountability, holding partners of equal weights *mutually* accountable for “development results” instead of focusing only on funding. In these two framing documents, government and donors are seen as accountable to each other, as well as to their citizens, via parliaments. As noted previously, the Busan Partnership of 2011 extended the concept of mutual accountability by more formally integrating the full range of partners already identified in the Paris Declaration and the Accra Agenda for Action, namely the private sector and civil society actors. In the education sector, the EFA consensus also allowed for the emergence of new actors, like networks of international nongovernmental organizations working for the achievement of EFA goals. This marked a significant change, given that they were not part of

¹⁰ In addition, we note that within the accountability literature, some authors distinguish between “horizontal” accountability, which implies a degree of reciprocity among equal partners (typically government and donors), and “vertical” or downward accountability (typically government and citizens). In this paper, we largely avoid this distinction, given that mutual accountability broadly reflects the spirit of horizontal accountability only.

any international development compact, despite the activism of international teachers unions and humanitarian and religious organizations from the early 20th century (Mundy, 2007).

But what constitutes mutual accountability in practice, especially in the context of JSRs? References to the JSR as a forum for monitoring and reporting on sector performance are common in the literature (Boak and Ndaruuhutse, 2011; GPE, 2012a; IHP, 2013a, 2014a, and 2014b). More broadly, according to Domingo et al. (2009), successful mutual accountability mechanisms entail three stages: (1) generating common goals and commitments; (2) monitoring and reviewing progress toward these commitments; and (3) engaging in dialogue around progress and next steps. This reflects the way a JSR is expected to function. This, in turn, serves to improve accountability among stakeholders.

SECTION 3: JSRs as a Mutual Accountability Platform

Education Stakeholders' View of Accountability in the Context of JSRs

A key observation emerging from the literature review in Section 2 was that a preponderance of researchers pointed to the importance of the JSR as a forum for monitoring and reporting on sector performance in light of assigned stakeholders' roles and responsibilities. Through this channel, JSRs are argued to serve as a vehicle for improving accountability among stakeholders. The value of JSRs as a platform for accountability is reinforced by outcomes of the GPE stakeholder survey, which explored the relative value that key stakeholders place on different purposes of JSRs. When aggregating across stakeholder groups,¹¹ all accountability functions taken together (government accountability toward donors and toward its citizens, as well as mutual accountability) were identified as among the most important outcomes of JSRs (21 percent of responses)—second only to the expectation that JSRs would create more country ownership over plan implementation (35 percent), as elaborated in Annex 1.

While the stakeholder survey (Table 2) shows that most key players agree that accountability is an important function of JSRs, it is not clear that stakeholders have a shared understanding of what the concept means in practice. In this section, we therefore use our stakeholder interviews in combination with the survey results to more clearly delineate the ways in which key education stakeholders (government, development partners and CSOs)¹² conceptualize the nature of accountability as operationalized in the context of JSRs, and how the global discourse of "mutual accountability" translates into realities on the ground. We examine these issues by considering each group of stakeholders in turn, highlighting the variation in views across and within each group. We begin by considering the inputs of development partners, given that the views of this group are relatively homogenous, typically aligning most closely with the rhetoric outlined in the international development compacts discussed in the previous section. We then consider the perspective of government, the stakeholder we regard as central to the planning and implementation of sector policy, before turning to consider the inputs of CSOs.

11 Government, development partners, NGOs, teachers unions, the private sector and academic organizations.

12 Note that these categories in the interviews do not directly align with those under the stakeholder survey. In particular, teachers unions are considered to be part of the CSO group for the interview analysis, while no private sector and academic stakeholders were interviewed, given their limited participation in JSRs.

TABLE 2 STAKEHOLDER EXPECTATIONS ON JSRs AS A PLATFORM FOR ACCOUNTABILITY BY STAKEHOLDER (% OF N)

	Government (N = 26)	Development Partners (N = 118)	NGO (N = 47)	Teachers Union (N = 6)	Private Sector (N = 3)	Academy (N = 3)
Government accountability toward donors	0	2	6	0	33	0
Government accountability toward citizens	12	9	15	0	33	0
Mutual accountability	4	9	13	17	0	0
All kinds of accountability functions	16	20	34	17	66	0

Notes: Table is based on original responses by all participants (not weighted) in the *stakeholder survey* (i.e., does not incorporate interview responses). The unit of analysis corresponds to the number of instances in which representatives from a stakeholder category marked a particular function as important, and is presented as percentage of total responses (N) for each stakeholder category. In this survey, other JSR functions were included. Percentages will never sum to 100 percent horizontally or vertically, since this table reflects responses related only to the accountability function. Results therefore cannot be compared across stakeholders (horizontally); only within the same category of stakeholders (vertically). For further details see Annex 1.

Our aim here is to compare how ideal aspirations for JSRs, as expressed in the stakeholder survey and aspects of the stakeholder interviews, might diverge from the on-the-ground realities articulated in the interview data. We acknowledge a significant limitation—namely that interview discourse is inevitably tainted by our respondents' institutional perspectives and associated objectives for JSR performance, and as such may not always reflect actual experience. However, we argue that where clear digressions from the theoretical ideal are highlighted by interviewees, these can provide useful insights into how JSRs might be improved. Given that in many instances respondents are likely to reflect an institutional bias in their discussion, we are of course cautious not to accept answers at face value, but instead evaluate stakeholder commentary against other data sources. This data is also useful in allowing us to triangulate observed JSR practices against what key players argue is the case.

Perspective of Development Partners

The position taken by development partners in the interviews aligns most closely with much of the commonly cited rhetoric around accountability, with the concept of mutual accountability frequently raised in conversation. This corresponds to the survey results, which show that for this group, JSRs represent an important opportunity to develop mutual accountability. The fact that development partners typically employ "mainstream" vocabulary in very similar ways is unsurprising given that

much of the mutual accountability discourse derives from the aid effectiveness agenda, which was primarily donor-driven.

“I believe it [a JSR] should be a platform for mutual accountability because the idea is to look collectively to all of the education stakeholders at the progress and the performance of the sector as a whole. And as we consider we all want the advancement of the education sector, then mutual accountability should be from all of us towards all of us.”

—Development partner representative

Development agencies argue that the JSR should not be seen as a static, one-way reporting mechanism in which government is the subject of interrogation by partners, but rather as an open forum where all stakeholders are held equally accountable. Mutual accountability is often flagged by development partners as being best fostered in an environment where government is not positioned in a defensive role. This provides some rationale for opening JSRs to a wide range of participants, whereby, at least in theory, both successes and failures stand to be distributed across multiple groups rather than blame being placed exclusively on government. The underlying assumption here is that if governments and development partners plan together, they must also implement together and ultimately also review together, learning from each other in the process. As one development partner noted, “We plan together and then we go and implement and [we] come together to look and critique.”

However, it is worth noting that the country-level label “joint donor review” employed by some donors within the development partners group we interviewed suggests that, despite the rhetoric of mutual accountability, in practice there is not always a clear move away from the static, one-way reporting line under which governments are held accountable for the expenditure of donor resources. This divide between the theory of mutual accountability as discussed by development partners versus the actual practice of one-way accountability is reinforced by the relative disconnect between the survey and much of the interview data. In the survey, development partners almost never highlighted the JSR as an instrument for the government to hold donors accountable. In contrast to our survey findings, during interviews some donors pointed to the ways in which JSRs could reinforce donor confidence by making it possible to evaluate the status of their investment when considering sector progress and monitoring whether their money is being used efficiently by governments.

A number of development partners report that the JSR provides an opportunity for them to present their own achievements against objectives and evaluate their contributions to the implementation of policy, specifically with respect to action plans. While on the one hand this may be regarded as positive—donors identifying themselves as accountable for key deliverables, thereby reducing pressure on government and enacting one interpretation of mutual accountability—it also reflects a shift away from government ownership of program implementation. Interestingly, however, none of the development partners interviewed discussed their responsibility either for aid disbursement and predictability or for the technical support they have committed to provide.

Since in most contexts the majority of expenditure for education is funded domestically by local taxpayers, development partners agree that JSRs should also promote accountability to citizens. This is mirrored in development partners’ survey results, where vertical accountability (i.e., government to its citizens) is cited with the same frequency as mutual accountability as an expected JSR outcome. To the extent one considers that CSOs serve as citizen representatives, development partners report

viewing CSOs (and therefore, by proxy, citizens) as integral to monitoring by providing a perspective on how the sector plan has translated into implementation on the ground. However, they do not directly address the ways in which CSOs form part of the mutual accountability contract.

Perspective of Government Officials

Shifting to consider the perspective of government, we also observe some degree of intra-group homogeneity, in particular in the way government officials understand the main lines of accountability. The latter typically challenges the mutual accountability rhetoric embraced by the development partners group, reinforcing the problems inherent in a simplistic portrayal of mutual accountability as a multi-way channel with responsibilities flowing equally in two or more directions. A more nuanced picture of mutual accountability can be drawn from the government interviews, in which the ministry is typically described as being at the center of the mutual accountability architecture, bearing primary responsibility for the delivery of sector results in partnership with others.

“Well, the Ministry has implemented the program and it talks frankly about it with partners: ‘That is what we planned, this is where we are.’ This allows us to take a critical look at our action and it allows us, the group in charge of the implementation, to be more attentive to certain aspects.”

—Ministry of Education official

Indeed, government officials most often describe accountability as operating in only one direction, with the government as the primary actor accountable to a number of other stakeholders. The most commonly discussed line of accountability across the interviews (and arguably the primary interest of the government group) is that JSRs allow government to attain a stamp of approval with respect to their performance on sector implementation.

“Donors have engaged with us around a program, so it is normal that every year we justify what we have done with the money. Which resources have we used? How? Why? What we do here is presenting [*sic*] results. That’s how we can see if their contributions really allowed us to move forward, and we also draw lessons from each other.”

—Ministry of Education official

Government officials typically feel that a JSR is a forum in which they can show development partners, and donors in particular, how resources linked to a particular implementation program have been used; an opportunity that in their view serves to reassure donors and builds trust in their capacity to develop robust plans and deliver on policy implementation. Aligning with what was expressed by development partners, government officials see JSRs as a moment in which development partners are able to monitor and evaluate the extent to which the government delivered on its promises. Holding partners, including donors, accountable in the same way is not typically an objective expressed by the government group. In an aid-dominated environment, they recognize that they have limited leverage to impose their own expectations on their financial partners. The lack of levers for enforceability built into the JSR mechanism seems to disadvantage governments more than donors: Donors can still use aid allocation and disbursements as a lever for enforcement if government commitments are not met, while government officials are arguably more reluctant to refuse external aid as a parallel means of compelling cooperation. This is reflected in the survey results, in which no government officials refer to JSRs as an accountability instrument for their donors. In theory, both donors and government

officials agree that JSRs are not meant to promote smaller-scale, one-way accountability; in practice, they also agree that government being accountable to financial partners is an essential element of a JSR for building mutual trust and strengthening collaboration within a partnership approach.

Rather than viewing government interview data as grounds for refutation of the mutual accountability principle, the interview data from government officials may be seen as providing a complementary interpretation. Under this alternative framework, governments are the key implementing agents within a partnership environment, providing necessary reinforcement for their role as ultimate owners of the sector plan. This does not necessarily need to translate into government being subject to the scrutiny of other partners who are exempt from any responsibilities; instead we argue not only for an expanded understanding of responsibilities, but also that there are alternative lines of responsibility for other stakeholders, including the provision of constructive inputs and technical support to assist government in working toward better service delivery. Donors and technical agencies in particular are critical to the development and enhancement of government capabilities. This arguably reflects a broader sense of mutual accountability than the simplistic two-way model often assumed in the literature, generating a more nuanced view of JSRs as an agreement among stakeholders. This was reflected by one Ministry of Education official, who observed: “The JSR is a contract of trust between partners working together. It is quite normal that at some point, once a year, we take stock of the situation.”

Government officials regularly pay lip service to the importance of accountability toward their citizens, and citizen accountability is identified by this group in the survey results as the most important function of JSRs. However, no respondents spent any time delineating the importance of this in the interviews, with the notable exception of one interviewee who was also the sole parliamentarian in the interview sample. The only way in which accountability to citizens is discussed is in reference to ensuring that CSOs have a voice and are able to challenge government reporting effectively.

Perspective of CSOs

Finally, we consider the perspective of CSOs with respect to accountability in JSRs. This is arguably the most heterogeneous group within the interview sample, representing a broad range of constituencies and views. Accountability in itself is not extensively discussed in these interviews, and is mostly referred to as an intangible concept that crystalizes the international community’s attention but does not echo country-sensitive and locally grounded concerns. This may be linked to the fact that systems-level thinking is less likely to be a feature of the dialogue within this group. While CSOs have long been active in shaping the international dialogue around education, it is only relatively recently, under the Busan Partnership (2011), that they have been fully and formally included in the international development compact. This may also help to explain why accountability rhetoric is newer to this group than to development partners and governments, and consequently is less exploited in the form of lip service.

However, the absence of articulation around the issue of accountability in the CSO interviews also mirrors the position taken by CSOs, who identify a tension between what they perceive as their crucial role in speaking to on-the-ground realities in education and their position as outsiders to the tight partnership formed by governments and development partners. The CSOs interviewed suggest that the responsibility of accountability falls primarily on government and development partners, although in different ways. In their view, these two partners are the key players responsible for the monitoring, evaluation and supervision of sector plan implementation. It is consequently not

surprising that there is no mention of the issues for which the CSOs should be responsible and held accountable. CSOs do not actively speak about the ways in which they might be held accountable, nor do the other stakeholder groups delineate clear lines of accountability for CSOs. This is in spite of the fact that in some cases, CSOs may be directly responsible for implementing parts of the sector plan.

Under the CSO model, accountability is more often viewed as a one-way channel, in which government should be answerable for any shortcomings in delivery. This reinforces the need for a broader understanding of how mutual accountability is actually practiced with CSOs playing a specific and unique role in the JSR arena. Their inputs often suggest that JSRs are an effective lobbying forum that allows them to draw attention to a set of specific issues that aligns with their particular agenda: from increased government allocations of domestic resources to education, to addressing the needs of the most vulnerable populations. Beyond this advocacy role, and perhaps more importantly, most CSO representatives interviewed saw their organization as performing a challenge function that is critical for moving away from the technocratic and relatively superficial exercise of JSR reporting. CSOs bring insights—most often in an anecdotal form, but on other occasions, when capacities are developed, built on a more robust evidence base, as a counter to government reporting or as additional evidence of what is happening on the ground. One CSO representative provided a useful example: “For instance, we really need to ask why gross enrollment rate was expected to be at 91 percent in 2015, and has not been achieved, while gender parity has been reached. These are questions to ask ourselves.”

CSOs challenge both government authorities and their financial partners to see beyond the surface of quantified indicators, pushing them to examine the root causes of implementation bottlenecks. Even if CSOs do not identify particular areas for which they are themselves accountable in the implementation of a sector plan, they reinforce a key accountability function by demanding “answerability” from the main stakeholders: governments and donors. This stands in contrast to the discourse analyzed above, where donors were explicitly painted by CSOs as the main actors holding governments accountable.

In conclusion, it appears that the lines of accountability are principally framed in terms of the responsibilities stakeholders have been assigned, or have assigned to themselves, for jointly contributing to the delivery of education results. The international discourse on mutual accountability may be understood as implying that the roles adopted by the three main categories of stakeholders (government, development partners and CSOs) carry equal weight; in this sense they bear the same level of accountability. Considering the ways in which stakeholders contribute to JSRs and perceive the question of accountability, mutual or otherwise, allows us to draw a more nuanced picture of the situation. From the interviews, it seems all stakeholders agree that in practice the main responsibility for improved education outcomes lies with governments. This line of accountability reinforces government ownership over sector plan implementation, with development partners and CSOs feeding into the chain of results more indirectly (by supporting government) than is articulated in the view of mutual accountability commonly proposed by the international rhetoric. Although multiple stakeholders may be engaged in aspects of plan implementation, we contend that this is ultimately the domain of government, with development partners primarily accountable for the provision of support to the government by supplying financial and technical support, while CSOs are primarily accountable for serving a strong challenge function grounded in constructive feedback to those responsible for policy implementation. The JSR then serves as the vehicle that creates space for each of these stakeholder groups to coordinate and assess their efficacy in fulfilling their respective roles.

SECTION 4: How Mutual Accountability Is Likely to Be Affected by JSR Effectiveness: An Empirical Assessment

In order for a JSR to serve as an effective input into an education sector policy cycle, it requires on-time, robust and frequently reported evidence (dimension 3 “based on evidence”) in order to assess and learn from past sector implementation (dimension 4 “monitoring tool”), against an agreed policy framework, typically the education sector plan, which includes shared sector objectives, targets, lines of responsibilities and activities (dimension 2 “aligned to shared policy frameworks”). This must be done in a participatory and inclusive manner, bringing together all key education stakeholders (dimension 1 “inclusive and participatory”), leading to the identification and prioritization of agreed remedial actions (dimension 5 “instrument for change embedded effectively into a policy cycle”) to address in the subsequent year(s). This serves as a useful basis on which to monitor, and possibly revise, the agreed policy framework (dimension 2), ideally in a virtuous circle of accountability (where greater accountability in monitoring reinforces greater accountability in planning and implementation), embedded into a flexible and adaptive learning cycle.

The five dimensions outlined above form the basis of our analytical framework, which proposes a set of characteristics and functions that an effective JSR should entail and perform. Our assumption is that JSRs demonstrating these dimensions are more likely to serve as a mutual accountability platform. We applied this analytical framework to 39 JSRs held between July 2014 and December 2015 in GPE partner developing countries. In this section, the proportion of JSRs is strictly described relative to the total number of *conclusive* cases (the denominator), which changes across items according to the availability of information in the JSR documents assessed.¹³ The findings of this variable-based analysis were also enriched with interview data to better delineate in what ways these effectiveness dimensions contribute to mutual accountability. When relevant, we included references to our literature review to support or contradict our key findings.

Dimension 1: Inclusive and Participatory

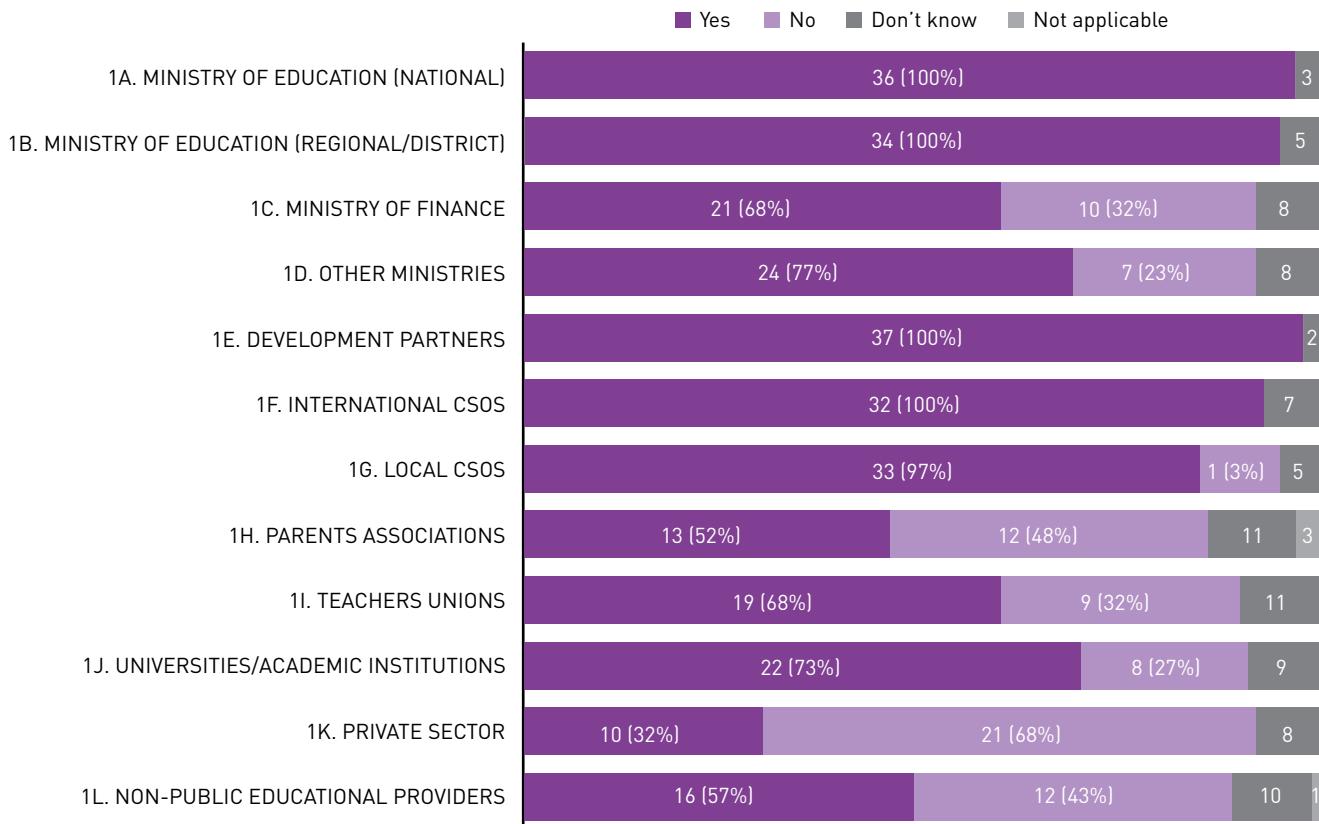
DEFINITION

A “joint” process necessarily implies representation of all stakeholders in the sector review. We see dimension 1 as representing key but distinct elements of representation, with *inclusion* reflecting who is at the table (are the right people in the room?) and *participation* reflecting the engagement and contribution of stakeholders to the proceedings. A truly participatory process also entails the provision of adequate support (technical, financial and logistical) to stakeholders who may not otherwise have the capacity to fully engage.

¹³ As discussed in our methodology (Section 1), in a number of JSRs, responses to particular items are inconclusive due to missing or unclear documentation. It is therefore not possible to report on whether these items were met or not.

FIGURE 1 DIMENSION 1: PARTICIPATORY AND INCLUSIVE

Interview questions: To what extent was the JSR inclusive? From which of the following stakeholder categories was there attendance during the JSR?



The desk-based nature of the variable analysis limits the extent to which true participation can be assessed within this framework; to illuminate this aspect, we draw on our stakeholder interviews.

ANALYSIS

Our variable analysis of stakeholder participation shows that the Ministry of Education, at both the national and regional levels, along with development partners and CSOs, are typically present (Figure 1). The high level of inclusion of national and international CSOs (1F and 1G) is encouraging, given that earlier research has suggested that CSOs were not systematically included in JSRs and that representation was often only partial (Packer, 2006, GPE, 2012a). The actual level of participation of CSOs depends heavily on country context, notably on the extent to which openness and transparency of dialogue is encouraged by governments as well as on CSOs' capacity to actively contribute. Our interview data suggests that CSOs usually see themselves as playing an observer role, often attending JSRs to make the same observations year after year, or to access critical information about sector performance that could reinforce their understanding of sector issues but does not necessarily contribute to dialogue on broader sector goals. The result of this arguably ineffective participation is to diminish sector accountability (Asuming-Brempong and Frimpong, 2014), particularly given the key role this group is envisaged to play under international development agreements, including EFA. Without meaningful input from CSOs, arguably the group closest to classroom realities, it is possible

that the proceedings of a JSR do not fully address key bottlenecks in implementation of the sector plan. Furthermore, high-level technical documents may also mask specific deficiencies, which CSOs may be best positioned to identify.

It is problematic that in a third of JSRs with available information,¹⁴ a representative from the Ministry of Finance was not in attendance (1C), underscoring limited buy-in to the idea of JSRs as a tool for financial accountability at the national level, and the extent of challenges in facilitating intra-governmental dialogue. Coherence between sector and national financing frameworks is central to collaboration on transparent, adequate and efficient education spending. This is particularly challenging in contexts where advocacy for increased domestic financing on education is key. Based on our stakeholder interviews, the relatively poor attendance of parents associations (in only half of JSRs with available information, 1H and 1I) and teachers unions (in only one-third of JSRs) may be anecdotally attributed to the opportunity costs, in terms of travel and foregone income. It may also translate into a lack of informed dialogue centered on school-level realities and teaching-and-learning matters. This highlights the troubling absence of a direct feedback loop from the primary beneficiaries of the sector policy.

Our findings suggest relatively high attendance of other ministries (1D)—greater, in fact, than Ministry of Finance attendance, with 24 JSRs (77 percent of those with available information) attended by ministries other than education or finance, compared to 21 JSRs (68 percent of those with available information) attended by the Ministry of Finance. However, with more than a fifth of JSRs not engaging with other ministries, further steps could be taken to increase participation by government actors outside of the line ministries in charge of education, so as to foster intra-governmental and cross-sector dialogue. Participation by the private sector and non-public educational institutions is also relatively low, and deserves attention.

The interview data and literature review suggest that the way a JSR is organized may impact effective participation. Logistical aspects, such as ensuring enough time and space for discussions, and having professional moderators who can effectively facilitate exchanges and ensure that all voices are heard, are among the practical issues highlighted here. Overall, the evidence points to a tension between representation of all stakeholders on the one hand, and participation effectiveness and quality sector dialogue on the other. While it is important that all stakeholders are represented in JSRs, a high number of attendees might limit the quality of discussions if proper mechanisms for engagement are not implemented (Holvoet and Inberg, 2009). An effective joint review that generates quality policy dialogue requires the perspectives, and sometimes the opposing views, of multiple actors. As such, the premise for effective participation identified by interviewees is that all stakeholders (in particular the government, in leading the JSR process) should be open to dialogue and criticism as well as self-reflection. The ability of different stakeholder groups to fully engage in technical discussions in those areas requiring some form of specialized knowledge (for example, finance and budgeting, or the interpretation of data and trends) is another aspect identified as key to the quality of the dialogue, highlighting the importance of capacity development for more meaningful engagement (IHP, 2013c). The “right people” and the “right number” of participants depend on country context, but determining these variables should be seen as central to facilitating quality policy discussions, and to the inclusion of multiple perspectives.

¹⁴ The desk-based nature of this analysis introduces a potential confounding variable: namely that, in some cases, it was not possible to determine whether particular items were met or not met based on the core documents obtained. These were coded as “inconclusive” (the “don’t know” category in the graphs), and should be clearly distinguished from standards or items that are “not met.”

Dimension 2: Aligned to Shared Policy Frameworks

DEFINITION

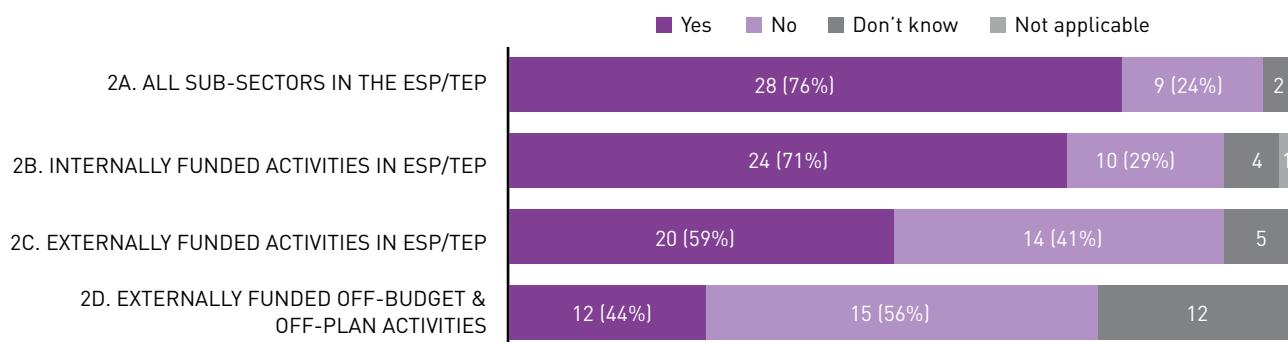
In our analytical framework, we view the national ESP as the shared policy framework defining the perimeter for the areas covered by a JSR. Here we consider not only the sector plan itself but also its operational subsets, such as the multiyear and/or annual action or operational plans. These form the basis for building common ground across the education development compact, as reflected in responses across all stakeholder groups: Most of those interviewed pointed to the ESP and associated policy documents as the key instrument for monitoring implementation and holding relevant stakeholders to account for assigned responsibilities.

ANALYSIS

Given that our variable analysis of this dimension (Figure 2) can only examine proxies for underlying dynamics, it provides us with a limited view of the extent to which JSRs are actually based on ESPs. Note that we consider comprehensive coverage of ESP to be an outcome of the JSR process, not simply an outcome of the review meeting. In our analytical framework, JSR coverage of the education sub-sectors and activities included in the ESP serve as basic proxies for determining whether the sector plan has been used as the main framework for driving the discussion and holding stakeholders accountable for possible shortcomings and results. Examination of agendas and implementation reports used during JSR meetings reveals that in more than 25 percent of cases, all sub-sectors included in the original ESP were not systematically covered in the implementation report, nor discussed during the JSR (2A). Performance is even weaker when looking at ESP activities by type of funding (2B and 2C); in particular, 34 percent of JSRs do not comprehensively track internally funded ESP activities, and that share is, unsurprisingly, larger (about 40 percent) when considering whether JSRs discuss activities funded by external partners. These findings speak to a troubling disconnect between the contents of an ESP and what JSRs actually monitor, implying poor alignment across these tools.

FIGURE 2 DIMENSION 2: ALIGNED TO SHARED POLICY FRAMEWORKS

Interview questions: *Did the JSR cover the full spectrum of the education sector as addressed in the ESP?¹⁵ Which of the following items apply?*



Note: Items coded as “no” include both those recorded as “no” and those recorded as “partial coverage.”

¹⁵ TEPs (or transitional education plans) are typically developed in contexts where development of a longer-term ESP is either technically unfeasible or inadvisable due to crisis or other challenges. The variable analysis questionnaire was designed to allow for analysis of the JSR in contexts where both ESPs and TEPs define sector plans. In the text, we use the term “ESPs” to describe both forms of the education plan.

Given the important role of the ESP in creating a shared platform for discussion, a critical question here is whether or not generating a JSR agenda based on the ESP translates into poor consensus around JSR outcomes for different stakeholder groups. Additional research should be conducted to understand whether this is due to a partial use of the ESP as the reference instrument, or whether other monitoring frameworks were used for conducting the sector review. In previous research by GPE (2012a), findings suggested that ESPs were not used as a “shared agenda” for JSR discussions in a majority of countries, either because the ESP was not seen as the most relevant document for this purpose, or because key stakeholders demonstrated little or no awareness of its contents. In countries where a joint financing agreement between government and donors was developed, examination of the plans outlined in these documents usually had preeminence over ESPs in defining the JSR agenda (GPE, 2012a). This is echoed by respondents in our interview data who pointed to the implementation of additional review mechanisms, for instance for dealing with disbursement-linked indicators in results-based financing programs, aside from the JSR. As outlined in our literature review, this is likely to contribute to the duplication of efforts and increased transaction costs in monitoring, reflecting poor coordination across donors and poor alignment of individual donors to country systems. This may undermine the development of robust country planning and M&E mechanisms.

Considering off-budget items, in 44 percent of JSRs with available information, activities implemented with external funding that were not included in the ESP were discussed during the JSR. This opens up a larger discussion about the relationship between the initial quality and comprehensiveness of policy frameworks and the effectiveness of sector reviews. This also points to the importance of ensuring the alignment of development partners to support national planning documents. In particular, it is important to consider whether the fact that at least half of JSRs covered activities not included in the ESP points to deficits in the initial planning documents. It also paves the way for further analysis under dimension 5 of our analytical framework, where we discuss JSRs as an opportunity to generate remedial mechanisms to address policy instrument shortcomings.

Dimension 3: Based on Evidence

DEFINITION

A robust evidence base is central to an effective JSR, allowing assessment of whether strategies are implemented as planned as well as measurement of the impact of policies on sector development (Droop, Isenman and Mlalazi, 2008). As outlined previously, in the context of JSRs, this requires the timely and widespread dissemination of quality data, both quantitative and qualitative, produced by a variety of stakeholders, and most commonly shared with key stakeholders or the wider public in the form of government annual implementation reports, mapped to annual or multi-year operational plans or ESPs in other instances. Other documents that also support the development of a strong evidence base for JSRs include financial reports (prepared by governments, development partners, CSOs and other stakeholders) on budgeted and actual expenditures; commissioned studies on specific sector themes or the impact of specific policies; and documents from past JSRs. Effective collection, production, analysis, dissemination and use of data through EMIS is a central ingredient here; as such, building data capacities is a key objective for governments, supported by donors under the mutual accountability compact. Reporting documents produced for and during JSRs should leverage country systems, in order to reinforce capacity-building in this area.

ANALYSIS

Government officials in our interview sample highlight the importance of JSRs as a mechanism for information sharing, and see this as inseparable from strong data production and reporting. The

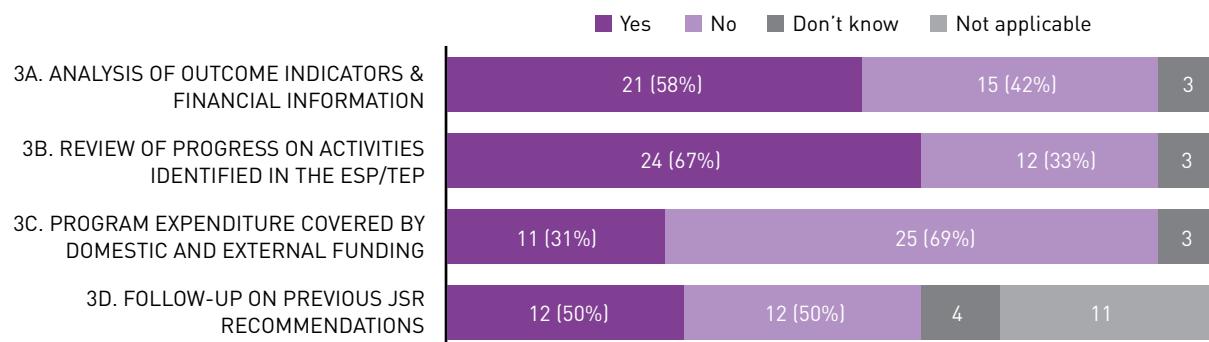
initial step of preparing JSR documents is seen as part of the mutual accountability function of JSRs, given that transparency and integrity of the information presented in reports determines the quality of discussions. Officials also pointed to the centrality of an annual implementation report as the basis for quality discussions and presentation of evidence on results achieved. However, they also emphasized the capacity challenges linked to collection and analysis of robust data. Development partners and Secretariat staff highlighted the tension between producing quality reports and ownership of reporting—international consultants are often viewed as having the skills to produce higher-quality reports, but this necessarily compromises what interviewees highlight as a more important outcome: country ownership. Another tension revealed by the interviews is the production of shadow reports and the active participation of stakeholders in JSRs. CSOs' inputs, sometimes formalized in "alternative" reports, are often viewed as useful resources from which to discuss issues identified by local communities and as ways of verifying data and analysis provided in government reports. Even where these additional contributions are regarded as valuable, they may simultaneously produce challenges for the reconciliation of divergent views and voices if CSO inputs do not align with JSR documents produced by government. In particular, CSO contributions may not be adequately reflected in JSR output documentation, which should in theory represent all stakeholders in constructing a joint analysis of the sector situation.

This dimension considers in the first instance the evidence base of a JSR when an annual implementation report is available, given the report's centrality to sector monitoring and evaluation functions. This means that the report should be of high quality and shared in a timely fashion with all stakeholders. An important caveat in reviewing findings from this dimension is that it was possible to obtain an annual implementation report in only 29 JSRs out of our sample of 39. Without this report, the items associated with this dimension could not be analyzed, so results for 10 JSRs (more than a quarter of the sample) were inconclusive. It is problematic that in so many cases no report was available, particularly given that this likely reflects that a robust and widely disseminated evidence base did not form the core of the JSR discussion in these cases. This dimension also considers whether a JSR is informed by evidence, including reliable education and financial data from the year under review, assessments of program implementation, documentary inputs and feedback from beneficiaries.

Figure 3 shows that while reporting against milestones at the program and activity levels as outlined in sector plans was relatively strong (3B), more than two-thirds of the JSRs with annual implementation reports available did not cover program expenditure, regardless of the funding source (3C).

FIGURE 3 DIMENSION 3: BASED ON EVIDENCE

Interview question: *Did the annual implementation report include the following?*



This suggests a clear gap in reporting and the need for greater provision of technical support for gathering appropriate financial evidence, in order to make discussions on expenditure evidence-based. The lack of reporting on expenditure is of concern, as this suggests the JSR's capacity to serve as a forum to discuss financial matters is not being fully exploited, limiting the extent to which it informs sectoral and national budget programming. A bottleneck in this aspect of sectoral review was reflected earlier, when under dimension 1 it was observed that Ministry of Finance representatives are often not in attendance at JSRs.

An area of particular weakness is that of presenting evidence on whether previous JSR recommendations are followed up on during the course of the year (3D),¹⁶ with half of the JSRs with available information not considering this during the review. This poses an obstacle for JSRs to serve as a vehicle for accountability to the extent that this finding implies the failure to incorporate recommendations stemming from a JSR into the next policy cycle, translating into poor delegation of responsibility for following up on recommendations. Furthermore, the annual implementation report included a brief situational analysis of the period under consideration with key sector indicators and financial information in fewer than 60 percent of JSRs (3A). This is similarly problematic for accountability, given that a robust assessment of current context both provides important information to key stakeholders and identifies areas where assigned responsibilities have not been adequately fulfilled.

As a result of data quality issues, JSRs do not always meet donors' reporting requirements, except in the cases in which development partners channel their funding through budget support—typically where they tend to view countries' M&E systems as adequate to meet the accountability and reporting needs of their own constituencies (Holvoet and Inberg, 2014a). This means that, as reinforced by interview data from development partners, JSRs with quality evidence build confidence in government by allowing donors to assess the status of their investment. Furthermore, if JSRs are seen as providing adequate grounds for donors to conduct an evaluation, this curbs the need for parallel monitoring and reporting mechanisms, reducing costly and inefficient duplication.

Despite mixed views in interviews on the usefulness of field visits for a JSR, if properly facilitated these can provide additional evidence on sector implementation at regional levels as well as important insights into realities at the school and classroom levels. Less clear is the role of field visits in contributing to the accountability function of JSRs. In principle, government officials interviewed suggested that field visits could increase transparency and contribute to the evidence base in JSRs. However, the quality of data collected depends on a multitude of factors, including who planned the field visits to ensure a representative sample of schools is selected and how they are conducted, which in turn reflects what type of information is collected, by whom, and its reliability. Our commentary on this aspect of JSRs is limited by the fact that our interview protocol did not consistently ask participants about field visits, so we cannot draw firm conclusions here.

Dimension 4: Monitoring Tool

DEFINITION

Monitoring of sector plan implementation is the most frequently reported function of JSRs across all sources used in this study. In our literature review we presented a broader understanding of sector monitoring than is typically proposed: More than simple “education accounting” or review of sector

¹⁶ Dimension 4 analyzes whether JSR discussions are based on previous JSR recommendations; dimension 5 the attributes of JSR recommendations.

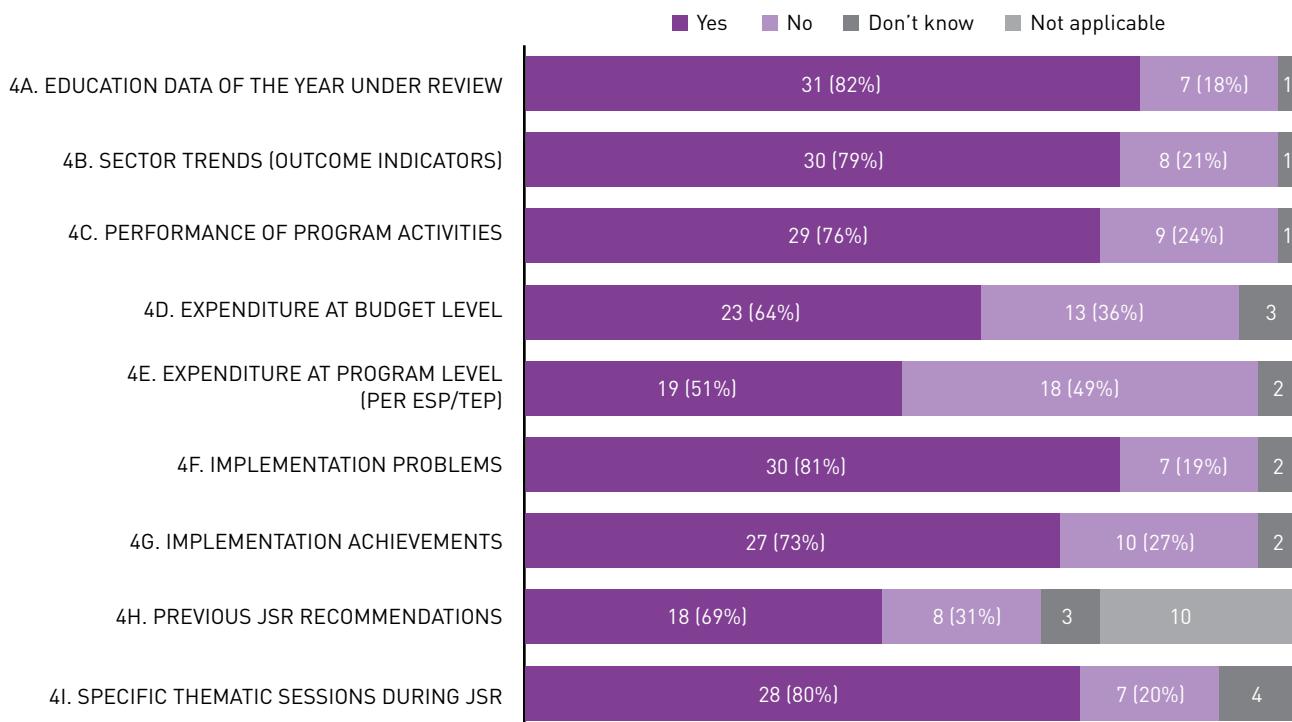
implementation, monitoring, we argue, should also encompass a learning function, where shortcomings in implementation are exploited for the lessons they offer about how to improve implementation moving forward. In our view, all stakeholders should contribute to this learning function of JSRs. JSRs also help to improve the M&E process by flagging urgent sector issues and deficits in data collection (Holvoet and Inberg, 2009).

ANALYSIS

Our variable analysis (Figure 4) shows a range of performance, with a large proportion of JSRs based on the discussion of significant items, including key outcome-level indicators that are relevant to identifying trends across the education sector (4B, 79 percent); implementation problems that call for a discussion on why targets are not achieved (4F, 81 percent); and the performance of core programs and activities, as measured by output-level indicators (4C, 76 percent). This suggests that the focus of JSR discussions is primarily organized around educational progress and results linked to implementation issues, but disconnected from the monitoring of corresponding financial issues. This echoes the analysis of dimension 3, which suggested a bias in favor of the presentation of activity-oriented evidence rather than evidence on expenditure. Only half of JSRs with available information discussed expenditure at the programmatic level (4E). Ideally, these aspects should be considered in tandem.

FIGURE 4 DIMENSION 4: MONITORING TOOL

Interview question: *Was the JSR based on the following items?*¹⁷



¹⁷ Due to a refinement of the questionnaire following the first round of data collection for the GPE results framework, questions 4B, 4G and 4K are only answered for JSRs conducted after July 2015. They are therefore omitted from this analysis.

This would allow examination of plan implementation with a dynamic optic that would most effectively identify remedial actions. This perspective, combining outcomes for activities with associated financial execution rates, would better support intra-sector dialogue around the appropriate (re)distribution of resources.

JSRs typically rely on at least some presentation of data for the year under review, but demonstrate a critical deficit with respect to discussion of previous JSR recommendations (4I). If JSRs are viewed as a formative learning process, this is of particular concern: As a reflective practice, reviewing past recommendations represents a potentially useful avenue for learning. Again reflecting the analysis of dimension 3, we see that in more than half of cases, there is no evidence that the review under consideration reflects on past JSR recommendations. This is especially pertinent in terms of accountability, as this may translate into limited follow-up on key weaknesses identified during the previous review.

Our findings suggest several ways in which the monitoring function of JSRs should be enhanced. First, JSRs should not be regarded as a one-time event, but instead as part of a continuous process that builds on year-round data collection, production, analysis, dissemination and use. All stakeholders should contribute to this process through the provision of data, strengthening of EMIS capacities, or production and review of necessary reports and studies. From this perspective, the participation of all stakeholders in the same monitoring tasks is essential. However, as highlighted under dimension 3, despite donor commitments to harmonize aid and align with government procedures, parallel reporting and monitoring exercises that are duplicative of JSRs persist (Hattori, 2009; Boak and Ndaruuhutse, 2011). As per the discussion of dimension 3, this is often attributed to quality concerns around data and analysis.

Second, learning is a recurring theme in our interviews with development partners, who highlight their aspirations for the JSR to shift away from a simple review, or a “show-and-tell” approach, where rote discussion of specific activities and associated expenditure dominates, and toward a more qualitative, strategic-level discussion. However, there is recognition that capacity issues limit the extent to which this is possible. This perspective is shared by government officials, who also articulate the value of the JSR as a learning opportunity focused on education results.

Finally, it is worth noting that monitoring in the context of JSRs results in the development of a mutually shared and agreed diagnosis of where the sector stands and how it is performing in the year under review. This is an important foundation for accountability, where everyone agrees both on the strengths and weaknesses of a system and on the basis on which remedial solutions should be reflected in partnership.

Dimension 5: Instrument for Change Embedded Effectively into a Policy Cycle

DEFINITION

We argue in our analytical framework that JSRs are an essential sector monitoring mechanism that should ultimately be exploited as a management and policy support tool for sector planning, programming, and budget preparation and execution. The JSR has potential to serve as a flexible and adaptive tool for operational planning. This stresses the importance of the process of formulating and following up on robust and actionable JSR recommendations, so JSRs can sustainably and effectively

feed into a dynamic policy cycle. If monitoring serves a backward-looking function of past sector implementation, embedding JSRs into a policy cycle serves a forward-looking function, informing strategic decision making for the next cycle. The JSR is a key lever for influencing the policy direction in a sector and planning for the shorter and medium terms. When embedded effectively into a policy cycle, JSRs can also support mutual accountability by designating responsible parties in charge of implementing and following up on JSR recommendations.

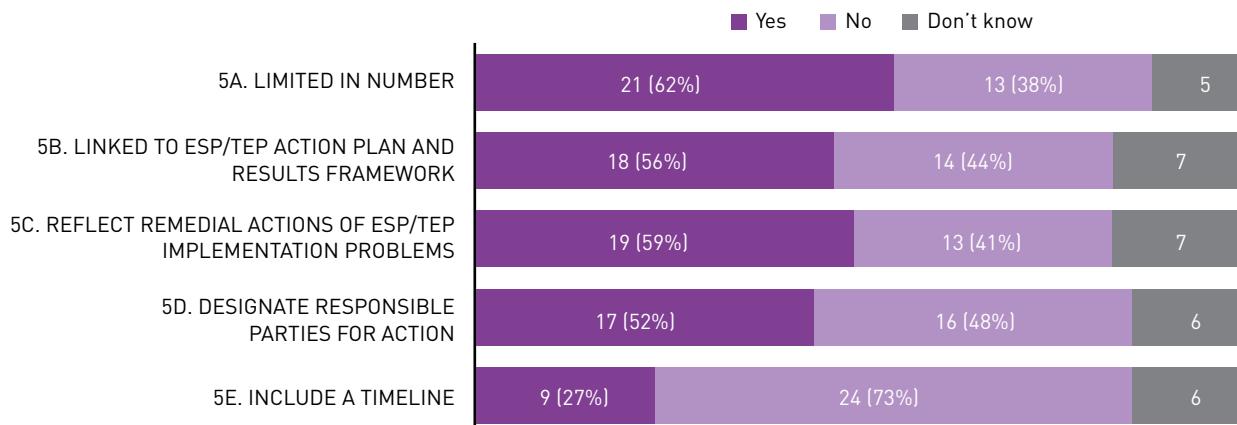
ANALYSIS

Figure 5 illustrates the results of the variable analysis for dimension 5. Interviewees typically agreed that JSR recommendations provide a record of JSR findings regarding priority issues identified and remedial actions proposed (5C). Recommendations also reinforce to which parties particular responsibilities are assigned (5D), and who should be held accountable for delivery at the subsequent JSR. This serves as a useful basis on which to consider revisions to the agreed policy framework (dimension 2), ideally in a virtuous circle of accountability (where greater accountability in monitoring reinforces greater accountability in planning and implementation), embedded into a flexible and adaptive learning cycle. However, as highlighted in the literature review (Section 2), linkages between JSR recommendations and *concrete policy changes* are weak, mainly because JSRs are poorly anchored into sector planning and budget exercises. *Timing* is crucial: Embedding JSRs into the appropriate phase of the policy cycle would allow relevant parties to more easily consider and translate recommendations into actions. Furthermore, recommendations are rarely prioritized, or even sufficiently actionable, within JSR documentation. As the literature review illustrates, there are also few clear mechanisms for follow-up, further hindering implementation of actions toward a forward-looking consensus.

This potentially poses serious challenges to the credibility of JSRs as a sustainable platform for generating mutual accountability. JSRs can only be considered sustainable when they are included in country-led and routine processes. If stakeholders do not see progress toward shared objectives, they are less likely to see the mechanism as a credible vehicle to influence policy, thereby calling into question its sustainability.

FIGURE 5 DIMENSION 5: INSTRUMENT FOR CHANGE EMBEDDED EFFECTIVELY INTO A POLICY CYCLE

Interview question: *What form do the JSR recommendations take?*



While “recommendation syndrome,” or an intractable proliferation of unfocused proposals in summary JSR reporting, is deplored throughout case study-based literature, in our variable analysis 62 percent of the JSRs examined do limit the number of recommendations. Furthermore, the absence of SMART recommendations is less pronounced in our sample, in which 51 percent of the JSRs with available information produced recommendations that detail roles and responsibilities (“who,” 5D), and 59 percent include specific remedial actions (“how,” 5C). However, only 27 percent included a timeline (5E), which arguably jeopardizes the likelihood of immediate operationalization. Our analysis sheds little light on how implementation of recommendations is operationalized, although interview data from government officials suggests that this group provides support for Hattori’s (2009) and Asuming-Brempong and Frimpong’s (2014) arguments that ideally follow-up occurs as soon as possible after completion of the review process.

The JSR process concludes with the drafting and validation of a JSR report, often referred to as a joint aide-mémoire. This typically records the shared sector assessment and remedial recommendations agreed upon among stakeholders (Penrose and Sackllokham, 2009). The JSR report is seen as the most tangible output of a JSR (Holvoet and Inberg, 2009), which, once validated and disseminated, holds education stakeholders accountable for jointly implementing a corrective course of action. Interviewees discussed the importance of having this report officially signed by high-ranking representatives of the various stakeholder groups in order to enforce its validity and hold the partnership to its mutual commitments. Packer (2006) notes that JSR reports are of central importance in an aid-driven environment, where the value of aid support is demonstrated by reflecting the overall progress of a recipient country toward achieving agreed-upon education goals. Interviewees cited the dissemination of the JSR report within and outside of the stakeholder group as a practice that reinforces transparency and contributes to building mutual trust.

While a JSR report is theoretically a joint document that, according to both the literature and our interviews, reflects agreement across the spectrum of stakeholders, the drafting of the report was largely discussed as a sensitive moment of political negotiation and tension, revealing divergent stakeholder interests and an extremely fragile balance of power that a mutual accountability mechanism would aim to equalize. Who among the partners should control the writing, selection of inputs and revision of the report appears to be an issue around which there is little consensus, and one that could ultimately compromise the mutual accountability compact. Both the literature and the interview respondents cite a number of instances in which it is donors rather than national governments who have drafted the aide-mémoire, or reflect on cases where particular stakeholders exerted pressure to insert what they considered to be important observations into the report in order to compensate for deficiencies in the earlier dialogue, or to reflect requirements demanded by their respective headquarters’ policies. This suggests that further analysis of the political economy of the JSR aide-mémoire drafting process could be fruitful when examining the intersection between JSRs and mutual accountability.

Discussion and Conclusion

This paper set out to explore the role of JSRs in supporting and developing mutual accountability. Our contribution is unique in its scope, combining variable-based data from 39 JSRs, and case-based data collected across multiple constituencies including interviews in five GPE partner countries. We selected the lens of mutual accountability, given both its prominence in GPE’s operating model and

the central role it plays in the development context, where the aid environment necessitates managing a complex web of interdependencies across key groups, including government, development partners and CSOs.

We examined the ways in which particular stakeholders view their contribution to support mutual accountability in JSRs. Our findings suggest that these groups do not necessarily share a common understanding of mutual accountability, nor does the widely accepted theoretical model articulated in international agreements apply in practice. In particular, we note that while the ways in which development partners talk about mutual accountability aligns closely with much of the development rhetoric and the notion of multidirectional channels across which responsibilities are equally distributed, governments observe a tension between shared responsibility and country ownership of (and therefore responsibility for implementation of) sector plans. CSOs often view themselves as not accountable to others but instead responsible for holding others to account, demanding answerability from governments in particular. Our significant theoretical contribution is the claim that this points to the need for an expanded view of responsibility in combination with a more nuanced understanding of how mutual accountability operates in practice. We propose a model in which governments are primarily responsible for the delivery of education sector results, but with support from development partners and CSOs: development partners play a key role in capacity building, while CSOs challenge the disconnect between policy and practice.

Recognizing that mutual accountability is best achieved where JSRs are effective, we then turned to consider how a sample of 39 JSRs conducted across GPE partner countries performs when evaluated against five key drivers of JSR effectiveness. We note that while the JSR mechanism is also a product of the aid effectiveness agenda, the value of this platform as a tool for developing stronger education systems, informed by quality sector dialogue, is contingent on its integration into a country-led policy cycle to support and improve sector planning and implementation. Notably, while some of the stakeholders interviewed for this research pointed to shortcomings in the current operationalization of JSRs, all attributed value to its role in improving sector outcomes in various forms: from providing a forum for the voice of those who might otherwise be marginalized in policy processes, to developing capacity for better sector monitoring and improving accountability. This speaks to the importance of assessing current performance to better understand how to strengthen the contributions JSRs make within the policy cycle.

Our research highlights that a key challenge to using JSRs as a tool for supporting a stronger sector policy cycle (dimension 5) is the relative disconnect between the content of an agreed policy sector framework that defines the perimeter of JSR discussions and provides the framework for monitoring (typically the ESP) and what is actually reviewed at JSRs. This reflects in part the poor quality (or even complete absence) of a subset of reporting and planning instruments, in particular annual implementation reports based on annual action or operational plans. In the absence of evidence-based reporting that is disseminated across all stakeholders, the quality of JSR discussions is inevitably compromised (dimension 3). It also underscores the reality that JSR discussions often neglect to give due attention to *expenditures* within the sector, favoring a focus on the implementation of *activities* against planned performance. It is hard to know whether the absence of Ministry of Finance representatives in many JSRs (dimension 1) is a product or the cause of this result. Whatever the underlying explanation, this weakens the capacity of the JSR to contribute to better budget monitoring and programming in order to address identified sector issues (dimension 3), and limits the extent to which JSRs function as a learning exercise to identify remedial actions where shortcomings in implementation are observed (dimension 4). In addition, it is troubling that JSRs often fail to monitor past JSR

recommendations (dimension 4) and that JSR recommendations do not translate into implementable policy, or operational, planning and budgeting decisions (dimension 5). Our research suggests that the problem lies less in recommendations being poorly formulated or cumbersome, but is instead attributable to the absence of levers to ensure JSR recommendations are appropriately followed up on. This poses a serious challenge to the mutual accountability function. The identification of these weaknesses offers a significant silver lining, however, in that, as noted, there are clear and tractable solutions for policymakers and practitioners to address a number of these challenges.

This paper also provides a springboard for future research into JSRs. Consideration of complementary sources of information, direct observations in particular, would allow a richer analysis of additional dimensions in the JSR effectiveness framework, including the role of national leadership in driving JSR performance, ownership and partner buy-in. Furthermore, the mutual accountability lens adopted here could be complemented by consideration of alternative but equally important JSR functions, such as improving cross-stakeholder coordination and developing national capacities.

Future research could also consider a more granular and qualitative analysis of the role of context for testing the robustness of the five dimensions and identifying the drivers that foster, or impede, the effectiveness of JSRs across different environments. Just as the current research undertaking resulted in refinements to the original GPE normative model of JSR effectiveness, such research would permit further enhancements. It is our belief that devising a rigorous JSR effectiveness model would be valuable both to policymakers responsible for the organization of JSRs and to those stakeholders who aim to support the strengthening of this mechanism.

Bibliography

- *Asuming-Brempong, S. and S. Frimpong. 2014. *Joint Sector Review Assessment, Ghana, Advancing Mutual Accountability through Comprehensive, Inclusive, and Technically Robust Review and Dialogue*. Accra: Government of Ghana, Ministry of Food and Agriculture.
- *Bartholomew, A., T. Takala and Z. Ahmed. 2010. *Avaliação a Meio Percurso da Iniciativa ‘Fast Track’ da Ept, Estudo Nacional: Moçambique*. n.p.: Cambridge: Cambridge Education, Mokoro and Oxford Policy Management.
- *Boak, E. and S. Ndaruhutse. 2011. *The Impact of Sector-Wide Approaches: Where From, Where Now and Where To*. Berkshire: CfBT Education Trust.
- *Buchert, L. 2002. “Towards New Partnerships in Sector-Wide Approaches: Comparative Experiences from Burkina Faso, Ghana and Mozambique.” *International Journal of Educational Development* 22(1): 69–84.
- *Busan Partnership. 2011. Busan Partnership for Effective Development Cooperation. Fourth High-Level Forum on Aid Effectiveness, Busan, Republic of Korea, 29 November to 1 December 2011.
- *Cassels, A. 1997. *A Guide to Sector-Wide Approaches for Health Development: Concepts, Issues and Working Arrangements*. Geneva: World Health Organization.
- *Chirwa, C. 2012. “Partnership for Education in Malawi: Power and Dynamics within the Education Sector Wide Approach.” PhD diss, University of Sussex.
- *Cruz, V. and B. McPake. 2010. “The ‘Aid Contract’ and Its Compensation Scheme: A Case Study of the Performance of the Ugandan Health Sector.” *Social Science & Medicine* 71(7): 1357–1365.
- *Davies, P. 2013. *Evaluation of Samoa Health Sector Management Programme (Health SWAp)*. Canberra: AusAID, Health Resource Facility.
- *de Kemp, A., J. Faust and S. Leiderer. 2011. *Between High Expectations and Reality: An Evaluation of Budget Support in Zambia (2005–2010). Synthesis Report*. The Hague: Ministry of Foreign Affairs of the Netherlands.
- Domingo, P., L. Wild, A. Hudson and C. Wathne. 2009. *Domestic and Mutual Accountability for Aid: Building Stronger Synergies. Literature Review and Conceptual Framework*. London: Overseas Development Institute.
- Droop, J., P. Isenman and B. Mlalazi. 2008. *Paris Declaration on Aid Effectiveness: Study of Existing Mechanisms to Promote Mutual Accountability (MA) between Donors and Partner Countries at the International Level. Final Report*. London: Oxford Policy Management.

* Indicates article included in literature review (Section 2) and for a frequency analysis on the type of literature, theme, sector, methodology, instruments, language and country cases in these articles (see details in Annex 1).

- *Foster, M. 2000. *New Approaches to Development Co-Operation: What Can We Learn from Experience with Implementing Sector Wide Approaches?* London: Overseas Development Institute.
- *Foster, M., A. Brown, A. Norton and F. Naschold. 2000. *The Status of Sector Wide Approaches*. London: Overseas Development Institute.
- *Foster, M., A. Brown and T. Conway. n.d. *Sector-Wide Approaches for Health Development: A Review of Experience*. London: Overseas Development Institute.
- *Gerster, R. and M. Chikwekwe. 2007. *Poverty Reduction Budget Support (PRBS) in Zambia Joint Annual Review 2007: Learning Assessment. Final Report*. n.p.: commissioned by the Government of the Republic of Zambia and the PRBS cooperating partners.
- *Global Partnership for Education (GPE). 2012a. *An Analysis of Joint Education Sector Review Reports*. Unpublished.
- *———. 2012b. *Making Education Aid More Effective: Monitoring Exercise on Aid Effectiveness in the Education Sector. Synthesis Report*. Washington, D.C.: Global Partnership for Education.
- *Global Partnership for Effective Development Co-operation. 2015. *Monitoring Guide 2015–2016: Monitoring the Effective Development Co-Operation Commitments*. n.p.: Global Partnership for Effective Development Co-operation.
- *Handley, G. 2009. *Sector Budget Support in Practice: Literature Review*. London: Overseas Development Institute.
- *Harding, A. and R. Gerster. 2004. *Learning Assessment of Joint Review 2004 Final Report. Report to the Government of Mozambique and Programme Aid Partners*. n.p.: Commissioned by the government of Mozambique and the G15 programme aid partners with financial support of the UK's department for international development and the Swiss State Secretariat for Economic Affairs.
- *Harrold, P. and Associates. 1995. *The Broad Sector Approach to Investment Lending: Sector Investment Programs*. World Bank Discussion Papers 302. Africa Technical Department Series. Washington, D.C.: World Bank.
- *Hattori, H. 2009. “Enhancing Aid Effectiveness in Education through a Sector-Wide Approach in Cambodia.” *Prospects* 39(2): 185–199.
- *Holvoet, N. and H. Rombouts. 2008. “The Challenge of Monitoring and Evaluation under the New Aid Modalities: Experiences from Rwanda.” *The Journal of Modern African Studies* 46(04): 577–602.
- *Holvoet, N. and L. Inberg. 2009. “Joint Sector Reviews: M&E Experiments in an Era of Changing Aid Modalities: Experiences from JSRs in the Education Sectors of Burkina Faso, Mali and Niger.” *Public Administration and Development* 29(3): 204–217.
- *———. 2014a. “Diagnostic Review of the Monitoring and Evaluation System of Uganda’s Education Sector: Selected Findings and Discussion.” *Journal of Education and Training* 2(1): 134–154.

- *———. 2014b. “Taking Stock of Monitoring and Evaluation Systems in the Health Sector: Findings from Rwanda and Uganda.” *Health Policy and Planning* 29(4): 506–516.
- *Institute for Health Sector Development. 2003. *Sector-Wide Approaches in Education*. London: Institute for Health Sector Development.
- *International Health Partnership (IHP). 2011. *Monitoring, Evaluation and Review of National Health Strategies. A Country-Led Platform for Information and Accountability*. Geneva: World Health Organization.
- *———. 2013b. *Joint Assessment of National Strategies: A Review of Stakeholders’ Needs*. Geneva: World Health Organization.
- *———. 2013c. *How to Conduct a Joint Assessment of a National Health Strategy (JANS) Based on Country Experience. A Working Document on Options for Conducting a JANS on a Health Sector or Sub-Sector Strategy*. Geneva: World Health Organization.
- *———. 2014a. *Progress in the International Health Partnership and Related Initiatives: 2014 Performance Report*. n.p.: International Health Partnership.
- *———. 2013a. *Joint Annual Health Sector Reviews: A Review of Experience*. n.p.: International Health Partnership.
- *———. 2014b. *Joint Annual Health Sector Reviews: Why and How to Organize Them*. Geneva: World Health Organization.
- International Monetary Fund (IMF), Organisation for Economic Co-operation and Development (OECD), United Nations and World Bank. 2000. *A Better World for All: Progress toward the International Development Goals*. Washington, D.C.: IMF.
- *Martínez, J. 2006. *Implementing a Sector Wide Approach in Health: The Case of Mozambique*. London: HLSP Institute.
- *Martínez, J., R. Dodd and V. Walford. 2013. *Do Health Sector Reviews Deliver Results?* London: HLSP Institute.
- *Mirzoev, T., A. Green and J. Newell. 2010. “Health SWAps and External Aid: A Case Study from Tajikistan.” *The International Journal of Health Planning and Management* 25(3): 270–286.
- Mundy, K. 2007. “Education for All: Paradoxes and Prospects of a Global Promise.” In *Education for All: Global Promises, National Challenges*, edited by D. P. Baker and A. W. Wiseman, 1–30. In Volume 8, International Perspectives on Education and Society Series. Oxford: Elsevier.
- . 2008. *Civil Society and Its Role in the Achievement and Governance of ‘Education for All’*. Background paper prepared for the *Education for All Global Monitoring Report*.
- Organisation for Economic Co-operation and Development (OECD). 1996. *Shaping the 21st Century: The Contribution of Development Cooperation*. Paris: OECD Publishing.

*———. 2005, 2008. *The Paris Declaration on Aid Effectiveness and the Accra Agenda for Action*. Paris: Organisation for Economic Co-operation and Development.

*Organisation for Economic Co-operation and Development (OECD) and the United Nations Development Programme (UNDP). 2014. *Making Development Co-operation More Effective: 2014 Progress Report*. Paris: OECD Publishing.

*Orpazó, J. 2008. *Documento de Referencia Encargado para el Informe de Seguimiento de la Educación para Todos en el Mundo 2009 Superar la Desigualdad: Por qué Es Importante la Gobernanza. Estudio de Caso: Honduras*. Paris: United Nations Educational Scientific and Cultural Organization.

*Packer, S. 2006. *Joint Monitoring Review Missions in the Education Sector*. Background paper prepared for the *Education for All Global Monitoring Report*. 2007/ED/EFA/MRT/PI/37. Paris: United Nations Educational Scientific and Cultural Organization.

*Penny, A., M. Ward, T. Read and H. Bines. 2008. “Education Sector Reform: The Ugandan Experience.” *International Journal of Educational Development* 28: 268–285.

*Penrose, T. and K. Sacklokham. 2009. *Education Sector Development Framework Performance Assessment Framework and Preparation for Implementation Plan. Lao PDR: Sector-Wide Approach in Education Sector Development*. ADB advisory technical assistance TA 4907-LAO.

*Riddell, A. 2007a. *Education Sector-Wide Approaches (SWAps): Background, Guide and Lessons*. Education Strategies and Field Support 12. Paris: United Nations Educational Scientific and Cultural Organization.

*———. 2007b. *The New Modalities of Aid to Education: The View from within Some Recipient Countries*. Background paper prepared for the *Education for All Global Monitoring Report 2008 Education for All by 2015: Will We Make It?* 2008/ED/EFA/MRT/PI/59 REV.

*Smith, G. 2009. *Sector Budget Support in Practice: Tanzania Health Desk Study*. London and Oxford: Overseas Development Institute and Mokoro.

Steer, L. and C. Wathne. 2009. *Mutual Accountability at Country Level: Emerging Good Practice. Background Note*. London: Overseas Development Institute.

*Ulikpan, A., I. Narula, A. Malik and P. Hill. 2014. “‘In the Driver’s Seat’: The Health Sector Strategic Master Plan as an Instrument for Aid Coordination in Mongolia.” *Globalization and Health* 10(1): 1.

*United Nations Development Programme (UNDP), United Nations Population Fund, United Nations High Commissionate for Refugees, United Nations Children’s Fund, UNWOMEN and World Bank. 2015. *Education 2030: Incheon Declaration and Framework for Action. Towards Inclusive and Equitable Quality Education and Lifelong Learning for All*.

United Nations General Assembly. 2000. United Nations Millennium Declaration. Resolution adopted by the General Assembly A/RES/55/2. New York City: United Nations General Assembly.

*Villancourt, D. 2009. *Do Health Sector Wide Approaches Achieve Results? Emerging Evidence and Lessons from Six Countries*. IEG Working Paper 2009/4. Washington, D.C.: World Bank.

*West, R. C. 2004. *Education and Sector-Wide Approaches (SWAp) in Namibia*. Paris: United Nations Educational Scientific and Cultural Organization, International Institute for Educational Planning.

*Williamson, T. and C. Dom. 2010. *L'Appui Budgétaire Sectoriel dans la Pratique. Rapport de Synthèse*. London and Oxford: Overseas Development Institute and Mokoro.

*World Bank. 2001. *Education and Health in Sub-Saharan Africa: A Review of Sector-Wide Approaches*. Washington, D.C.: World Bank.

ANNEX 1: Methodological Annex

JSR Variable Analysis

This research included an analysis of the JSRs that were conducted between July 2014 and December 2015 in GPE partner countries, based on a questionnaire that identifies five key JSR dimensions relevant for effectiveness set out in a number of variables (“variable analysis”). While JSRs are typically conducted at the country level, in some instances they may be conducted within a different geographic zone (e.g., under federal systems, this may happen at the state or provincial level). In our analysis, the term *state* is used as a generic descriptor for the geographic zone in which a JSR was held; this can mean a semi-autonomous regional unit, a state, a province or a country.

The analysis was conducted via a desk-based review of key JSR documents by GPE country leads, with answers recorded in a template specifically designed for this purpose. Direct observations of JSRs were not included in this analysis. The rationale for conducting a desk review was to ensure that the assessment of the effectiveness of JSRs would be standardized and based strictly on objectively verifiable evidence. This means that some of the questions asked to assess the dimensions of JSRs were limited to proxies that were intended to capture the underlying concepts behind the relevant attributes.

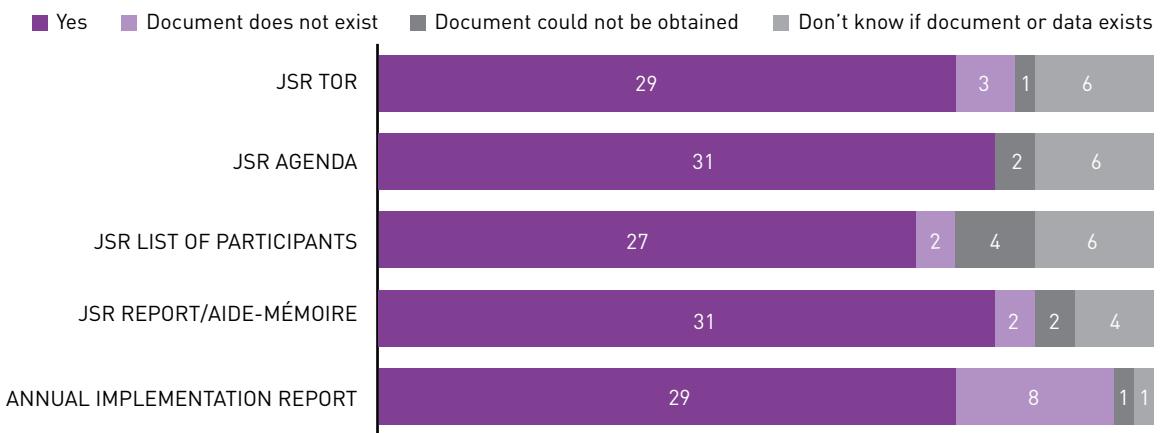
The five key dimensions of JSRs in the questionnaire are that they are inclusive and participatory; aligned to shared policy frameworks; based on evidence; a monitoring tool; and an instrument of change embedded effectively into the policy cycle. Each dimension was translated into one question comprising sub-questions, or items that together described the dimension being explored, per the JSR Questionnaire in Annex 2. GPE country leads responded to one questionnaire per JSR, after obtaining and reviewing the necessary JSR documentation, including JSR terms of reference, agendas, participant lists, annual implementation reports and JSR reports or aide-mémoires.

The desk-based nature of this analysis introduces a potential confounding variable, namely that in some cases it was not possible to determine whether the key characteristics within a JSR dimension were covered or not, based on the availability of core documents needed for the analysis. These dimensions were coded as *inconclusive*, and should be clearly distinguished from the key characteristics in a dimension coded as *not covered*. An inconclusive dimension suggests a lack of evidence, meaning that documentation was poor. Figure 6 highlights the documents used to conduct the desk-based analysis. On average, about a quarter of the JSRs were missing key pieces of information, either because documents were not produced or were not readily available to the country leads.

The sample comprised 51 JSRs in total, of which 12 were conducted in the same state in successive years. A selected sample of 39 JSRs (in 36 countries)¹⁸ was used for the variable analysis, in which only one JSR per state was included, so as to ensure that all states received equal weighting, as shown in Table 3. For the 12 states that conducted two JSRs over the period under consideration, the JSR with the fewest inconclusive dimensions in the JSR questionnaire was retained. If both JSRs had an equal number of inconclusive dimensions, the JSR with the fewest inconclusive items was retained. If the JSRs also had the same number of inconclusive items, the most recent JSR was retained.

After country leads responded to this questionnaire, the information was validated to verify that there were no inconsistencies within and across responses to questionnaires. A consolidated database

¹⁸ Pakistan is represented twice (Balochistan and Sindh provinces), while Somalia is represented three times (the regions of Somaliland and Puntland, as well as at the federal level).

FIGURE 6 DOCUMENTS USED IN DESK-BASED REVIEW

recording this information was constructed to allow analysis. Data treatment included the following steps:

1. Categorization of all response items (item-level coding). A subset of items was identified as the key items describing a dimension, i.e., as necessary in order to ensure that the JSR demonstrated a particular attribute (these are in boldface type in the questionnaire, in Annex 2). The additional items describe aspects of the dimension beyond the core aspects, and were treated as enhancements or “good practice” items. Depending on the possible answers to each dimension, the questionnaire required country leads to assess whether the items in each question were *covered*, *not covered* or *only covered partially*, or if the *available information was insufficient or not applicable*.
2. Analysis of a JSR dimension (question-level coding), based on the answers for all items associated with each question or dimension, followed a four-value categorical scale, defined as shown in Table 4.
3. This data was used to produce summary statistics for analysis, first counting the number of dimensions with key characteristics covered, not covered, and inconclusive per JSR. Next, the aggregated number of JSRs per dimension within these categories was counted, as reported in this paper. We also explored correlations across various JSR rounds, dimensions and items; however, we found little value in reporting these in the main body of this paper. Not only did the small sample size limit their value as statistically significant observations, but there also were few strong correlations observed. Similarly, an examination of the relationship between the questionnaire variables and external variables (including key sector indicators, such as primary completion rate, and aspects of aid dependency, such as relative volume of external aid and number of active donors) did not prove fruitful, and so findings are not reported in this paper. This remains an area for future research.

Literature Review

Articles were collected during January and February 2016, mainly from Google Scholar, and thereafter from the bibliographic references of these articles, until saturation was reached (i.e., until the articles were being identified repeatedly). The search was based on the following 10 keywords, in an attempt to capture variations on the JSR concept: *joint sector review*, *joint annual review*, *annual joint review*, *annual meeting of progress review*, *joint review meeting*, *joint monitoring*, *joint evaluation*, *joint review*

TABLE 3 JSRs INCLUDED IN SAMPLE

JSRs in Round 1 (Jul 2014–Jun 2015)

*Benin	*Madagascar	Somalia Federal
*Burkina Faso	*Malawi	*Somalia Puntland
*Burundi	*Mali	*Somalia Somaliland
*Cambodia	*Mauritania	*South Sudan
*Cameroon	*Mozambique	*Sudan
*Chad	Nepal	*Tanzania Mainland
*Comoros	*Nicaragua	*Vietnam
*Djibouti	*Pakistan-Balochistan	*Yemen
*Democratic Republic of the Congo	*Pakistan Sindh	*Zambia
*The Gambia	*Rwanda	Zimbabwe
Ghana	*Senegal	

JSRs in Round 2 (Jul-Dec 2015)

Democratic Republic of the Congo	Mauritania	Sudan
The Gambia	*Nepal	*Tajikistan
*Ghana	Rwanda	Tanzania Mainland
*Guinea	*Sierra Leone	*Togo
*Lao People's Democratic Republic	*Somalia Federal	*Uganda
*Liberia	South Sudan	*Zimbabwe
Malawi		

* JSRs included in the discussion of the selected sample of 39 JSRs.

mission, sector monitoring and sector implementation monitoring. These keywords were searched in English, with the words *education sector, health sector* and *agriculture sector* added, in order to collect articles from various sectors. The keyword searches were also done in French, Spanish and Portuguese, languages that are used for official reporting in the majority of developing countries. We also did a search using the country-specific names of education joint sector reviews, in their original languages.¹⁹ The titles,

19 It is worth noting that out of the 29 different country names identified in four languages, 25 share the mention of a "joint" "sector" "review."

TABLE 4 QUESTION-LEVEL CODING

Category 0	JSR does not cover the key characteristics in a particular dimension
Category 1	JSR covers the key characteristics in a particular dimension
Category 2	JSR covers more than the key characteristics in a particular dimension
Category 98	Information obtained is insufficient; the JSR in a particular dimension is then categorized as “inconclusive”

TABLE 5 THEME AND SECTOR IDENTIFIED IN ARTICLES

Theme	Articles by Sector					Articles by Theme
	Education	Health	Agriculture	Multi-Sector	None	
Joint Sector Reviews	4	4	2			10 (20%)
Sector-Wide Approaches	5	8		4		17 (35%)
Aid Effectiveness	4	2			4	10 (20%)
Monitoring and Evaluation	2	4				6 (12%)
Sector Budget Support		1		3		4 (8%)
Donor-Government Reviews					2	2 (4%)
	15 (31%)	19 (39%)	2 (4%)	7 (14%)	6 (12%)	49 (100%)

abstracts and lists of content for the first 60 articles of each keyword search were screened, in order to determine whether or not they discussed JSRs. When in doubt, the text was briefly reviewed. Any articles that discussed JSRs in a central or peripheral way were retained for a full reading; the others were discarded.

The literature review ultimately included 49 articles. As Table 5 shows, 10 articles (20 percent) primarily discuss JSRs; four of them are in the education sector, four in the health sector and two in the agricultural sector. The rest of the articles discuss JSRs only in a peripheral way. Approximately a third of the articles (17, or 35 percent) focus on SWAps, 10 of them (20 percent) on aid effectiveness and the rest on monitoring and evaluation (12 percent), sector budget support (8 percent) and donor-government reviews (4 percent). Almost a third of the articles refer to the education sector

(31 percent), and slightly more than a third refer to the health sector (39 percent); only 4 percent refer to the agricultural sector, and the rest refer to no sector or to multiple sectors.

Figure 7 reveals that most of the existing knowledge on JSRs comes from gray literature (82 percent), which is produced outside of traditional commercial or academic publishers. In this case, most of the gray literature referred to publications by international organizations. Only 16 percent of the articles came from peer-reviewed journals. Figure 8 shows that the preferred methodology in these studies was single-case studies (35 percent), followed by multiple-case studies (33 percent); the rest were manual-type documents, assessment tools, international declarations, discussion papers or literature reviews. Figure 9 demonstrates that the preferred instrument for data collection and analysis was a documentary review (37 percent), followed by interviews (29 percent) and field observations (13 percent). Few of the articles used surveys or international meetings (8 percent), and in 12 percent of cases there was no mention of the research instrument used.

A total of 57 different country cases were studied in the articles; on average, five countries were discussed per article. However, as Table 6 shows, some of the same country cases were discussed in various articles: Cambodia, Bangladesh, Ghana, Mozambique, Uganda and Vietnam were the most frequently discussed, either in the eight articles focusing primarily on JSRs or in the full sample of 49 articles. Figure 10 shows that the vast majority of articles were published in English (88 percent); there was one article published in French, another in Spanish and another in Portuguese. Three of the articles—the international declarations—were published in multiple languages.

FIGURE 7 LITERATURE TYPE

Number of articles (% of 49)

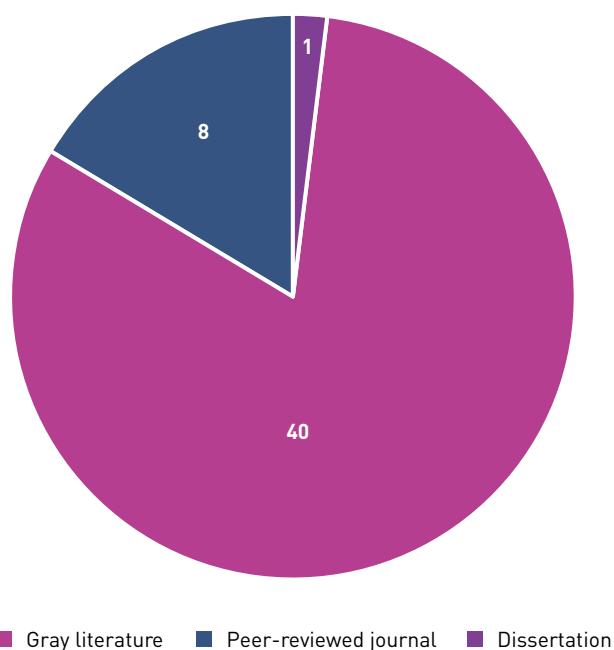
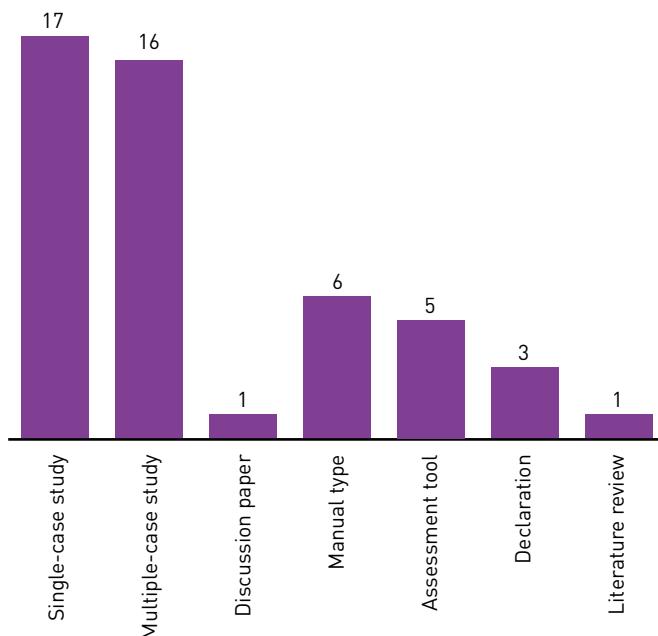


FIGURE 8 METHODOLOGY

Number of articles (% of 49)

**FIGURE 9 INSTRUMENT USED**

Number of articles (% of 49)

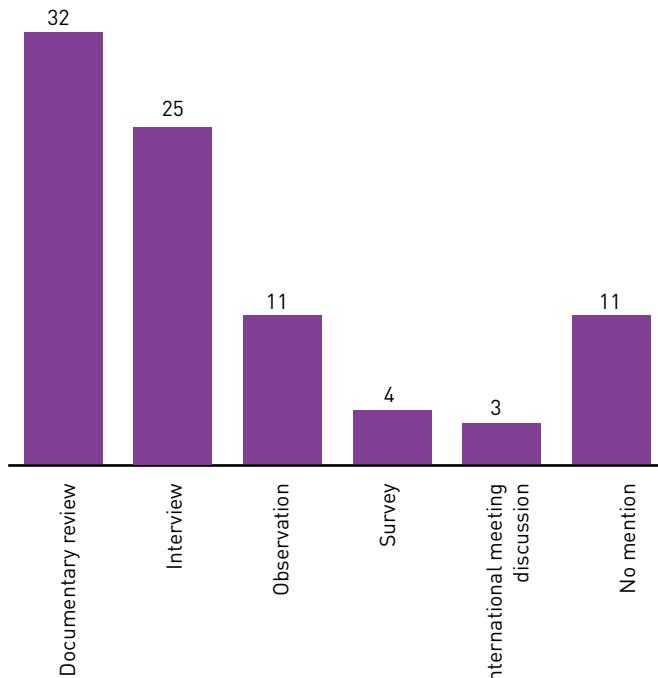
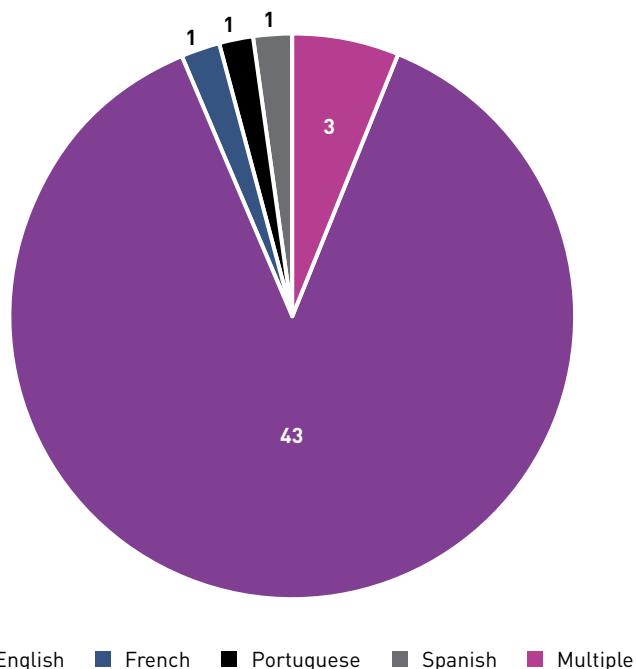


TABLE 6 COUNTRY CASES MOST FREQUENTLY COVERED IN THE ARTICLES

	In All 49 Articles	In the 8 Articles Focused on JSRs
Cambodia	10	5
Bangladesh	6	4
Ghana	9	4
Mozambique	13	3
Uganda	9	3
Vietnam	9	3
Ethiopia	8	
Rwanda	5	
Burkina Faso	5	
Nepal	5	

FIGURE 10 DISTRIBUTION OF ARTICLES BY LANGUAGE

(number of articles indicated)



Interviews with Key Country Stakeholders

As part of GPE's larger efforts to investigate JSR effectiveness, 32 semi-structured interviews with open-ended questions were conducted. Interviews were conducted in five countries—Bangladesh, Burkina Faso, Guinea, The Gambia and Zambia—during Secretariat missions between March and June 2016 where JSR attendance was highlighted as a core mission objective. The interview protocol asked about the experiences of the interviewees; their roles in their respective organizations; their involvement in the JSR they were attending; their understanding of a JSR, and its functions or purposes; their opinions on an ideal JSR; and the areas of success and challenges that were identified in the JSR they were attending.

Interviewees were purposely selected. Among the 30 country-level stakeholders, 11 represented various services within ministries of education at the national and regional levels; nine represented bilateral or multilateral donor organizations; eight represented CSOs and education CSO coalitions; one represented a teachers union; and one was a member of parliament. Two additional interviews were conducted with the country leads at the Secretariat who were responsible for two of the countries studied. However, the current paper reports only on the country-level stakeholders, in a first attempt to conceptualize JSRs from the perspective of the stakeholders who organize, attend and follow up on JSRs, and who are accountable within the sector. Table 7 shows the list of interviews conducted, by country and by stakeholder category.

Survey on JSR Function Preference

In September 2016, an online survey was administered to 28 developing country partners by GPE to collect information on perceptions around the roles and responsibilities of GPE actors. The question for this research was the following: *From your perspective, what do you expect from a well-functioning JSR? Please select one to three items from the list [provided below in Table 8] that best reflect the areas you consider to be the most important.*

Seventy respondents from the following stakeholder categories replied to the survey: government officials (10 respondents representing 14 percent of sample), mostly from ministries of education, but one from a Ministry of Social Affairs; donors from bilateral and multilateral agencies (40 or 57 percent); nongovernmental organizations (16 or 23 percent); teachers unions (2 or 3 percent); one private sector participant (1 percent); and one participant from an academic institution (1 percent). Since respondents selected between one and three responses, the number of responses per respondent was variable, producing 203 responses in total (only five respondents selected fewer than three choices).

Table 8 summarizes overall stakeholder preference for the suggested JSR functions, identified across four stakeholder categories (government, donor, NGO and teachers union). As the table shows, the function of ownership over plan implementation was ranked highest by stakeholders (35 percent responses), followed by the accountability function, which includes enhanced government accountability toward donors and citizens as well as mutual accountability (21 percent). The third rank was equally allocated to three functions: improving stakeholder capacity for sector planning and monitoring; harmonization or coordination; and improving transparency of resource allocation and execution (10 percent each).

TABLE 7 INTERVIEW LIST

Country	Stakeholder Category	Number of Interviews
Bangladesh	CSO/NGO	1
Bangladesh	Donor/Technical Agency	3
Bangladesh	Government	2
Burkina Faso	CSO/NGO	3
Burkina Faso	Donor/Technical Agency	2
Burkina Faso	GPE Country Lead	1
Burkina Faso	Government	4
The Gambia	CSO/NGO	4
The Gambia	Donor/Technical Agency	2
The Gambia	Government	5
Guinea	Donor/Technical Agency	1
Guinea	GPE Country Lead	1
Zambia	CSO/NGO	1
Zambia	Donor/Technical Agency	1
Zambia	Government	1

Additional Data

As part of the larger JSR effectiveness research, some additional data was collected and informed our thinking in producing this research, although the data was not analyzed in depth for this paper. This includes observations collected between March and June 2016 at the five JSRs where GPE's Secretariat conducted interviews, as above (Bangladesh, Burkina Faso, Guinea, The Gambia and Zambia). The objective was to gather information on the dynamics among participants and identify key characteristics of JSRs, including the nature of participation; organization and logistics; discussion dynamics; content; and resources used for presentations, among others. Also, two focus groups were conducted in July and September 2016 with GPE country leads and education specialists who have experience in JSRs. The aim was to collect their opinions on JSR purposes; factors for success and limitations; and areas for capacity building related to JSR efficacy, from the perspective of GPE. To complement the GPE vision on JSRs, four semi-structured interviews were conducted with the Secretariat Leadership Team.

TABLE 8 JSR FUNCTION PREFERENCE

Weighted stakeholder responses (as a % of total)

Function	%
Increased stakeholder ownership over plan implementation	35
Accountability functions:	21
<i>Enhanced government accountability toward donor</i>	3
<i>Enhanced government accountability toward its citizens</i>	9
<i>Enhanced mutual accountability</i>	9
Improved stakeholder capacity in sector planning and monitoring	10
Better coordination and harmonization of stakeholder efforts and interventions	10
Improved transparency with regard to resource allocation and execution (both internal and external funding)	10
Strengthened leadership in educational development	5
Mobilization of additional financial resources	2
Learning from every actor involved in sector development	2
Better alignment of interventions with country systems	1
Reduced transaction costs via joint reporting and monitoring, and information sharing	1
Improved national capacity on data analysis and reporting	1
	100

Notes: Given that a larger proportion of donors and fewer NGOs and teachers unions responded to the survey when compared to government respondents, this table is based on weighting the responses of each of the four stakeholder categories (government, donor, NGO and teachers union) equally; each category therefore represents the equivalent of 25 percent of respondents.

ANNEX 2: JSR Questionnaire Used for Variable Analysis

Dimension	Question	Possible Data Source(s)	Response Items
General Information	Please answer to the following items	<input checked="" type="radio"/> ESP/TEP <input checked="" type="radio"/> JSR agenda	<p>Beginning date of the JSR (mm/dd/yyyy):</p> <p>Ending date of JSR (mm/dd/yyyy):</p> <p>First education JSR of ESP/TEP implementation period?</p> <p>Beginning year of ESP/TEP period covered (yyyy):</p> <p>Ending year of ESP/TEP period covered (yyyy):</p> <p><i>Sub-sectors covered in the ESP/TEP:</i></p> <ul style="list-style-type: none"> Early childhood education Primary education Lower secondary education Upper secondary education TVET Higher education Nonformal education and adult literacy <p>ESP/TEP is sector-wide</p>
Control Question	What documents among those listed did you use to answer the questions in this assessment?		<p>JSR terms of reference</p> <p>JSR agenda</p> <p>JSR list of participants. If yes, <u>please indicate below</u> if this was the tentative or actual list:</p> <ul style="list-style-type: none"> Tentative list of participants Actual list of participants <p>JSR report/aide-mémoire (this includes any document produced as an output of the JSR)</p> <p>Previous JSR report/aide-mémoire (this includes any document produced as an output of the JSR)</p> <p>Education sector plan/transitional education plan</p> <p>Annual action plan of the ESP/TEP</p>

Dimension	Question	Possible Data Source(s)	Response Items
			<p>Annual implementation report (any document that tracks comprehensively the education sector progress would qualify as an equivalent of the annual implementation report, and can be used for this assessment)</p> <p>Additional document(s) used for this assessment. <i>Please list.</i></p>
1. Inclusive and participatory	<p>To what extent was the JSR inclusive? From which of the following categories was there participation during the JSR?</p>	<ul style="list-style-type: none"> ● List of JSR participants ● JSR terms of reference 	<p>Ministry/ministries in charge of education at the national level. If yes, please indicate below which of these ministries participated:</p> <ul style="list-style-type: none"> Basic education TVET Higher education Social affairs/gender <p>Any other ministry(ies) related to education at the national level that participated. <i>Please list.</i></p> <p>Ministry of Education at the regional/district level and/or regional authorities in charge of education</p> <p>Ministry of Finance and/or budget</p> <p>Other ministries: planning, health, civil service, etc.</p> <p>Development partners: multilateral, bilateral agencies</p> <p>International CSOs</p> <p>Local CSOs</p> <p>Parents associations</p> <p>Teachers unions</p> <p>Universities/academic institutions</p> <p>Private sector: companies funding education, foundations</p>

(continued)

Dimension	Question	Possible Data Source(s)	Response Items
			<p>Nonpublic educational providers (private school networks)</p> <p>Any other category(ies) of participants that may be missing from the listing above. <u>Please list.</u></p>
2. Aligned to shared policy frameworks	<p>Did the JSR cover the full spectrum of the education sector? Which of the following items apply? These can be discussed during the JSR (as per in the JSR agenda) or covered in the annual implementation report.</p> <p>The spectrum should include all the sub-sector programs/activities, including those funded both by domestic and external sources and addressed in the ESP/TEP.</p>	<ul style="list-style-type: none"> ● JSR agenda ● JSR report/ aide-mémoire ● Annual implementation report <p><u>Additional sources include:</u></p> <ul style="list-style-type: none"> ● ESP/TEP ● Annual action plan 	<p>JSR (or annual implementation report) systematically covers all sub-sectors, as reported in the ESP/TEP</p> <p>The JSR (or annual implementation report) discusses internally funded activities/programs, as reported in the ESP/TEP</p> <p>The JSR (or annual implementation report) discusses externally funded activities/programs, as reported in ESP/TEP</p> <p>The JSR (or annual implementation report) discusses externally-funded activities/programs that are <i>off-budget</i></p> <p>The JSR (or annual implementation report) discusses externally-funded activities/programs that are <i>off-plan</i></p>

Dimension	Question	Possible Data Source(s)	Response Items
3. Based on evidence	<p>Did the annual implementation report include the following?</p> <p>* Examples of financial information at trend/outcome level (ideally all indicators should be covered):</p> <ul style="list-style-type: none"> ● Execution percentage of total government budget ● Execution percentage of education budget ● Share of education expenditure over the total domestic expenditure disaggregated by capital and recurrent spending ● Education recurrent expenditure disaggregated by wage and nonwage 	<p>● Exclusively the annual implementation report</p>	<p>Brief situational analysis of review period with <u>key sector indicators at the outcome level</u> of the year under review</p> <p>Brief situational analysis of review period with <u>key financial information</u>* of the year under review (<i>see notes in question</i>)</p> <p>Sector progress and results achieved through the implementation of the ESP/TEP action plans (review of programs/activities)</p> <p>Expenditures at program/activity level** covered by <u>domestic funding</u> (see notes in question)</p> <p>Expenditures at program/activity level** covered by <u>external funding</u></p> <p>Follow-up on previous JSR recommendations</p> <p>Any additional element(s) that the annual implementation report includes that it is not listed above. <i>Please list.</i></p> <p>Beyond the annual implementation report, other document(s) other documents demonstrating JSR is evidence-based. <i>Please list.</i></p>

(continued)

Dimension	Question	Possible Data Source(s)	Response Items
	<ul style="list-style-type: none"> ⦿ Education expenditure disaggregated by major source: <ul style="list-style-type: none"> (i) Regular budget (ii) on-budget aid, earmarked traceable (iii) off-budget aid (major source of; i.e., ODA) ⦿ Share of external aid vs. domestic allocation ⦿ Projected domestic allocations <u>**Expenditures at program/activity level:</u> ⦿ Expenditures by programs and activities classified in the annual action plan (which should be able to integrate a full budget perimeter, including wage, recurrent non-wage and capital) 		

Dimension	Question	Possible Data Source(s)	Response Items
4. Monitoring tool	Was the JSR based on the following items?	<ul style="list-style-type: none"> ● JSR agenda ● JSR report/aide-mémoire ● Annual implementation report <p><u>Additional sources include:</u></p> <ul style="list-style-type: none"> ● ESP/TEP ● Annual action plan ● Results framework 	<p>Education data of the year under review. <i>If yes, please indicate in the comments column [H] the year of the data used.</i></p> <p>Total budget expenditures (national budget)</p> <p>Sector trends (outcome-level indicators)</p> <p>Performance of program/activities (output-level indicators)</p> <p>Education expenditures at the budget level</p> <p>Education expenditures at the program level (as defined in the ESP/TEP)</p> <p>International aid commitments and actual disbursements to the education sector</p> <p>Implementation problem areas and weaknesses</p> <p>Implementation achievements and good practices</p> <p>Previous JSR recommendations</p> <p>Monitoring of the indicators selected for the GPE variable part</p> <p>Specific thematic session(s) during the JSR. <i>If yes, please indicate below which ones:</i></p> <ul style="list-style-type: none"> Equity Learning Efficiency Aid modalities Aid effectiveness agenda SDG4 and FFA Any other thematic session(s). <i>Please list.</i> <p>Any other item(s) on which the JSR was based that was not listed above. <i>Please list.</i></p>

(continued)

Dimension	Question	Possible Data Source(s)	Response Items
	What form do the JSR recommendations take?		They are limited in number (judged as manageable by the CL)
5. Instrument for change	<ul style="list-style-type: none"> ➊ JSR report/aide-mémoire ➋ Annual action plan 	<p>They are clearly linked to the components of the ESP/TEP action plan and results framework</p> <p>They are in the form of remedial actions to address the weaknesses in the ESP/TEP implementation</p> <p>They designate responsible parties for implementation or action</p>	<p>They include a timeline for implementation and/or for leading to planning documents revision</p>
			They replicate previous-year recommendations