RISE Working Paper 20/057
December 2020

System (In)Coherence: Quantifying the Alignment of Primary Education Curriculum Standards, Examinations, and Instruction in Two East African Countries

Julius Atuhurra and Michelle Kaffenberger

Abstract

Improvements in instructional coherence have been shown to have large impacts on student learning, yet analysis of such coherence, especially in developing countries and at a systems level, is rare. We use an established methodology, the Surveys of Enacted Curriculum (SEC), and apply it to a developing country context to systematically analyze and quantify the content and coherence of the primary curriculum standards, national examinations, and actual teaching delivered in the classroom in Uganda and Tanzania. We find high levels of incoherence across all three instructional components. In Uganda, for example, only four of the fourteen topics in the English curriculum standards appear on the primary leaving exam, and two of the highest-priority topics in the standards are completely omitted from the exams. In Tanzania, only three of fourteen English topics are covered on the exam, and all are assessed at the "memorization" level. Rather than aligning with either the curriculum standards or exams, teachers' classroom instruction is poorly aligned with both. Teachers tend to cover broad swathes of content and levels of cognitive demand, unrelated to the structure of either the curriculum standards or exams. An exception is Uganda mathematics, for which standards, exams, and teacher instruction are all well aligned. By shedding light on alignment deficits in the two countries, these results draw attention to a policy area that has previously attracted little (if any) attention in many developing countries' education policy reform efforts. In addition to providing empirical results for Uganda and Tanzania, this study provides a proof-of-concept for the use of the SEC methodology as a diagnostic tool in developing countries, helping education systems identify areas of instructional (in)coherence and informing efforts to improve coherence for learning.

UKaid from the British people
Australian Aid
Bill & Melinda Gates foundation
Oxford Policy Management
Blavatnik School of Government
University of Oxford

System (In)Coherence: Quantifying the Alignment of Primary Education Curriculum Standards, Examinations, and Instruction in Two East African Countries

Julius Atuhurra
Japan Society for the Promotion of Science; Twaweza East Africa

Michelle Kaffenberger
Blavatnik School of Government, University of Oxford

This is one of a series of working papers from "RISE"—the large-scale education systems research programme supported by funding from the United Kingdom's Foreign, Commonwealth and Development Office (FCDO), the Australian Government's Department of Foreign Affairs and Trade (DFAT), and the Bill and Melinda Gates Foundation. The Programme is managed and implemented through a partnership between Oxford Policy Management and the Blavatnik School of Government at the University of Oxford.

Please cite this paper as:
Atuhurra, J. and Kaffenberger, M. 2020. System (In)Coherence: Quantifying the Alignment of Primary Education Curriculum Standards, Examinations, and Instruction in Two East African Countries. RISE Working Paper Series. 20/057. https://doi.org/10.35489/BSG-RISE-WP_2020/057

I. Introduction

With the rapid increase in schooling attainment in recent decades, education systems in many developing countries are primarily coherent for schooling – getting more children in school for more years. To achieve (nearly) universal enrollment in primary school, systems prioritized and were aligned to undertake the large-scale logistical tasks of building schools, procuring supplies, and training and hiring teachers. As the learning crisis has gained prominence there is increasing recognition that this coherence for schooling is not sufficient, on its own, for achieving learning for all. Many children who are in school are learning very little. To ensure all children receive an adequate education, education systems need to become coherent for learning (Pritchett, 2015). This presents a challenge: how to measure (in)coherence in a system and diagnose and improve areas not coherent for learning.

An education systems framework, posited by Pritchett (2015), defines elements in a system across which incoherence can arise. The framework is made up of system design elements (delegation, finance, support, information, and motivation) and relationships of accountability, including accountability through politics, through bureaucratic management, and accountability to parents, among others. In an ideal system, all design elements and relationships of accountability would be coherent with each other and coherent with learning. In reality, incoherence often abounds.

Incoherence can arise across elements of the systems framework. For example, schools may be delegated by the Ministry of Education to teach a new curriculum but not given the adequate finances or support to procure the new curriculum and conduct teacher training. Communities may be delegated the authority to hold schools accountable for certain outcomes (such as through local school committees), while not being given the information necessary, such as on learning outcomes, to do so effectively.

In this paper, we use an established methodology, the Surveys of Enacted Curriculum (SEC), to examine the coherence of instructional components in a system, specifically the content of the curriculum standards, national examinations, and actual teaching delivered in the classroom. Schools and teachers are expected to complete many tasks, among them completing the prescribed curriculum within an allotted timeframe and preparing children for national exams, especially the high stakes primary leaving exams at the end of primary school in many African countries. In the terminology of the systems framework, teachers (and schools) are delegated multiple tasks, which may or may not be coherent with each other, and may or may not be coherent for learning.

Improvements to the coherence of instructional components has been shown to have large impacts on student learning. In a review of education interventions in developing countries, Crouch and DeStefano (2017) find that interventions focused on certain tightly focused and aligned ("coherent") pedagogical practices can have large effect sizes, and that these effects are often larger than interventions focused on structural factors (such as school autonomy or results based teacher pay) or inputs (such as infrastructure or money). One example they highlight is Tusome, a program implemented by the Kenyan government in collaboration with RTI International that includes a coherent package of instructional materials, books for students, and training and coaching for teachers all of which is calibrated to learners' learning levels. This program has

achieved effect sizes ranging from 0.6 to 1.0 standard deviations in English and Kiswahili (Freudenberger & Davis, 2017; Piper et al., 2018). Another example is Pratham, an Indian NGO that runs a "Teaching at the Right Level" program. This program focuses on bringing instruction into coherence with student learning levels and has seen effect sizes ranging from 0.15 to 0.70 (Banerjee et al, 2016).

Coherence of formal curriculum standards, examinations, and teacher instruction are associated with increased student learning in developed countries. Using the same tools as in the current study, Smithson and Collares (2007) and Gamoran et. al (1997) find positive associations between alignment measures and student achievement in reading and mathematics. Porter (2002) states that alignment of curriculum standards and examinations is important for providing a "coherent set of expectations to teachers". In describing what they call (in a negative sense) a "hopeful pedagogy", as opposed to a deliberate effort to achieve a well-aligned system, Nicholas and Raider-Roth (2016) conclude that a misaligned system will greatly undermine efforts to improve children's learning, teacher engagement, institutional improvement and accountability.

Little analysis has been done, however, of the alignment of formal curriculum standards, examinations, and teacher instructional content in developing countries, especially Sub-Saharan Africa (Twaweza, 2015). To date studies that systematically analyze the content and progression of curriculum standards and the alignment of examinations and classroom instruction with the standards are rare (Burdett, 2017).

The non-governmental organization Twaweza East Africa recently undertook an innovative approach that quantifies the content progression and level of coherence of instructional components in Uganda and Tanzania. Using the Surveys of Enacted Curriculum (SEC), a team of experts developed a taxonomy of topics and subtopics across four primary-level subjects (mathematics, English, social studies, and science). They coded the competences in the curriculum standards for each primary school grade and items in the national primary leaving exams (PLEs) against these comprehensive lists of topics and subtopics. They rated the "levels of cognitive demand" for each competence (or item), which range across five levels from memorization to application to non-routine problems. The team also piloted an effort to ascertain teachers' instructional content coverage using the same structure of topics, subtopics and levels of cognitive demand. The result is a rich dataset that allows descriptive analyses of the content that is embedded in the curriculum standards, national exams, and teachers' instruction, including the progression of content across grades as learners advance in school, and measurement of the levels of coherence (or incoherence) of the components with each other.

This paper analyzes the progression and coherence of the primary curriculum standards, national primary leaving exams and teacher instructional content for primary mathematics and English in Uganda and Tanzania. It empirically shows areas of coherence and incoherence and by doing so suggests areas that, if brought into greater coherence, could improve children's learning. The paper makes four main contributions, including two main sets of findings.

First, this study provides structured, country-specific taxonomies of topics, subtopics, and levels of cognitive demand that can be used to evaluate and diagnose existing curriculum standards,

examinations, and instructional content and inform revisions to or development of new instructional components such as textbooks or teacher training. Such taxonomies do not exist for many (or most) developing countries, making it difficult or impossible to develop instructional materials with consistent, suitable structures to enable children's learning within and across grades and schooling cycles.

Second, the study finds that content coverage in Uganda and Tanzania varies across the curriculum standards, national exams, and teacher instruction, from being very narrowly focused on a few topics, as with the primary leaving exams, to broadly covering a full suite of topics, as with teachers' classroom instruction. In both countries, the English curriculum standards follow a non-systematic "drop and reinstate" content coverage pattern in which a topic is covered in one grade, not covered ("dropped") in the next grade or two, and then reinstated sometimes at a (much) higher level of cognitive demand. This type of inconsistent progression not only limits content continuity from grade to grade, but can make it difficult for children to fully engage with the material in the higher grades as they have not had the opportunity to master the necessary intervening skill levels. Taken to the extreme, the Ugandan English standards do not cover the foundational topics of phonics or phonemic awareness in the early primary years, and only introduce phonemic awareness in grade six expecting children to engage with it at a relatively high cognitive skill level.

The national exams tend to prioritize the two lowest-order skill levels (memorize and perform procedures), curriculum standards prioritize lower order skills (Tanzania) or mid-level skills (Uganda), and teachers' instruction often covers all five levels of cognitive demand. While the breadth of skill levels covered by teachers may be productive if it ensures children gain procedural and conceptual mastery of competencies, it could also indicate an overemphasis on breadth over depth of knowledge and abilities. In Uganda, rural teachers tend to prioritize lower-order skills of recall/memorization more than urban teachers.

Third, the analysis reveals high levels of incoherence across many components. National exams are, typically, poorly aligned with the curriculum standards, indicating (using the language of the systems framework) that teachers are being delegated competing tasks. In Uganda, for example, only four of the fourteen topics in the English curriculum standards appear on the primary leaving exam, and two of the highest-priority topics in the standards are completely omitted from the exams. In Tanzania, the alignment between the English exam and curriculum standards is 0.11 on a scale ranging from 0 to 1, with 0.50 considered reasonably well-aligned. Ugandan mathematics is an exception with an alignment of 0.57 between curriculum standards and exams.

Teacher instruction generally aligns poorly with the curriculum standards as well. In Tanzania, two mathematics topics (number sense and measurement) constitute two-thirds of the curriculum standards content but make up less than half of teachers' instructional focus. Conversely, teachers place 25% of their mathematical instructional focus on operations, which accounts for only 10% of the curriculum standard's content. While the curriculum standards, in both countries and for both subjects, tend to prioritize low- and mid-level cognitive skills, teachers' instruction covers a broad set of skill levels, often stretching from the lowest order (memorization) all the way to the highest order (application to non-routine problems). In Uganda, alignment between teachers' instruction and curriculum for English is only 0.15 on the 0 to 1 scale. In Tanzania teacher

instruction is poorly aligned with exams, with alignment measures of 0.33 for mathematics and 0.11 for English.[1]

Finally, this study demonstrates a methodology that education systems can use to diagnose areas of incoherence in their own systems, which in turn can inform consistent, well-designed curriculum standards and actions to improve coherence across instructional components and learning outcomes. We have focused here on three components of instruction, but the methodology can be used to diagnose coherence across any number of instructional components. For example, an education authority could evaluate new textbooks or teacher support materials for their coherence with curriculum standards. This is in addition to the uses demonstrated here, such as evaluating the content and progression of curriculum standards, to ensure a consistent set of learning objectives and logical progression as children advance from one grade to the next, and diagnosing the coherence of these standards with the expectations put forth in national exams.

By providing a structured systematic tool for quantifying the content of instruction across multiple instructional components, the SEC methodology can provide valuable insight to education systems to improve instructional coherence and ultimately learning outcomes.

II. Curriculum standards, national examinations, and potential sources of incoherence in Uganda and Tanzania

II.A. Basic Education in Uganda and Tanzania

Basic education in Uganda constitutes the first 12 years of formal schooling, including at least one year of pre-primary, seven years of primary schooling, and four years of lower secondary.[2] At age six children are required to enroll in primary school. The seven years of primary school are divided into three distinct learning cycles: lower primary grades one to three (P1 – P3), transition grade four (P4), and upper primary grades five to seven (P5 – P7). In 1997 the Government of Uganda (GoU) introduced its Universal Primary Education (UPE) policy, abolishing tuition fees for government primary schools. A significant minority (about 20%) of primary school students attend private, fee-charging schools (MoESTS, 2014). At the end of grade seven, children sit for a high-stakes national examination that determines entry into lower secondary school.

While access to primary school has increased substantially in recent years, with net enrollment in primary school over 90%, learning improvements have not kept pace. Several studies have found that even within the East and Southern Africa region, basic educational performance in Uganda is significantly lower than other countries, including Kenya and Tanzania (Byamugisha & Ssenabulya, 2005; Ward, Penny, & Read, 2006; USAID/RTI, 2014; Jones, 2015; Atuhurra, 2016; Uwezo, 2016). On the SACMEQ III assessment, Uganda averaged 479 for reading, and 482 for mathematics, well behind its East African neighbors Kenya and Tanzania (Bashir et. al., 2018; SACMEQ, 2020).

[1] In Uganda the sample of teachers did not allow for measuring alignment between teacher instruction and examinations.

[2] Recent evidence suggests that only about 27% of 3- to 5-year-olds attend pre-primary education, however.

Uganda has undertaken five primary curriculum reforms since 1965, the most recent being in 2007-2010 (Ezati, 2016). The current primary curriculum is organized in three distinct cycles which align with the three learning cycles in primary school (NCDC, 2006a; NCDC, 2006b). In P1-P3, the curriculum is thematic, aimed at achieving quick development of foundational numeracy and literacy, adopts a learner-centered approach to teaching and assessment, includes teacher guides and resource books, and is taught in the child's local or familiar language. Starting from grade four and through upper primary the curriculum transitions from being theme-based to being subject-based and from using a child's home language to English as the medium of instruction. Effective learner-centered instruction requires adequate support and training for teachers, and while the adoption of this approach in Uganda has been praised for the good intentions in practice implementation has proven difficult and frustrating for teachers (Altinyelken, 2010), and an emphasis on high-stakes summative exams remains.

In Tanzania, the Education and Training Policy 2014, which came into force in 2015, marked the fifth major curriculum reform since independence (Nzima & Mkumbo, 2017). Prior to 2015 compulsory education consisted of seven years of primary school. Tanzania is in the process of restructuring the system to include 12 years of free compulsory schooling, a process which is expected to be completed in 2021 (MoEST, 2018). The new structure begins with one year of pre-primary for children who are six years old, followed by seven years of primary and four years of lower secondary schooling. This can then be followed by two years of non-compulsory upper secondary school. While the restructuring is in progress, lower secondary school is fee-free but not yet compulsory, and admittance continues to be determined by passage of the primary leaving exam (MoEST, 2018). Enrollment in primary school in Tanzania increased dramatically in the early 2000s following decentralization and the abolition of primary school fees, rising from 4.9 million children attending primary school in 2001 to 8.4 million in 2010 (Matete, 2016), and the net primary enrollment rate now stands at 84% (MoEST, 2018).

Similar to Uganda, while access and enrollments have increased rapidly in recent decades, learning outcomes have remained low. National exams in 2014 showed that only 8% of standard two students could read properly, and similarly only 8% could subtract (UNICEF, 2017). The World Bank's Service Delivery Indicators indicated some improvements in learning outcomes between 2014 and 2016, with the percent of primary school children able to identify a letter rising from 61% to 70%, and the percent who could read a paragraph rising from 25% to 30% (Asim et al, 2019). The percent who could subtract double digit subtraction rose from 42% to 50%. But these outcomes indicate many children lack basic skills.

Traditionally instruction has been in Kiswahili in the primary years, transitioning to English instruction in secondary school, though recent policies have promoted the use of both languages at all education levels (Vuzo, 2018).Tanzania launched a new primary curriculum in 2015 which is more competency-based than its predecessor (MoEST, 2018). Part of this reform, the "3R" reform, included prioritizing reading, writing, and arithmetic in grades one and two, shifting 80% of instructional time to these core literacy and numeracy subject areas (USAID, 2015; Mbiti and Rodriguez-Segura, forthcoming). As part of this reform Kiswahili was prioritized over English instruction in grades one and two.

II.B. Development of curriculum standards and national examinations in Uganda and Tanzania

Both countries pursue a standards-based education system that requires teachers to teach content that is prescribed in the curriculum standards and use instructional materials approved by the curriculum institutes. Likewise, the national examinations bodies are required to assess students' learning achievements based on the intended content structure spelled out in the curriculum standards. The curriculum standards and national examinations, however, are developed through separate processes.

In Uganda, a government white paper (GoU, 1992) sets out the broad objectives to be achieved at each level of basic education. These broad objectives, however, are not translated into a more specific taxonomy of topics, subtopics, or cognitive demand and sequencing that would be required to reach objectives (Atuhurra and Alinda, 2018). The curriculum standards are set by the National Curriculum Development Centre (NCDC), a "corporate autonomous statutory body" under the Ministry of Education and Sports (MoES). The statutory act establishing the NCDC (GoU, 1973) specifies twelve specific roles, most notably to investigate, initiate, revise, reform and evaluate curricula at primary, secondary, tertiary and other education institutions, as well as designing and developing teaching aids and instructional materials to aid teaching and learning. The NCDC is also jointly responsible, with other teaching and examination organizations, to devise, test and evaluate examination questions and methods of examining students.

The high-stakes primary leaving exam is the purview of a separate body, the Uganda National Examinations Board (UNEB). The UNEB is responsible for conducting end-of-cycle primary, secondary, technical and other examinations and awarding certificates to successful candidates (GoU, 1983). In practice, the over-arching objectives of conducting these exams relate mainly to the selection and sorting of students for further education and future careers (Allen, Elks, Outhred & Varly, 2016). UNEB also regularly conducts the relatively lower-stakes National Assessment of Progress in Education (NAPE) involving grades three and six children.

Reforms of curriculum standards in Uganda almost always enlist high-level political engagement and oversight. This level of oversight is typically missing from the design and conduct of exams (at the school, district, and national level). Political scrutiny of national exams primarily occurs only when performance results are published, with a focus on identifying the best and worst performers, whether among learners, teachers, schools, or districts. Throughout the parallel and mostly separate processes it is largely assumed (rather than ensured) that the exams align well with the curriculum standards, presenting an opportunity for incoherence. (Ward, Penny & Read, 2006; World Bank, 2012; Allen, Elks et. al, 2016; Luigi Giussani, 2017; Munene, 2017)

In Tanzania, the Tanzanian Institute of Education (TIE) is primarily responsible for curriculum development, and was responsible for the most recent round of reforms in 2015 (MoEST, 2018). It is a public institute under the Ministry of Education, Science, and Technology, with responsibility for designing and developing curricula for pre-primary, primary, secondary, and teacher training education levels (GoT, 1975). TIE also carries out in-service and pre-service

training for teachers and provides and oversees quality assurance for teaching methods, objectives, and materials.

Many other agencies, however, are also involved in curriculum development interventions, including the National Council for Technical Education (NACTE), National Examinations Council of Tanzania (NECTA), School Quality Assurance Division of the MoEST, and the Agency for Development of Education Management (ADEM), in addition to partner-supported programs such as EQUIP-T, Tusome Pamoja, and others (MoEST, 2018). While TIE has established a National Curriculum Framework for Basic Education and Teacher Education, intended to regulate curricula and "avoid overlaps and oversights" (MoEST, 2019), in practice these efforts are largely uncoordinated, resulting in a lack of a comprehensive and holistic approach to curriculum in Tanzania (MoEST, 2018).

Changing policies also affect curriculum standards. Following Tanzania's 2014 Education and Training Policy which instituted six years of primary school, a six-year primary curriculum was developed and implemented in 2015 by TIE. The government has since changed tacks, reverting to seven years of primary school, requiring TIE to again reform the curriculum (MoEST, 2018).

The National Examinations Council of Tanzania is the government institution responsible for the administration of all national examinations in Tanzania (GoT 1973; GoT 2019). In particular, NECTA designs and administers the primary school leaving exam (PSLE), supervises the grading of the exam, and issues performance reports. Similar to Uganda, this is intended to be in accordance with the curriculum developed by TIE (NECTA, n.d.). NECTA has a "moderation panel" composed of curriculum developers, school inspectors, university lecturers, and experienced teachers who are responsible for ensuring items and questions on the exam are "representative of the syllabus content and objectives and are of the appropriate difficulty for the level being assessed" (NECTA, n.d.). The poor coordination of curriculum interventions (MoEST, 2018), however, makes this difficult in practice.

II.C. Potential sources of incoherence

The multitude of actors in the curriculum and examinations processes in each country provides opportunities for incoherence across curriculum standards, national exams, and in turn, teacher instruction. To formalize our analysis of the sources of incoherence, we apply the systems framework from Pritchett (2015). The framework includes four principal-agent relationships and five system design elements. Incoherence can arise within a relationship of accountability across design elements (down a "column" of the framework); across relationships of accountability within a design element (across a "row" of the framework); or between two relationships of accountability (between two columns of the framework) (Figure 1).

Figure 1. Education systems framework of accountability

Five design elements	Principal-agent relationships of accountability			
	Politics (Citizens to Executive Authority)	Compact (Executive Authority to Education Authorities)	Management (Education Authorities to Front-line providers)	Voice/ Client power (Citizens to Front-line providers)
Delegation			**A_1, B_1**	**B_2**
Finance				
Support			**A_2**	
Information			**A_3**	
Motivation				

The current study examines two main sources of (in)coherence within the framework. The first is within the Management relationship, involving what education authorities delegate schools and teachers to do (A_1 in Figure 1), how well supported teachers are to achieve the delegated tasks (A_2), and the information education authorities have to hold schools and teachers accountable (A_3). The second is between the Management and Voice/Client Power relationships, involving the potential for incoherence in what education authorities delegate schools and teachers to do (B_1) versus what parents delegate (demand) schools and teachers to do for their children (B_2).

For the first source of incoherence, within the management column, in Uganda the NCDC delegates to schools and teachers to implement the curriculum standards as designed, with a key objective of ensuring children learn the intended content. At the same time, the UNEB delegates to schools and teachers to prepare children for the primary leaving exam. As described in Section II.B., poor coordination means these two tasks being delegated to education providers could be incoherent, for example if the content of the curriculum does not align well with the content on the exam. Similarly, in Tanzania, TIE delegates the curriculum for schools and teachers to cover, while NECTA sets and delegates preparation for the primary leaving exam, leaving schools and teachers with potentially conflicting or even contradictory delegated tasks.

Additional actors involved in curriculum development and implementation provide additional opportunities for incoherence. For example, a partner-sponsored program may intervene with a special curriculum for one or a couple grade levels in a subset of schools. This presents opportunities for incoherence between what teachers are delegated to instruct through the intervention and the national standards. If the special curriculum has different sequencing and progression of topics than the national curriculum it could present additional incoherence for children's learning in subsequent grades.

The support provided to teachers to implement the delegated curriculum, such as through textbooks, teachers' guides, and training, also has important implications for coherence of instruction in the classroom. Even if the curriculum and exams are well-aligned with each other, if the supporting materials and training provided to teachers are poorly aligned with both, then their classroom instruction may differ from the intended content. While analyzing the coherence of support inputs with other instructional components is beyond the scope of the current study, the SEC methodology can be applied to this type of analysis and this important area for future work is addressed in the Discussion Section.

A final potential source of incoherence in the management column is the limited information education authorities such as NCDC and TIE have for measuring the actual implementation of the curriculum standards as designed. For example, there are no mechanisms to measure whether schools implement the curriculum as specified, cover the prescribed topics, and whether and how much remedial instruction is needed. Education inspection systems, for example, are not typically equipped to gather detailed data on instructional content and practices in the classroom. Without this kind of information, the curriculum authorities cannot hold schools accountable for completing the curriculum, nor do they have information needed to guide future curricular reforms.

Second, across the delegation row, schools and teachers can also experience incoherence between what they are delegated to do by the education authorities (in the management relationship) and what they are delegated to do by parents, students, or other citizens (in the voice/client power relationship). For parents and students, passing the high-stakes primary leaving exams may be a top priority, leading parents to pressure schools and teachers to emphasize exam preparation over curriculum mastery. This is coherent with one part of the management relationship, the examination bodies, but it could be incoherent with another part of the management relationship, the curriculum bodies, if the curriculum and examinations are not well-aligned. Further complicating these relationships is that citizens may be more able to hold schools accountable for exam preparation than for learning gains if they themselves do not fully understand what constitutes high quality education in primary grades or have not been given adequate information on children's learning gains. Exam scores and pass rates by comparison are more straightforward to understand and interpret, which could further citizen's focus on these.

Schools and teachers, delegated competing tasks and limited time, may be forced to choose whether to align their instruction more with curriculum standards, exam preparation, or some other alignment. For example, if teachers believe the curriculum is inappropriately leveled or paced for the children in their classroom, rather than align their instruction with either the curriculum or exams they may align with children's learning levels and needs. Teachers could also follow tradition or earlier training rather than aligning their instruction with the most recent adaptation of curriculum standards.

III. Methodology: Using data to describe instructional content

III.A. The Surveys of Enacted Curriculum Methodology

Quantitative analyses of education curricula are rare in developing countries. At best, curriculum analysis relies on the use of observational methods to identify and address gaps in content and

pedagogy in the classroom. In the United States however, for over twenty years, the Surveys of Enacted Curriculum (SEC) has provided a set of practical and reliable data collection, analysis and reporting tools that are used by researchers, as well as state and local education agencies to systematically analyze the academic content of the education curriculum and measure the degree of alignment between instructional elements (Blank, Porter & Smithson, 2001; Porter, 2002; Smithson 2013). SEC has also been used to facilitate education content reform as well as provide professional development informing teachers' instructional content and practices in the classroom (Porter, 2002).

In a standards-based education model, in which curriculum standards are specified and teachers are expected to teach to those standards, as in both Uganda and Tanzania, it is necessary to have high-quality, well-designed curriculum standards that facilitate and enable children's learning. With high quality standards in place, an important goal is then to achieve alignment with these standards across instructional components. Classroom instruction, learning assessments, instructional materials and teacher professional development should all originate from and tightly link to the well-designed curriculum standards (Porter, 2002).

The SEC tools are used in multiple ways to help achieve these aims. One is providing a systematic analysis of curriculum standards which can be used by curriculum experts in a country to compare the content and progression with instructional good practices informed by cognitive science and research on how children learn. A second, and the main focus of the current study, is the use of SEC tools to diagnose the degree of alignment across instructional elements, identifying areas of incoherence.

Implementing the SEC methodology involves coding and rating instructional content on two dimensions: the topics and subtopics covered, and the cognitive demand required for each. The topics and subtopics that fall under each academic subject are context specific. For example, the appropriate topics and subtopics for literacy in Uganda versus in the United States may differ. To implement the methodology a first step is to identify (if one exists) or develop (if one does not exist), a taxonomy of the relevant topics and subtopics for the analysis.

For cognitive demand, a consistent five-level scale is used across applications. Cognitive demand ranges from least demanding, as in memorize or recall, to most demanding, evaluate or apply to non-routine problems.[3] An illustrative definition (using performance descriptors) for each level is provided in Table 1.

[3] Note that levels of cognitive demand are not synonymous with levels of difficulty. A memorization problem can be difficult, while an analysis problem could be easy. The levels of cognitive demand represent differences in required thinking, connecting, and doing to complete.

Table 1. Levels of cognitive demand and illustrative definitions of each

Level of cognitive demand	Illustrative definition
Memorize/Recall	Recognize, identify, or recall facts, definitions, or formulas
Explain/Perform Procedures	Perform procedures, solve routine problems, do computations, make observations, take measurements
Generate/Demonstrate understanding	Communicate ideas, explain findings from analysis, explain reasoning
Analyze/Conjecture	Make and investigate conjectures, infer and predict
Evaluate/Apply to non-routine problems	Apply and adapt strategies, solve novel problems, make connections

The SEC approach relies on experts to code and rate curriculum standards, instructional materials, and assessments. For teacher instructional content, teachers are trained on the methodology, the topics and subtopics, and levels of cognitive demand. The teachers then complete self-report surveys on the topics and levels of cognitive demand covered in their instruction. Surveys are typically conducted at the end of a school year and cover content taught in the year that just ended.

The experts or teachers are asked, for each subtopic, for the level of coverage (none, slight, moderate, or sustained) and the relative emphasis of each level of cognitive demand for that subtopic (no emphasis, slight, moderate, or sustained) (Porter, 2002). These data are then transformed into proportions of total instructional time for each cell in the two-dimensional matrix. The output is a three-dimensional content map with topics/subtopics on the Y axis, levels of cognitive demand on the X axis, and percent of total coverage on the Z axis (Smithson, 2015). (See Figure 2, in Section IV, for an example of the visual content maps.) These three-dimensional figures provide informative visual displays of the content in the instructional component being analyzed and can be read like topographical maps.[4] Utilizing a simple two-step process to interpret the maps, specific data points are first identified at the intersection of topic/subtopic and cognitive demand level, and then emphasis is determined by variations in color schemes for each data point (Atuhurra & Alinda, 2017).

Two levels of SEC analysis are possible, coarse grain at the topic level, and fine grain at the subtopic level. Coarse grain content analyses are summative analyses that portray relative emphasis on topics and cognitive demand. Such analyses provide the distribution of content across topics and performance expectations. Thus, coarse grain content maps and marginal charts display the topic level content emphasis structure. Fine grain content analyses on the other hand facilitate formative analyses at the subtopic level. They underlie the coarse grain descriptions and enable a more detailed diagnostic analysis of a topic to ascertain the specific distribution of emphasis across the subtopic content areas in each topic. After obtaining the summative alignment picture, the fine

[4] One key difference between the SEC outputs and topographical maps is that topographical maps (and similar visual displays) are based on continuous data, while the results of the SEC analysis are based largely on discrete data. The focus in interpreting SEC results should be on the intensity at intersection points on the content map.

grain maps are used to point to specific subtopic areas that need to be addressed to improve alignment (Smithson, 2013).

A final step is analysis of the degree of alignment across instructional components. As described in more detail in Porter (2002), the alignment analysis compares the proportion of coverage in each cell of one content map (e.g. describing an assessment) with that in another (e.g. describing curriculum standards). The results are reported as alignment indices. Alignment index values range from zero to one, with 1.0 indicating perfect alignment and zero indicating no alignment or perfect misalignment.

When analyzing alignment of curriculum standards and assessments, the expectation is not for perfect alignment, as assessments cover a sample of items from a domain and therefore may not cover every topic and subtopic included in the curriculum. Furthermore, when analyzing alignment of teacher instructional content with curriculum standards and assessments, perfect alignment is again not the appropriate expectation. Teachers face varied learning levels and needs in their classrooms and so it is reasonable that they would make some adjustments to the content being taught. At the coarse grain, topic level a value of 0.5 on the alignment index is typically considered reasonable alignment between instructional components (Smithson, 2015).

It is also common that alignment is lower at the fine grain, subtopic level than at the coarse grain level. This is for two primary reasons. The first is that, when analyzing alignment of teacher instructional content with other instructional components, more divergence would be expected among subtopics. For example, for a lesson on the topic of "measurement", a teacher may follow with the curriculum-prescribed topic of measurement, but diverge in the specific subtopics depending on the level of mastery in their classroom, their personal preference or previous training for subtopics, or some other factor influencing how they cover the material. The second is that there are greater opportunities for disagreement among raters of curriculum and assessments, as well as teachers, at the fine grain level. Within a subject, such as primary mathematics, there may be 14 topics across which raters must allocate topics, and the distinctions across topics are relatively clear. Within several topics, however, there may be as many as 16 subtopics, and the distinctions may be much less clear. At the fine grain level, therefore, a rule of thumb is to expect the fine grain alignment measure to be at least half the coarse grain measure to represent good alignment. More detailed examination of fine grain content maps, beyond the summary alignment index measure, can provide greater insight on subtopic coverage as well.

The SEC methodology provides multiple advantages for understanding the content and alignment of instructional components. First, it provides a common content matrix (or template) that can be used across different instructional components, making it possible to draw comparisons that would otherwise be ill-defined. Second, it allows independent and replicable descriptions of the content using robustly generated content analysis data. Third it allows systematic, quantified measures of alignment (Porter 2002). Finally, by using only two dimensions, topics/subtopics and cognitive demand, it is relatively less complex to implement and therefore can be used in more contexts (Atuhurra and Alinda, 2018). The ability to analyze teachers' instructional content adds a further contribution to curricula studies since such data reveals the opportunities students have had to learn

the intended content and enables policy makers to explore the extent to which curricula reforms are being implemented in the classroom (Kurz et. al., 2010).

The SEC approach also has limitations. Its reliance on self-reported surveys of teachers for the analysis of teachers' instructional content could be problematic in a couple ways. If teachers are not adequately trained on the methodology, they could report incorrect information. Completing the surveys accurately requires understanding the topics, subtopics, and levels of cognitive demand about which information is collected. Furthermore, teachers may be tempted to intentionally report incorrect information if they feel there is a "correct" answer, or they believe they may be judged based on their answers. The SEC approach addresses this in a couple ways. Teachers must undergo a comprehensive orientation session before completing the surveys, to foster a basic understanding of the methodology, clarify the goals of the survey, and thoroughly discuss the cognitive demand concept. Teachers are also shown that the SEC tools are designed to promote their professional growth and facilitate self-reflection on their work in the classroom and enable peer-level dialogues on how they can improve their practices. Many teachers report appreciation for the insight the tools provide following training. SEC data is never used for evaluation of teacher performance.

Classroom observations, commonly used by other methodologies aiming to understand teachers practices in the classroom[5], are inappropriate alternatives to teacher self-report surveys in the case of instructional content analysis. Classroom observations occur at a point in time during the school year, and even if repeated would give only snapshots of instructional content observed while the researcher was in the classroom. Analyzing alignment of instructional content with curriculum standards requires information on the content taught throughout the whole year, and so gathering information from teachers on all their instructional content is necessary for this kind of analysis.

A second limitation of the SEC methodology is that, by design, SEC is a diagnostic tool which in and of itself does not provide a normative conclusion on the "best" or the "right" form for the content of instructional components. SEC provides, for example, a systematic analysis of the alignment between curriculum standards and an exam, but it is then up to experts in the relevant context to use the results to determine whether the exam should be brought into better alignment with the standards, the standards into better alignment with the exam, or both reformed to better align with children's learning and needs.

The SEC tools have served multiple additional purposes in education research, policy and practice, including informing design of teacher professional development programs, evaluating curriculum reforms, analyzing instructional materials such as textbooks, and explaining observed variances in learner achievement gains on tests and exams (Porter, 2002). In this study, we utilize SEC data to assess the degree of systemic coherence of instructional components at the primary education level in Uganda and Tanzania. To our knowledge, these data represent the first time this methodology has been used in developing countries.

[5] See, for example, the CLASS (https://teachstone.com/class/) and TEACH (https://www.worldbank.org/en/topic/education/brief/teach-helping-countries-track-and-improve-teaching-quality) methodologies.

III.B. SEC contextualization in Uganda and Tanzania

Twaweza, a regional NGO working on improving education in East Africa, used the SEC tools to conduct studies of basic education curricula effectiveness in three East African countries – Kenya, Tanzania and Uganda. With technical support from the SEC project at the Wisconsin Center for Education Research (WCER), at the University of Wisconsin-Madison, these efforts yielded extensive data on curriculum standards, national examinations, and teacher instructional content in each country. Here we focus on two of the countries – Uganda and Tanzania.

In 2015, Twaweza organized consultative forums in both Uganda and Tanzania, drawing participants from stakeholder organizations including universities, curriculum development and examination bodies, schools, and other civil society organizations working in the education sector, to provide input to the study. Select members of the consultative forums, as well as additional individuals chosen for their specific expertise, were engaged as experts to generate the comprehensive topic taxonomies and conduct the coding and rating of the curriculum standards and national examinations in each country. In Uganda, eleven experts were drawn from two national universities, the national curriculum body, primary teacher's colleges, schools, and the school inspectorate department of the education ministry.[6] In Tanzania, nine experts were drawn from the national curriculum institute, national examination council, two national universities, three schools and Twaweza.[7] The experts received training on the SEC methodology and the application of the tools to analysis of curriculum standards and exams.

In Uganda, the SEC studies covered four primary-level subjects including mathematics, English, integrated science, and social studies. In Tanzania, five subjects, mathematics, English, Science, social studies and biology (at secondary level only), drawn from both primary and lower secondary, were covered.

Existing subject-specific taxonomies could not be identified in either country, so the panel of experts in each country first developed comprehensive taxonomies of subject-level content covered at the basic education level. The development of these subject taxonomies was informed by the academic content in each country's curriculum standards, as well as other theoretical and evidence-based information about the important components of each subject in question. Content lists were developed to reflect both general content areas (topics) and the more specific content areas (subtopics). In Uganda the contextualized primary education mathematics taxonomy has 13 topics and 96 subtopics, and the English taxonomy has 14 topics and 133 subtopics. In Tanzania the contextualized basic education mathematics taxonomy had 16 topics and 180 subtopics, while the one for English had 14 topics and 133 subtopics. While there have been no major primary level

[6] The detailed composition in Uganda's case was as follows: three experts from the National Curriculum Development Centre (NCDC), one expert from Makerere University's school of education, one expert from Kyambogo university's teacher education department, one expert from the ministry's directorate for school inspections, two primary school teachers, one expert from Nakaseke core primary teachers' college and two experts from Twaweza East Africa - Uganda.

[7] The detailed composition in Tanzania's case was as follows: one expert from the Tanzania Institute of Education (TIE), one expert from the National Examinations Council of Tanzania (NECTA), one expert from the University of Dar es salaam (UDSM), two experts from Dar es salaam University College of Education (DUCE), two primary school teachers, one secondary school teacher and one expert from Twaweza East Africa - Tanzania.

curriculum reforms conducted in Uganda since the development of this taxonomy, curricular changes have been effected covering grades three to seven in Tanzania since 2017.

In each country surveys of teachers' instructional content and practices were conducted in two districts, one urban and one rural. While not representative, the two districts in each country were purposively selected to provide a comparison between rural and urban settings.

In Uganda one hundred public schools from each of Wakiso (urban) and Iganga (rural) districts were selected for inclusion. Teachers were selected from three grade levels: primary three, five, and six, and a total of 600 teachers were included in the study. In Tanzania the survey involved teachers from both public and private schools at primary and secondary levels. A total of 622 teachers from 41 schools participated in this survey – 25 schools from the urban Ilala district and 16 from the rural Mkuranga district[8]. All teachers were oriented to and trained on the SEC methodology and taken through practice exercises involving coding and rating learning objectives in their subject curriculum and items from exams and tests before completing the survey tools.

IV. Results

IV.A. Instructional Coherence in Uganda

i) The content and progression of the primary mathematics curriculum standards

Three topics dominate the Ugandan primary mathematics standards: number sense, properties and relationships; operations; and measurement (Figure 2). Content maps, such as those in Figure 2, display the results of the SEC analyses throughout this paper. The content maps list the topics (or subtopics) on the vertical axis, roughly in order of increasing topical sophistication moving down the axis. Levels of cognitive demand are on the horizontal access, in order of increasing levels of cognitive demand moving to the right along the axis for each map. They are read as topographical maps, with the colors indicating the level of emphasis at each intersection of a topic and cognitive demand level. So, for example, Panel 1 of Figure 2 shows that in the Ugandan P1-P3 mathematics curriculum standards, number sense at the procedures and demonstrate levels of cognitive demand and operations at a procedures level of cognitive demand are the three most emphasized points on the content map.

In each of the four panels in Figure 2, over sixty percent of total emphasis is devoted to the three most emphasized topics. Another set of three topics follows the first set, accounting for about twenty percent of overall coverage throughout the primary cycle – geometric concepts; data displays; and special topics.

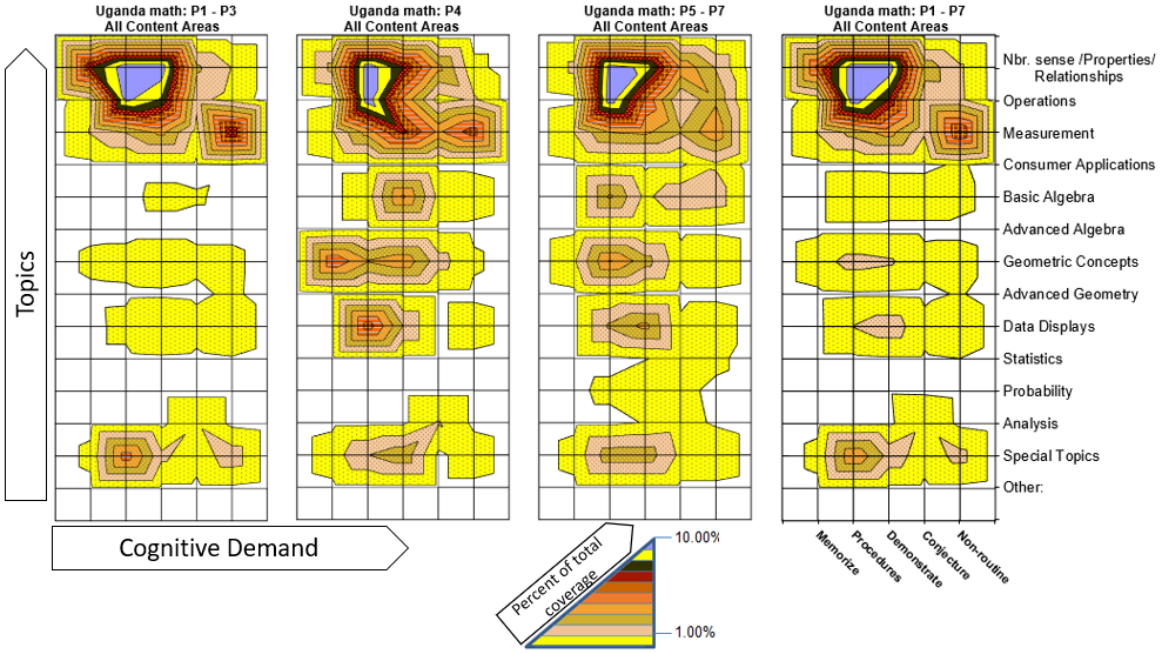
In terms of cognitive demand, the primary mathematics standards primarily emphasize developing learner abilities to perform mathematical procedures and demonstrate understanding of concepts. For measurement however the curriculum standards emphasize solving non-routine problems.

The lower primary classes focus almost entirely on the three foundational topics of number sense, operations, and measurement. The content coverage structure then becomes broader and more

[8] A detailed breakdown of the achieved sample of primary mathematics and English teachers is given in Appendix A.

complex in upper primary, especially in the last three grades (P5 – P7) where all topics except advanced algebra and advanced geometry are covered (Panels 2 and 3 of Figure 2).

Figure 2: Ugandan topic level mathematics curriculum standards: Lower (P1-P3), transition (P4), and upper primary (P5-P7), and aggregation (P1-P7)



Between-grade content alignment analyses give the alignment of content coverage from one grade to the next. In Table 2 these are shown with coarse grain (topic level) alignment indices above the main diagonal, and fine grain (subtopic level) indices below. Coarse grain alignment indices between grades and schooling cycles range between 0.54 and 0.76 (Table 2), suggesting fairly smooth progression on the prescribed content topics as learners advance through the grades. These measures depict, at topic level, the spiral outlay of the prescribed content i.e. alignment is neither too high to imply a lack of content stretch between grades, nor too low to suggest little to no connection between topics covered in successive grades which might make it hard for learners to keep up with the prescribed material.

At the more detailed fine grain level, alignment measures are lower, as expected. This indicates that within an individual topic, there is greater variation in the subtopic content intended for coverage across grades. At 0.30, the alignment measure between lower primary and P4 implies some considerable progression on prescribed content. For example, the prescribed geometric concepts content shows several new subtopics introduced in P4 with greater emphasis than was the case in the lower primary classes (Appendix B, Figure B.1.). Further, as discussed in Section III.A., it is usual for fine grain analysis to have lower alignment measures than coarse grain, and most of the measures in Table two are at least half of the respective coarse grain alignment measure.

Table 2: Primary mathematics standards - Uganda: Content progression alignment measures

	Grade / Cycle	Coarse grain alignment measures										
		P1	**P2**	**P1-P2**	**P3**	**P1-P3**	**P4**	**P5-P7**	**P5**	**P6**	**P6-P7**	**P7**
Fine grain alignment measures	**P1**	**1.0**	0.76									
	P2	0.56	**1.0**		0.54							
	P1-P2			**1.0**	0.56							
	P3		0.46	0.47	**1.0**		0.66					
	P1-P3					**1.0**	0.63					
	P4				0.32	0.30	**1.0**	0.76	0.72			
	P5-P7						0.40	**1.0**				
	P5						0.36		**1.0**	0.70	0.70	
	P6								0.38	**1.0**		0.57
	P6-P7								0.36		**1.0**	
	P7									0.31		**1.0**

ii) The content in the standardized end-of-primary mathematics examination

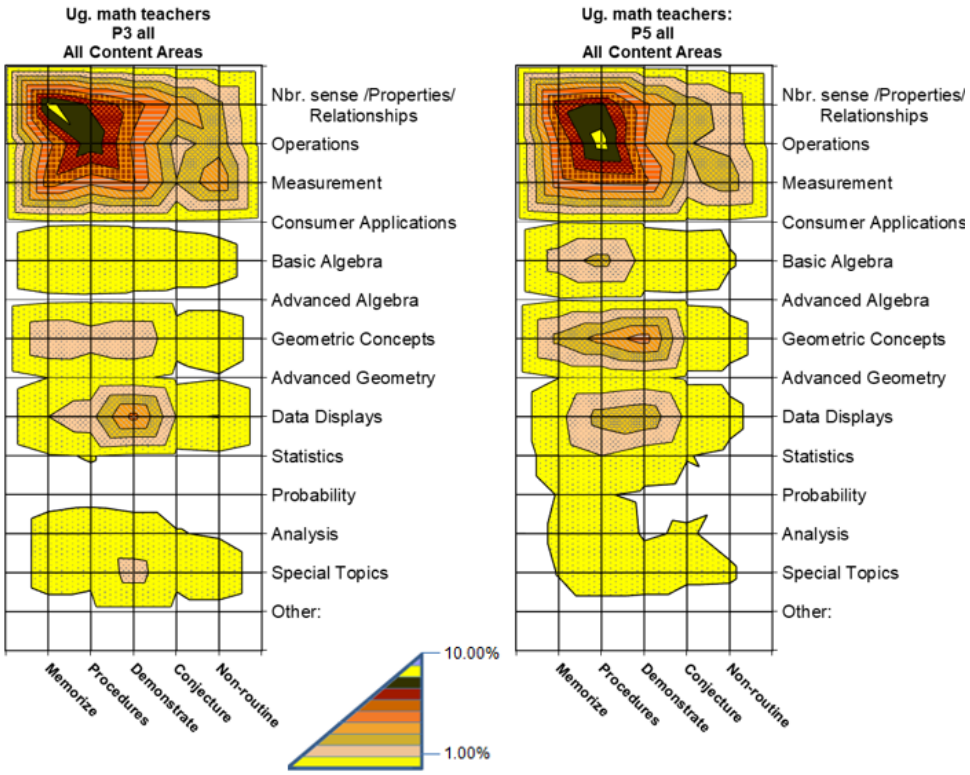
The three consecutive, end-of-primary national mathematics examinations for the years 2013 to 2015 portray high degrees of internal alignment. Coarse grain alignment measures are as high as 0.71, indicating similar topic level content is covered in the exams across years. This is not surprising as these national-level exams are intended to be standardized. At the fine grain level however, the extent of overlap is lower with the alignment indices ranging between 0.34 and 0.50.

To assess the typical coverage of the examinations, we aggregate the three years. While these exams cover the breadth of topics that are prescribed in the curriculum standards, three topics dominate – number sense, operations and measurement (Panel 2 of Figure 4). About two-thirds of total content coverage in the exams is dedicated to these three. Furthermore, measurement is the most emphasized topic, attracting more than one-third of total emphasis. On cognitive skills, perform procedures, demonstrate understanding, and solve non-routine problems (in that order) are most emphasized. Memorize/recall and analyze/conjecture together account for only six percent coverage.

iii) The content of primary mathematics instruction

Turning to teachers' instructional content, we analyze mathematics in grades three and five, for whom two hundred mathematics teachers were surveyed in Uganda. For both sets of teachers (P3 and P5), the most coverage is given to the same three most emphasized topics prescribed in the curriculum standards, accounting for about two-thirds of total coverage (Figure 3). Accounting for about a fifth of total coverage, three other topics follow: basic algebra, geometric concepts and data displays. Both sets of teachers mostly emphasize the first three cognitive skill levels, with ability to perform mathematical procedures the most emphasized.

Figure 3: Uganda mathematics teachers' instructional content – grades 3 and 5



Location-based disparities in instructional practices appear between rural- and urban-based teachers. For both classes, rural-based teachers place much more emphasis on the more foundational skills of number sense and operations than measurement. For the more advanced topics of basic algebra and geometric concepts, rural-based teachers focus less on the two content areas than urban-based teachers.

In terms of cognitive demand, much more emphasis is given to the lowest level, recall, in rural than urban areas. In fact, the emphasis structure on cognitive demand is reversed in the two locations, with rural teachers prioritizing recall followed by perform procedures and demonstrate, whereas urban teachers prioritize the more advanced skills of demonstrate, then perform procedures and lastly recall. In both locations teachers place relatively little emphasis on the two highest-order skills of conjecture and application to non-routine situations.

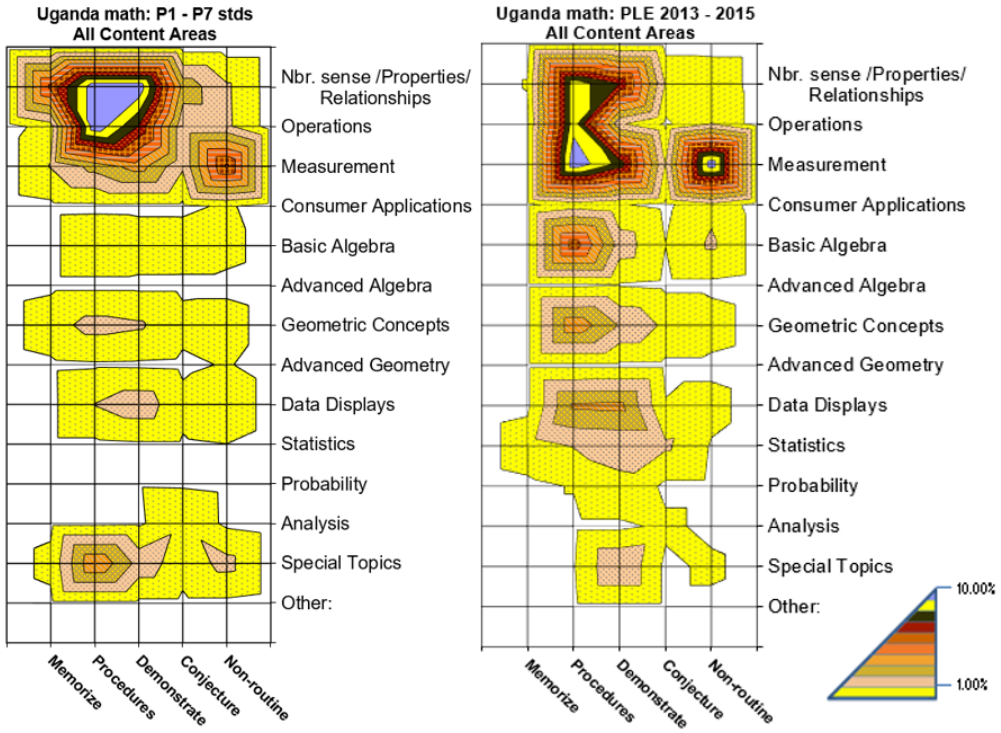
iv) Measuring coherence between mathematics standards and examinations

At the coarse grain (topic) level, the P1-P7 curriculum standards and primary leaving examinations have an alignment score of 0.57, above the threshold for acceptable alignment of 0.50. The same three topics attract the most coverage in both. However, measurement receives disproportionate coverage on the examination – 35% of the coverage, compared with only 15% in the standards. Similarly, number sense and operations receive more than 60% of the coverage in the standards

yet only 30% on the examinations. Some coverage imbalances are also observed mainly on four topics – basic algebra, geometric concepts, data displays and special topics (Figure 4).

Cognitive demand is well-matched between the two components with most coverage given to perform procedures, followed by demonstrate understanding and then application to non-routine situations. However, the exams do place more emphasis on perform procedures and application to non-routine problems, than do the standards.

Figure 4: Coherence between Uganda mathematics standards and exams – coarse grain.



At the more detailed fine grain level the alignment index is relatively low at 0.33, but still is more than half of the coarse grain measure. The alignment indices increase slightly when only content prescribed for the upper primary grades (P4 to P7) is considered; the coarse grain index rises to 0.68 and fine grain to 0.43 (Table 3).

Table 3: Primary mathematics – Uganda: Coherence across standards, examinations and instruction

Grades/ Cycles	Overall Alignment indices							
	Standards vs. Exams		Standards vs. Instruction		Standards vs. Instruction			
					Rural		Urban	
	Fine	*Coarse*	*Fine*	*Coarse*	*Fine*	*Coarse*	*Fine*	*Coarse*
P1-P7	0.33	0.57						
P4-P7	0.42	0.68						
P5-P7	0.43	0.67						
P6-P7	0.41	0.67						
P3			0.37	0.62	0.36	0.61	0.35	0.62
P5	0.37	0.62	0.30	0.64	0.29	0.63	0.30	0.65
P6	0.34	0.60						
P7	0.31	0.59						

Note: The sample for the teacher survey did not allow analysis of coherence between instruction and exams.

v) Measuring coherence between mathematics instruction and standards[9]

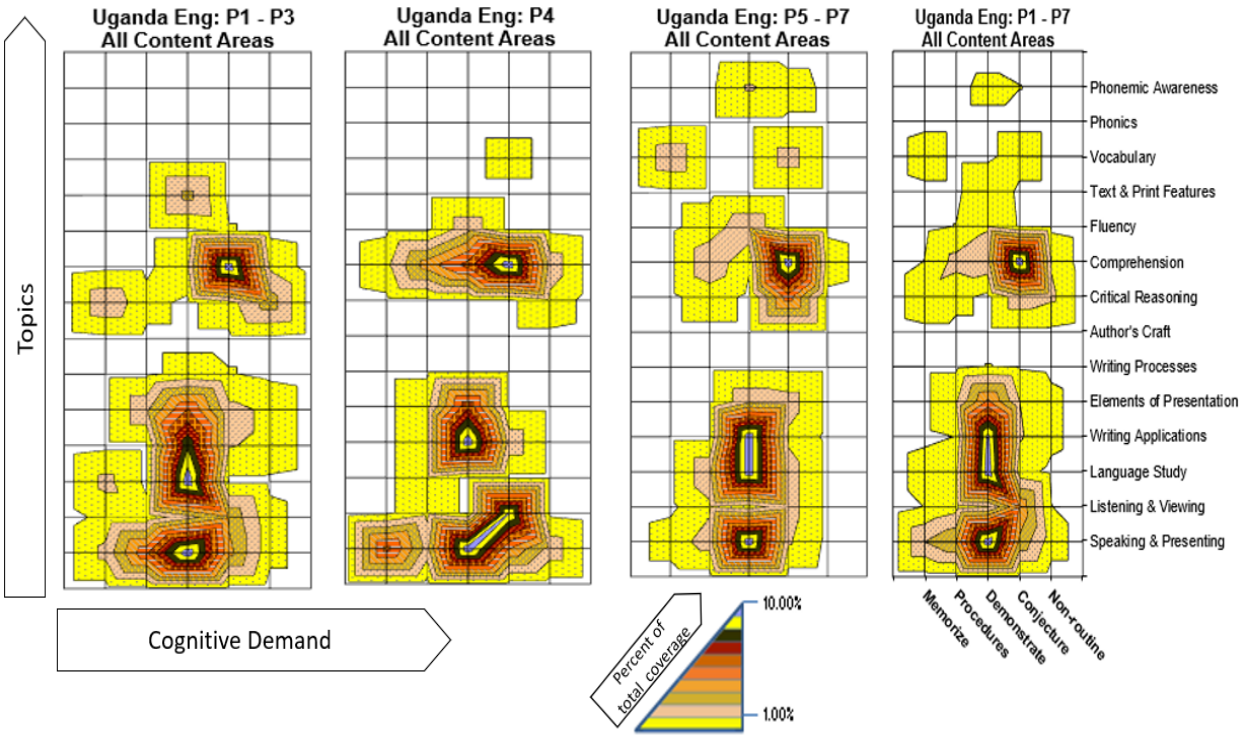
The alignment measures between the curriculum standards (the intended curriculum) and instructional content (the taught curriculum) for both grades (P3 and P5) are relatively high at the coarse grain level (Table 3). For both grades coarse grain alignment passes the threshold for reasonable alignment, at 0.62 (P3) and 0.64 (P5). Fine grain alignment is near or surpasses half of coarse grain measure for both grades.

vi) The content and progression in the primary English curriculum standards

Covering two-thirds of the overall focus, most coverage in the Ugandan English standards goes to three topics – speaking and presenting, comprehension, and language study (Figure 5). Attracting ten percent of total coverage, writing applications follows in a distant fourth place. Surprisingly, the standards completely omit two would-be critical foundational language topics in lower primary, i.e. phonics and phonemic awareness, only for the latter to show up weakly in the penultimate grade (P6). The eight topics that are not in the top three are covered sporadically, in a seeming attempt to ensure all topics are covered, but without a clear progression from grade to grade. Many follow a "drop and reinstate" pattern in which they are covered one year, dropped for a year or two, and reinstated, often at a higher level of cognitive demand. For example, while nine topics are prescribed for coverage in P4, six of these are dropped in the P5 standards and then all are reinstated in P6 together with another completely new topic making ten topics in this grade. Then the topics are reduced back to nine in P7.

[9] Because the teacher instruction sample covers only grades 3 and 5, and the primary leaving exam occurs after grade 7, we do not report the coherence between teacher instruction and the exam as interpretation would be difficult in terms of appropriate expectations for coherence of grades 3 and 5 instruction with the PLE.

Figure 5: Primary English curriculum standards.



For cognitive demand, the primary curriculum standards are skewed towards demonstrate, followed by conjecture which accounts for a quarter of the emphasis. Perform procedures follows in a distant third place with about a tenth of the total emphasis. This emphasis structure is steadily maintained in the curriculum standards for all the seven classes.

The language content coverage in P1 seems to focus on facilitating young learners to freely express themselves among peers as they try to fit into a relatively new environment, with much focus on public speaking and oral presentation. More than half the emphasis on cognitive demand in this grade is directed at demonstrating understanding. Conjecture and memorize attract equal emphasis, jointly explaining about one-third of total cognitive skills emphasis.

Table 4 shows the alignment measures between grades and learning cycles, with coarse grain alignment indices (range 0.06 – 0.63) appearing above the main diagonal and fine grain measures (range 0.05 – 0.43) below. The wide ranges in the indices are driven by P5, which is an outlier and poorly aligned with the others.[10]

Overall, the content progression alignment measures for English are lower than for mathematics.[11] At the topic level, the earlier noted alternating (drop-and-reinstate) content coverage structure

[10] The P5 curriculum standards focus almost entirely on just one topic: language study.

[11] English and mathematics are very different subject domains, with math being much more sequential and tighter in its descriptive language. English is a more diffuse subject where students are typically engaged in multiple aspects of the content domain at one time, and so the descriptive language for it tends to be less tight and sequential. As a result, it is usual for English to have more disperse progression and lower progression alignment than mathematics.

serves as a key reason for the low measures. Lower fine grain measures reflect a higher degree of content variation between consecutive grades or cycles, but most surpass half of their coarse grain equivalent.

Table 4: Primary English standards - Uganda: Content progression alignment measures

	Grade / Cycle	Coarse grain alignment measures										
		P1	**P2**	**P1-P2**	**P3**	**P1-P3**	**P4**	**P5-P7**	**P5**	**P6**	**P6-P7**	**P7**
Fine grain alignment measures	**P1**	**1.0**	0.46									
	P2	0.28	**1.0**		0.48							
	P1-P2			**1.0**	0.44							
	P3		0.18	0.25	**1.0**		0.63					
	P1-P3					**1.0**	0.61					
	P4				0.43	0.34	**1.0**	0.48	0.06			
	P5-P7						0.32	**1.0**				
	P5						0.05		**1.0**	0.11	0.14	
	P6								0.07	**1.0**		0.52
	P6-P7								0.13		**1.0**	
	P7									0.34		**1.0**

vii) The content in the standardized end-of-primary English examinations

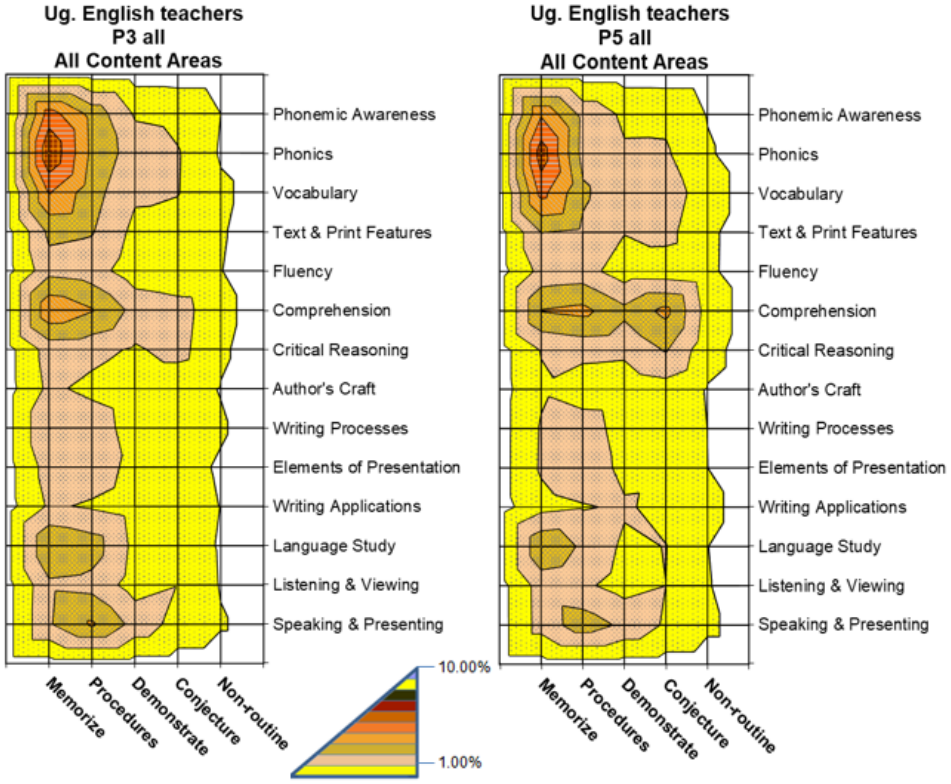
The three successive end-of-primary English exams for the years 2013 to 2015 are highly aligned with each other. Fine grain alignment ranges from 0.66 and 0.67 and coarse grain from 0.74 and 0.85.

Only four of the fourteen topics covered in the curriculum standards appear on the exams for each of the three years – vocabulary, comprehension, critical reasoning and language study (Figure 7). Two topics, comprehension and language study, constitute three quarters of the total content covered on the exams. Cognitive demand in the exams mirrors the structure in the curriculum standards, with demonstrate and conjecture highly dominant.

viii) Primary school English instruction

Teacher instructional coverage reveals a wide spread of content coverage with all fourteen English topics covered in both third and fifth grade regardless of the prescriptions laid out in the curriculum standards – no areas are excluded during instructional delivery (Figure 6). Although some topics can be identified as receiving greater emphasis in the classroom the overall picture for both classes is that most topics receive relatively equal coverage. The emphasis range for both sets of teachers falls between three and thirteen percent for each topic. This is quite different from the mathematics instruction in which just three topics receive two-thirds of the coverage by teachers. The two most emphasized cognitive demand levels are memorize and perform procedures, together representing about 60% of the focus while the other three levels share the remainder – with non-routine application being the least emphasized.

Figure 6: Uganda English teachers' instructional content – grades 3 and 5



No systematic disparities are identifiable relating to topic coverage between rural and urban teachers. However, for both grades, the focus on cognitive demand differs significantly between rural and urban teachers. Similar to mathematics instruction, while rural teachers emphasize the lower order skill of memorize/recall, urban teachers' instructional practices place less emphasis on memorize/recall in preference to the higher order perform procedures and conjecture/analyze skills.

ix) Measuring coherence between English standards and examinations

Only four of the fourteen topics in the curriculum standards feature in the standardized national exams (Figure 7). Two of the four, vocabulary and critical reasoning, jointly receive only 6% of the coverage in the curriculum standards while receiving 25% of the coverage in the examination. Further, two high-emphasis topics in the curriculum standards are completely missing from the exams – writing applications and speaking & presenting, which jointly account for forty percent of the content in the primary English standards.[12] On cognitive demand, the two are more aligned, giving large emphases to demonstrate and conjecture.

[12] The omission of these topics from the exam may be unsurprising, as they are harder and more expensive item types to develop and score. This raises important questions as these are important topics for children to master, and accordingly feature in the curriculum standards, but if they are not assessed will they still be taught.

Reflecting these disparities, coarse grain alignment is only 0.36 and fine grain is 0.26 (Table 5). By contrast, the grade 5 curriculum standards, when analyzed alone, has a coarse grain alignment with the exam of 0.42. This is still less than the 0.50 considered well-aligned, but suggests the standards in that grade are more tailored to exam content than other grades. In this grade one topic dominates the standards: language study.

Figure 7: Coherence between standards and exams – coarse grain.

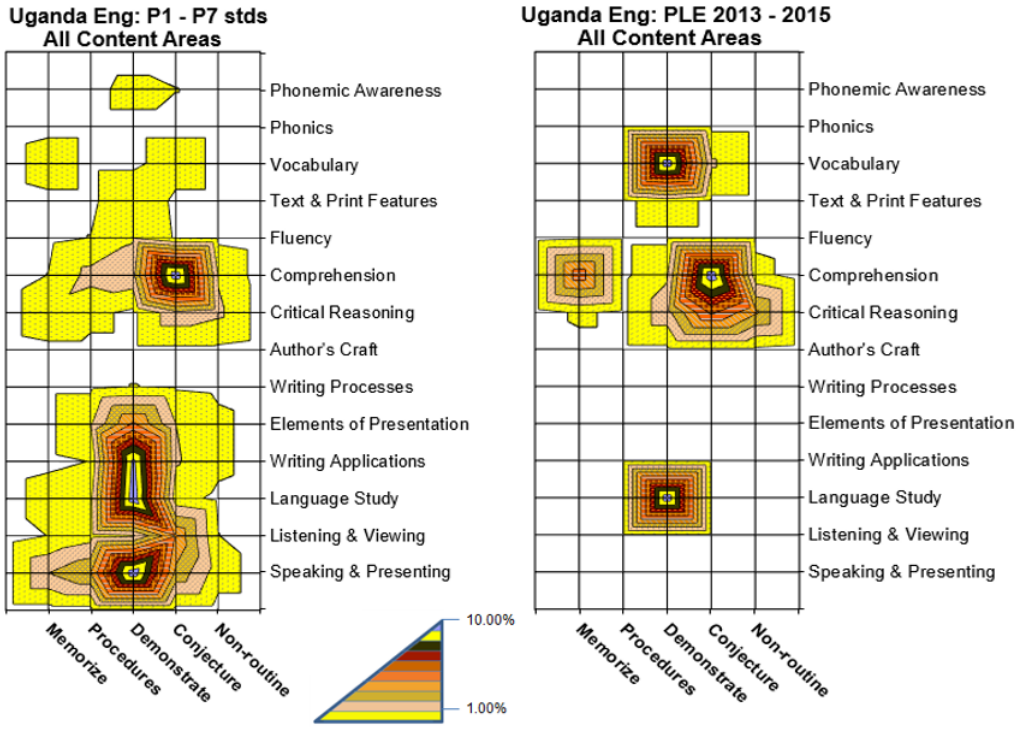


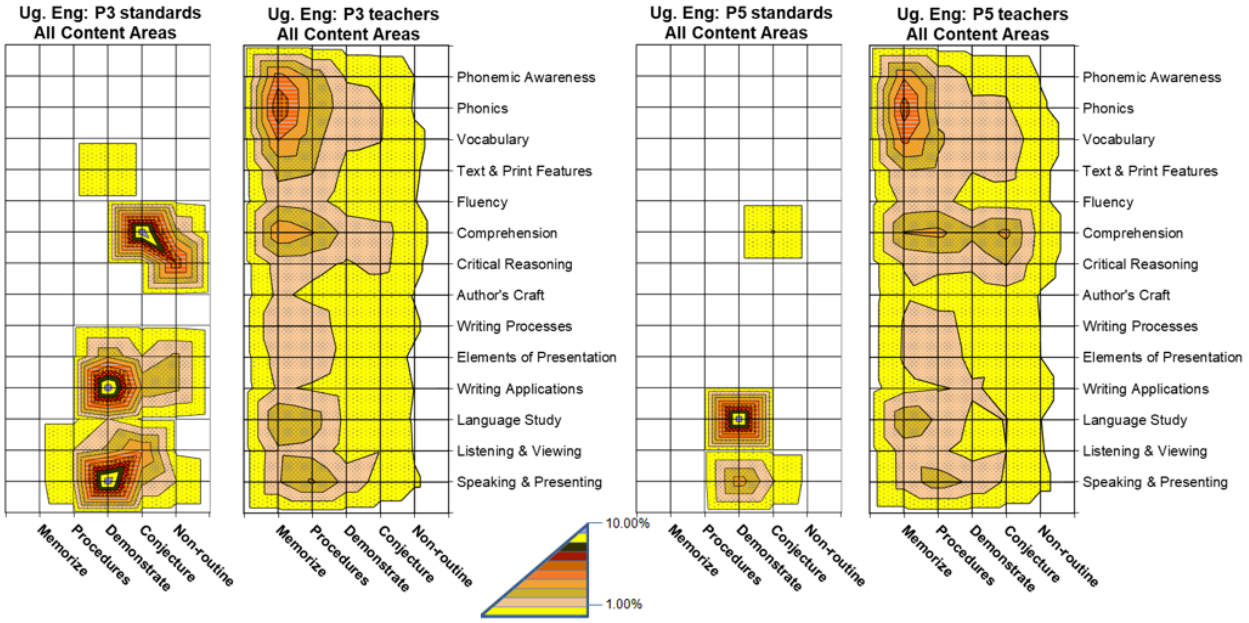
Table 5: Primary English – Uganda: Coherence across standards, examinations and instruction

Grades/ Cycles	Overall Alignment indices							
	Standards vs. Exams		Standards vs. Instruction		Standards vs. Instruction			
					Rural		Urban	
	Fine	*Coarse*	*Fine*	*Coarse*	*Fine*	*Coarse*	*Fine*	*Coarse*
P1-P7	0.26	0.36						
P4-P7	0.26	0.37						
P5-P7	0.36	0.52						
P6-P7	0.17	0.35						
P3			0.05	0.15	0.04	0.15	0.05	0.14
P5	0.42	0.42	0.01	0.05	0.01	0.05	0.01	0.06
P6	0.14	0.25						
P7	0.19	0.41						

x) Measuring coherence between English instruction and standards

For both classes, P3 and P5, the alignment measures for instruction and standards are extremely low, an indication of differing content emphasis between the curriculum standards and actual instruction in the classrooms (Figure 8). The alignment is poor on all three dimensions, topic/subtopic coverage, focus on cognitive skills development, and amount of emphasis per content area and cognitive skill level covered. Coarse grain alignment between P3 teacher instruction and standards is 0.15, while for P5 it is 0.05. Fine grain is even lower at 0.05 and 0.01 respectively.

Figure 8: Uganda English – coherence between standards and instruction – grades 3 and 5



Despite the curriculum standards' focus on demonstrate and conjecture abilities, P3 teachers mostly prioritize the first two low-level cognitive categories, memorize and perform procedures. In P5, the standards prescribe most emphasis to be demonstrate, but teacher's instruction primarily emphasizes memorize and perform procedures skills. P5 standards also place most emphasis on language study, while teachers allocate less than a tenth of total coverage to this topic.

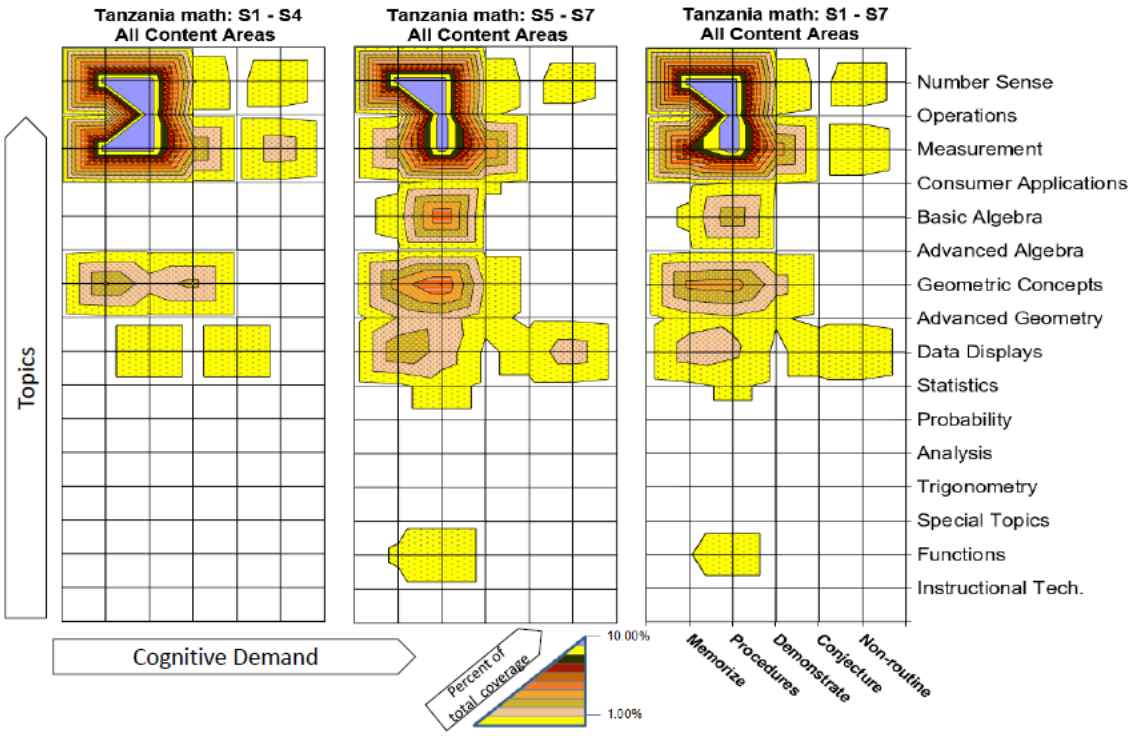
IV.B. Instructional Coherence in Tanzania

i) The content and progression of the primary mathematics curriculum standards

The ten topics covered by the Tanzania primary mathematics curriculum standards can be categorized into three distinct sets. First, three topics constitute most of the coverage: number sense, operations and measurement (Figure 9). Four-fifths of total coverage is attributed to these three content areas. Second, another three topics follow at some distance jointly accounting for seventeen percent of total emphasis. These are basic algebra, geometric concepts and data displays. The final four topics are the least emphasized and account for a combined total of less than four

percent. Five topics are covered during the first four years. In upper primary (S5 – S7[13]) coverage is broader and more complex with ten topics.

Figure 9: Primary mathematics curriculum standards.



Ninety percent of the cognitive demand emphasis is on developing the two lower-order thinking skills of memorize/recall and perform procedures. This cognitive demand structure is true for all the prescribed topics and grades in the cycle.

Analysis of content progression across grades shows that coarse grain alignment ranges from 0.51 to 0.79, suggesting a moderately smooth stretch between topics prescribed in each progressive grade (Table 6).

At the more specific subtopic level, the progression alignment measures are lower, ranging from 0.17 to 0.54, and decline in progressive years. This indicates increasing variation in the specific subtopics prescribed, particularly in the upper primary years. The very low fine grain progression measure of 0.17 between S6 and S7 is because four number sense subtopics are introduced for the first time in S7.

[13] In Tanzania, grade levels are referred as standards, so "S" abbreviates standards. "Standard 5" becomes "S5".

Table 6: Primary mathematics standards - Tanzania: Content progression alignment measures

	Grade / Cycle	Coarse grain alignment measures										
		S1	**S2**	**S3**	**S4**	**S1-S4**	**S5**	**S6**	**S7**	**S5-S6**	**S6-S7**	**S5-S7**
Fine grain alignment measures	**S1**	**1.0**	0.79									
	S2	0.54	**1.0**	0.55								
	S3		0.41	**1.0**	0.51							
	S4			0.38	**1.0**		0.65					
	S1-S4					**1.0**	0.61					0.66
	S5				0.38	0.34	**1.0**	0.60			0.64	
	S6						0.31	**1.0**	0.59			
	S7							0.17	**1.0**			
	S5-S6									**1.0**		
	S6-S7						0.34				**1.0**	
	S5-S7					0.35						**1.0**

ii) The content in the standardized end-of-primary mathematics examinations

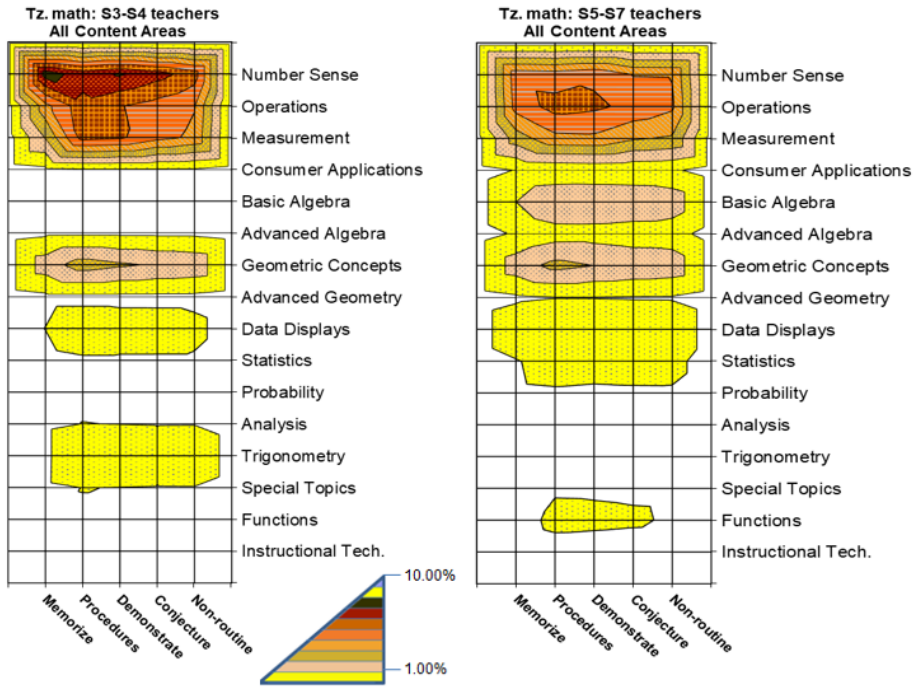
The mathematics exams from 2013-2015 are highly aligned across years, with coarse grain alignment ranging from 0.74 to 0.85 and fine grain alignment ranging from 0.47 to 0.52. Aggregating the content across the three exams, three topics dominate: operations, measurement and basic algebra. These three explain more than four-fifths of the content covered. On cognitive demand, the main focus in the exams is learners' abilities to perform procedures, followed at some distance by demonstrate understanding. The higher order skills of conjecture and application to non-routine problems are not covered.

iii) The content and progression of primary mathematics instruction

In the five grades (S3 – S7) for which teacher instructional data was collected[14], three topics are most emphasized in the classroom: number sense, operations and measurement (Figure 10). A fourth topic follows at some distance – geometric concepts. In the two lower primary grades (S3 – S4), number sense dominates topic level coverage with a third of total coverage. In upper primary grades operations is most emphasized with twenty-five percent of total coverage. Instruction in upper primary is also broader with the introduction of consumer applications, basic and advanced algebra, statistics, and functions.

[14] Teachers for the first two primary-level classes were not included in the teacher survey in Tanzania.

Figure 10: Tanzania mathematics – Teachers' instructional content coverage in lower and upper primary grades



For cognitive demand, teachers attach almost equal importance to all five cognitive categories. This is true for lower and upper classes, as well as rural and urban-based teachers.

Alignment of content from grade to grade is high, ranging from 0.49 to 0.82 at fine grain and 0.76 to 0.93 at coarse grain (Table 7). Such high indices reveal little to no variation between the instructional content experiences learners are exposed to as they move up classes. Teachers tend to cover all the listed subtopics without regard to what is prescribed for which grade. Although this could be a good practice if it meant that students have many opportunities to master the same content, it can also mean that instruction is devoid of depth due to the breadth of topics and subtopics being covered in a given class.

The progression alignment measures are higher for rural-based teachers, suggesting that their students receive fewer opportunities to learn and master the prescribed primary mathematics content, due to the breadth of topics covered each year.

Table 7: Primary mathematics teachers - Tanzania: instructional progression alignment measures

	Classroom instruction					
	All		***Rural***		***Urban***	
	Fine	*Coarse*	*Fine*	*Coarse*	*Fine*	*Coarse*
S3 vs. S4	0.62	0.80				
S3-S4 vs. S5	0.55	0.82				
S4 vs. S5	0.54	0.76				
S5 vs S6	0.78	0.89	0.79	0.94	0.62	0.78
S6 vs. S7	0.82	0.93	0.79	0.94	0.70	0.84
S3-S4 vs. S5-S7	0.49	0.80				
S5 vs S6-S7	0.79	0.91	0.79	0.94	0.66	0.83

iv) Measuring coherence between mathematics standards and examinations

Nearly all the content areas covered in the curriculum standards also feature in the end-of-primary examinations (Figure 11). There are, however, large differences in emphasis for both topics and cognitive demand. At 0.44, alignment to the full primary mathematics standards falls short of the 0.50 threshold for acceptable alignment (Table 8). The alignment to just the highest two grades, S6 and S7, the years right before the exam, reaches 0.50.

Two major findings explain the poor alignment between the standards and exams. First, major topic level emphasis disparities are prominent. Operations and basic algebra together contribute only 15% of the content in the primary curriculum standards, yet half the content on the exams is from these two topic areas. On the other hand, only seven percent of the content on the exams is attributable to number sense and geometric concepts, which comprise half of the content in the curriculum standards. Second, while both the standards and the exams prioritize the cognitive skill of perform procedures, a large portion of the exams also targets demonstrate understanding, a skill level that is among the least emphasized in the curriculum standards. Instead, the standards focus a third of overall emphasis on memorize/recall, a cognitive demand level that is completely left off the exams.

Figure 11: Tanzania mathematics – coherence across standards, instruction and exams

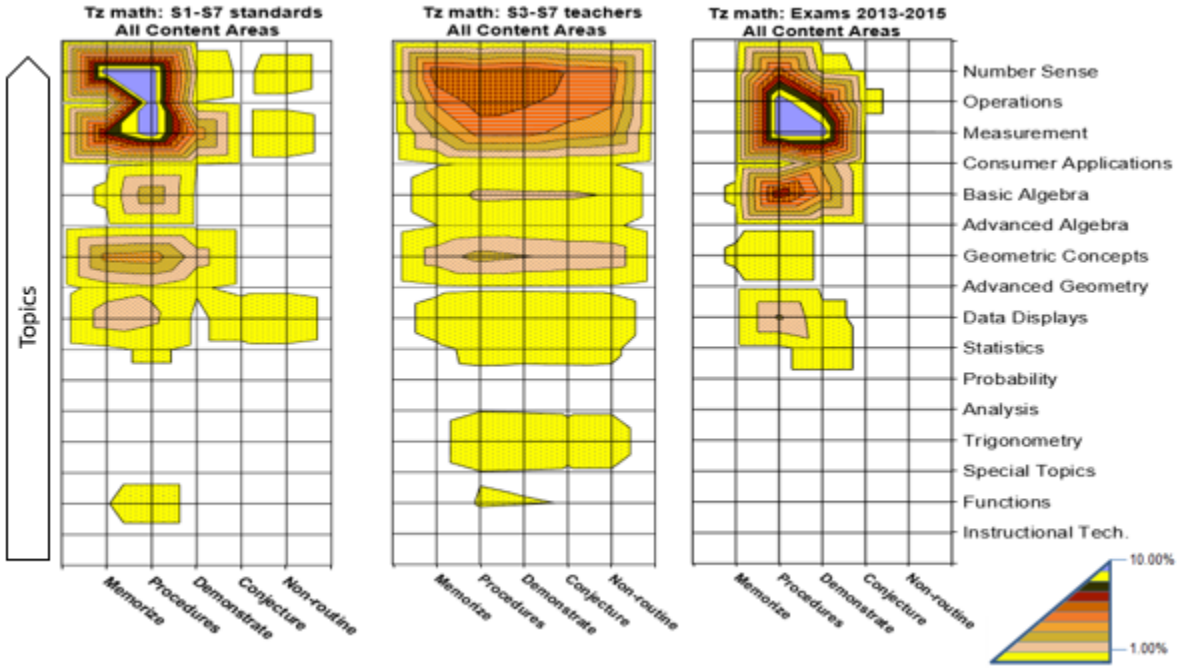


Table 8: Primary mathematics - Tanzania: Coherence across standards, examinations and instruction

	Overall Alignment indices					
	Standards vs. Exams		Standards vs. Instruction		Instruction vs. Exams	
	Fine	*Coarse*	*Fine*	*Coarse*	*Fine*	*Coarse*
S1-S7	0.29	0.44				
S3-S7	0.29	0.47	0.26	0.44	0.15	0.33
S5-S7	0.32	0.48	0.22	0.41	0.15	0.33
S6-S7	0.34	0.50	0.14	0.34	0.14	0.32
S3			0.20	0.35		
S4			0.16	0.42		
S5			0.16	0.40		
S6			0.09	0.33		
S7	0.21	0.51	0.08	0.27	0.15	0.33

v) Measuring coherence between mathematics instruction and standards

Teachers' instructional content in the classroom is very poorly aligned with the curriculum standards. Analyzing the standards and instruction for S3 – S7 (for which we have teacher instructional data) we find fine grain alignment ranging between 0.08 and 0.26, and coarse grain alignment between 0.27 and 0.44 (Table 8). There are no significant location-based differences.

Although similar topics are covered by both instruction and the curriculum standards, their emphases differ considerably (Figure 11). First, two topics in the standards, number sense and

measurement, receive two-thirds of total content coverage, and yet less than half of teachers' instructional focus goes to these two content areas. Conversely, with just a tenth of the total emphasis in the curriculum standards, teachers allocate twenty five percent coverage to operations. Finally, the disparities in levels of cognitive demand are also quite large. While the standards overwhelmingly emphasize the first two lower-order cognitive categories, teachers allocate their instructional focus almost equally across all the five cognitive demand categories.

vi) Measuring coherence between mathematics instruction and examinations

The alignment of classroom instruction and exams[15] are low, at 0.15 fine grain and 0.33 coarse grain (Table 8). The content mismatches originate primarily from differences on topic level emphasis. Major discrepancies are present for all five highly emphasized topics – number sense, operations, measurement, basic algebra and geometric concepts. The exams prioritize operations and measurement, while teachers' instruction spreads coverage more evenly across the five (Figure 11).

At the subtopic level, because teachers allocate almost equal emphasis to all subtopics within each topic they cover (regardless of the prescribed subtopic coverage in the curriculum standards) they place disproportionately little emphasis on the few subtopics that are most emphasized on the exams. The breadth of what teachers cover differs from the depth of focus on a smaller subset of topics/subtopics in the exams. This discrepancy is most prominent in the upper primary grades (S5 – S7) and best illustrated by the topics number sense, consumer applications, basic algebra, and geometric concepts.

On cognitive demand, the exams are entirely focused on perform procedures and demonstrate understanding, while teachers assign nearly equal emphasis to all five levels of cognitive demand. This means less than half the cognitive skills development by teachers coincides with the focus of the exams.

vii) The content and progression in the primary English curriculum standards

Content from three topics dominates Tanzania's primary English curriculum standards, accounting for 80% of total coverage (Figure 12): comprehension, language study, and speaking and presenting. In upper primary grades S5 to S7, emphasis on two additional topics, writing applications and critical reasoning, increases substantially and speaking and presenting declines. Phonemic awareness is completely omitted from the curriculum standards. Topical coverage throughout the seven grades uses a drop-and-reinstate structure as in Uganda. A topic may be covered one year, dropped for a year or two, then reappear.

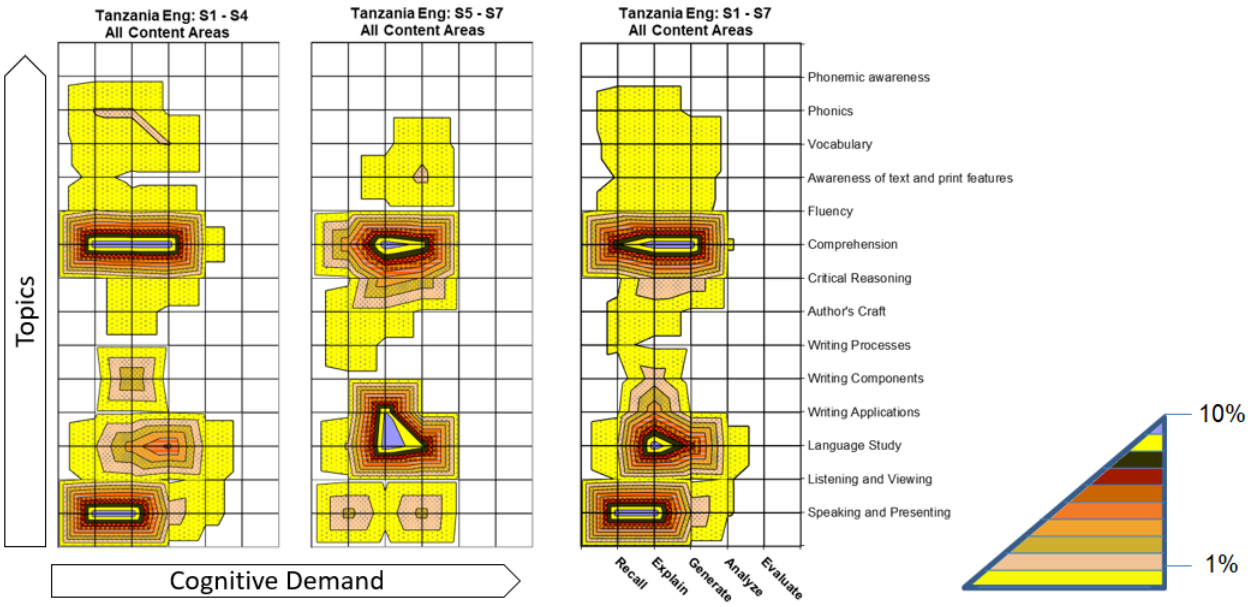
More than half the emphasis for cognitive demand is on explain, with the remaining emphasis equally divided between recall and generate[16]. The higher order analyze and evaluate skills are not

[15] Comparing the available aggregated classroom instructional content (S3 – S7) to the aggregated three-year PSLEs for 2013-2015.

[16] English in Tanzania uses different labels for the levels of cognitive demand, but conceptually they refer to the same difficulty level. Explain is the equivalent of perform procedures; generate of demonstrate understanding; analyze of conjecture; and evaluate is the equivalent of apply to non-routine problems.

emphasized in the primary standards. This cognitive demand structure in the curriculum standards is consistent throughout all seven grades.

Figure 12: Tanzanian Primary English curriculum standards.



In standard one, more than two-thirds of the content in the English curriculum standards belongs to only one topic, speaking and presenting. Surprisingly, the foundational language topic phonics is not prescribed until standard two. Accounting for about twenty percent of total coverage, comprehension is the next most emphasized content area in standard one, and then language study with ten percent. The subtopics covered under speaking and presenting in standard one suggest an intention to facilitate young learners to fit into their new school surroundings with a focus on clear and confident self-expression and presentation to others[17]. The two lowest order cognitive skill levels, recall and perform procedures, constitute the entire focus on development of learner performance abilities in this grade.

In the lower primary years, S1 – S4, alignment from year to year is low, ranging from 0.09 – 0.30 fine grain and 0.24 – 0.57 coarse grain (Table 9). For the upper grades S5 – S7 progression alignment is moderately higher, ranging from 0.35 – 0.43 fine grain and 0.47 – 0.65 coarse grain. The lower measures for lower primary grades suggest relatively steeper stretches in between-grade content prescribed for coverage in these classes. At 0.09 fine grain and 0.24 coarse grain, the progression from S2 to S3 suggest the highest level of intended content variation.[18] Such low

[17] The subtopics attracting most emphasis for the topic speaking and presenting in standard one are: public speaking & oral presentation, demonstrating confidence, situational & cultural norms of expression, conversation & discussion, debate & structure of arguments, dramatics & creative interpretation, and interviewing.

[18] For example, two topics, vocabulary and author's craft, are introduced for the very first time in S3 with three subtopics to be covered. In addition, five new subtopics are introduced under the two most emphasized S3 topics of comprehension and language study.

alignment, and high variation in topic coverage from grade to grade, could make it difficult for children to keep pace, especially in the early schooling years when learning foundational skills.

Table 9: Primary English standards - Tanzania: Content progression alignment measures

	Grade / Cycle	Coarse grain alignment measures										
		S1	**S2**	**S3**	**S4**	**S1-S4**	**S5**	**S6**	**S7**	**S5-S6**	**S6-S7**	**S5-S7**
Fine grain alignment measures	**S1**	**1.0**	0.46									
	S2	0.30	**1.0**	0.24								
	S3		0.09	**1.0**	0.57							
	S4			0.26	**1.0**		0.47					
	S1-S4					**1.0**	0.42					0.45
	S5				0.35	0.35	**1.0**	0.65			0.66	
	S6						0.43	**1.0**	0.54			
	S7							0.39	**1.0**			
	S5-S6									**1.0**		
	S6-S7						0.46				**1.0**	
	S5-S7					0.36						**1.0**

Relatively low progression alignment measures characterize the transition from lower primary to upper primary, from S4 to S5. The variation is driven by differences in both topic coverage and cognitive skill focus. Two topics that had been dropped from the curriculum standards after S1 and S3 are reinstated in S5, text and print features, and author's craft. Three S4 topics are dropped from the S5 standards – fluency, writing components and listening & viewing. Although both comprehension and language study remain the most dominant topics in S5 as was the case in S4, the two topics attract much higher emphasis in S5 and there is a major reorientation in their cognitive skills focus, moving away from the mid-level skill of generate in S4 backwards to the lower order skill of explain in S5.

viii) Content in the standardized end-of-primary English examinations

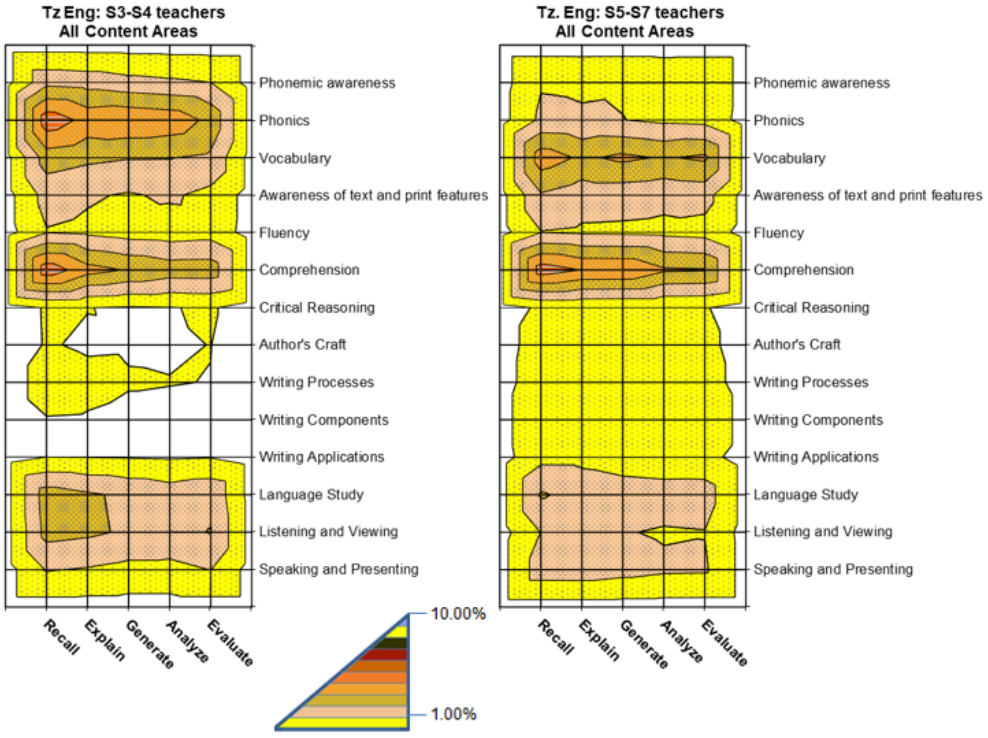
Consistent with earlier subsections, we aggregate all content from the exam items appearing on three successive end-of-primary national examinations, for the three years 2013 – 2015 (see panel 3 of Figure 14). These exams are highly aligned across years, with alignment indices ranging from 0.54 to 0.65 at fine grain and 0.72 to 0.83 at coarse grain.

As was observed in Uganda's case, a striking feature of the English exams is the very high level of topic selectivity for inclusion in these exams. In Tanzania's case, only three topics are covered in the exams: vocabulary, comprehension and language study. On cognitive demand, all emphasis in the exams is targeted at the lowest order skill of recall/memorize.

ix) Content and progression of primary English instruction

Four of the fourteen English topics account for the majority of teachers' instructional content during grades S3 – S7 (the grades for which we have teacher instructional data): comprehension, vocabulary, phonics and language study. Of the four, phonics and comprehension attract most emphasis in lower primary (S3 – S4) while vocabulary replaces phonics in the top-two in upper primary (S5 – S7) (Figure 13). More topics are also covered in the upper primary classes, and overall emphasis therefore is more spread, indicating more breadth of topics but less depth in the upper primary years. Although phonemic awareness was completely omitted in the curriculum standards, teachers include this topic in the content delivered in classrooms.

Figure 13: Tanzania English – Teachers' instructional content coverage in lower and upper primary grades



Teachers' instructional practices reflect an almost equal distribution of emphasis across the five cognitive demand categories. While recall and perform procedures attract slightly higher focus in both lower and upper primary, the differences with the other three cognitive levels are negligible. When aggregated across grades, no clear rural-urban instructional disparities are observed.

Progression of teacher instruction is highly aligned from grade to grade (Table 10). With indices ranging from 0.61 to 0.70 fine grain and 0.71 to 0.81 coarse grain, the delivered instructional content across subsequent grades is more similar than varied. Contrary to the content coverage stipulated in the curriculum standards, teachers are covering more topics in each grade.

Additionally, for each topic covered, they tend to cover all the listed subtopic content regardless of the curriculum coverage requirements.

Table 10: Primary English teachers - Tanzania: instructional progression alignment measures

	Classroom instruction					
	All		***Rural***		***Urban***	
	Fine	*Coarse*	*Fine*	*Coarse*	*Fine*	*Coarse*
S3 vs. S4	0.65	0.74				
S4 vs. S5	0.62	0.74				
S5 vs S6	0.64	0.75				
S6 vs. S7	0.70	0.81	0.72	0.83	0.57	0.70
S3-S4 vs. S5	0.61	0.71				
S3-S4 vs. S5-S7	0.68	0.77				
S5 vs S6-S7	0.68	0.77				

Note: Sample sizes for teacher instruction disaggregated by grade and location are only large enough for analysis of progression from S6 to S7.

At 0.62 fine grain and 0.74 coarse grain, the transition from S4 to S5 signifies the highest between-grade variation in instructional content. One S4 topic, critical reasoning, is completely dropped in S5 and significant reductions in instructional emphasis are made for three topics i.e. phonics, comprehension and listening & viewing. On the other hand, significant increases in instructional emphasis are made for both speaking & presenting and author's craft. But even with these changes this is a highly aligned transition.

When disaggregated by grade level, there are significant rural-urban differences in the progression alignment measures between S6 and S7. Instruction between S6 and S7 is more highly aligned in rural areas than in urban areas. Looking at S7 in particular, two topics, vocabulary and comprehension, account for one-third of instructional coverage for rural teachers and yet more than half of the coverage for urban teachers.

x) Measuring coherence between English standards and examinations

Only three of the fourteen curriculum standards topics are covered in the exams: vocabulary, comprehension and language study (Figure 14), translating into the very low coherence measures shown in Table 11. Coarse grain alignment is only 0.11, and fine grain alignment is 0.01. Each of the three topics in the exams garners much more emphasis in the exams than the standards. Although speaking and presenting is the second most highly emphasized topic in the curriculum standards, it is completely missing in the exams.

Figure 14: Tanzania English – coherence across standards, instruction and exams.

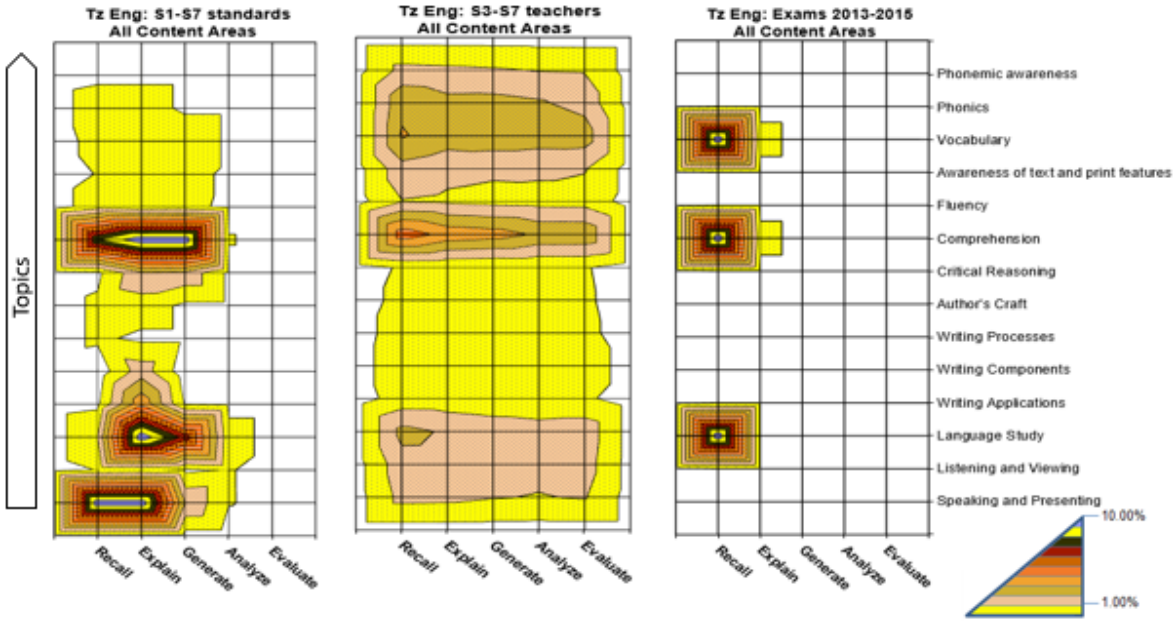


Table 11: Primary English - Tanzania: Coherence across standards, examinations and instruction

	Overall Alignment indices					
	Standards vs. Exams		Standards vs. Instruction		Instruction vs. Exams	
	Fine	*Coarse*	*Fine*	*Coarse*	*Fine*	*Coarse*
S1-S7	0.01	0.11				
S3-S7	0.01	0.07	0.11	0.32	0.05	0.11
S5-S7	0.00	0.06	0.07	0.27	0.05	0.11
S6-S7	0.00	0.05	0.06	0.24	0.06	0.12
S3			0.05	0.30		
S4			0.04	0.27		
S5			0.04	0.27		
S6			0.03	0.16		
S7	0.00	0.10	0.05	0.22	0.04	0.10

The content coverage discrepancies are even more severe at the fine grain subtopic level. The English standards and exams are so poorly aligned all the fine grain alignment measures round to zero.

On cognitive demand, the exams focus entirely on the lowest order skill of recall while the curriculum standards prioritize explain, followed by recall and generate.

xi) Measuring coherence between English instruction and standards

Ranging from 0.03 – 0.11 fine grain and 0.16 – 0.32 coarse grain, the alignment between instruction and standards is also very low (Table 11). These low alignment measures persist through all grades. There are no significant rural-urban disparities.

While all thirteen topics in the standards are also covered during classroom instruction, the amount of emphasis accorded to each topic differs significantly between the two. For example, the overwhelming focus on comprehension and language study in the standards is not matched during instruction[19] (Figure 14). At the fine grain level, large discrepancies also exist between the numbers of prescribed and covered subtopics.

In terms of cognitive demand, the standards prescribe a majority of the focus on perform procedures and demonstrate understanding. In practice however, teachers are distributing their efforts almost equally across the five cognitive demand categories, with some slight edge given to recall/memorize.

xii) Measuring coherence between English instruction and examinations

The alignment measures between instruction and the exams are very low, pointing to significant incoherence. Fine grain alignment ranges from just 0.04 to 0.06, and even at the coarse grain level only reaches 0.12 (Table 11). Low alignment persists regardless of teacher location (rural or urban).

Only three of the fourteen topics covered in the classroom are covered in the exams. For each of the three exam topics therefore, there exists a large difference in the amount of emphasis allotted between instruction and exams. Comprehension provides the most extreme example, with a twenty-five percentage point greater emphasis on the exams than during instruction. Finally, with the exams focusing only on recall/memorize skills, a major imbalance in cognitive demand exists between the two, contributing immensely to the low alignment measures.

V. Discussion and Conclusion

Our analysis of curriculum standards, examinations, and teachers' instructional content have four key findings related to the coherence of instructional components in Uganda and Tanzania, with lessons that are broadly relevant for education systems aiming to improve instructional coherence for learning. While the exact structure of standards, examinations, and instructional content, and sources of incoherence differ somewhat across countries, in both countries incoherence is more common than coherence.

The strength of a standards-based education system rests on the robustness of the standards. Curriculum standards prescribe the content deemed necessary and appropriate for children to learn and the progression through which they should learn it, and teachers' instruction and examinations

[19] While only 28% of emphasis during instruction goes to these two topics, they account for 70% of the emphasis in the curriculum standards. In P5, the combined difference in the amount of emphasis given to these two topics between the curriculum standards and instruction amounts to a huge 52% points.

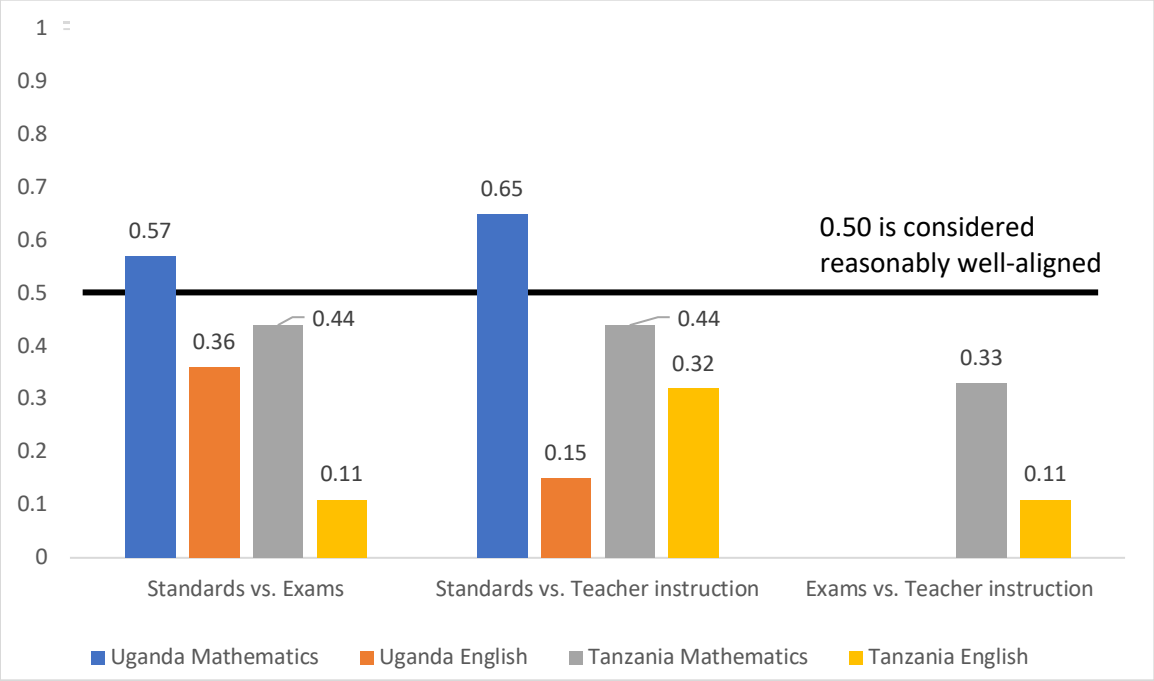
should follow from the standards. Therefore, the first question our study addresses is whether the curriculum standards in Uganda and Tanzania are structured in a way that promotes learning.

Our analysis of the content, pace, and sequencing of the curriculum standards suggests there is room for improvement in making them coherent for learning. In both countries, the alignment of mathematics content from grade to grade encouragingly shows consistency for students while also indicating a level of stretch as new topics and levels of cognitive demand are added. In Tanzania, limited attention is given to higher order cognitive skills, however, which may leave children lacking full conceptual mastery. In English, progression alignment is lower in both countries, and important foundational topics are sometimes missing. The Ugandan standards omit the foundational topics of phonics and phonemic awareness from grades one through four, likely expecting that children have gained these skills in preschool. With less than 30% of children attending preschool however, these topics may need to be included in the primary school standards to ensure all children have the opportunity to learn them (Atuhurra and Alinda, 2017). In Tanzania, just three topics constitute 80% of the English content coverage in the primary years, indicating a very narrow focus over seven years of schooling. Finally, the "drop and reinstate" pattern observed in both countries, in which a topic is covered in one year, dropped in the subsequent year or two, then reinstated, indicates non-systematic content coverage, which may make it difficult for children to build on previous knowledge.

Second, while teachers are delegated the task of teaching the content stipulated in the curriculum standards, they face many, often competing demands, including reconciling the demands of the standards with the reality facing them in the classroom. From handling large, overcrowded classes to teaching children of varying learning levels and backgrounds, the requirement to cover the prescribed content can prove difficult. The content and pacing of curriculum standards can also be overambitious for student learning levels and paces of learning (Pritchett & Beatty, 2012). The success of programs that adapt instruction to children's learning levels, such as "teaching at the right level" type approaches suggest that typical instruction is often poorly aligned for the average student to learn. Furthermore, teachers may not have the training or support needed to implement the standards as intended, and may often be under pressure to prepare children to do well on exams which may or may not be well aligned with the standards.

We indeed find low alignment between curriculum standards and teachers' instructional content in both countries (see middle panel of Figure 15). In only one of the four subject/country combinations that we analyze does the alignment of teacher's instruction with curriculum standards surpass the 0.50 threshold for being considered reasonably well aligned (Uganda mathematics – coarse grain alignment is the exception). In both countries, teachers often cover broad swathes of content which is unrelated to the structure in the curriculum standards. This could reflect a lack of a planned overall content coverage structure, or could indicate rebellion from inappropriate or poorly designed and sequenced curriculum standards. Further research with teachers could help uncover the reasons for the observed content coverage patterns, which in turn could inform reforms to the curriculum standards or to the support teachers receive to cover the curriculum. The observed breadth of teachers' content coverage likely reduces the depth of knowledge children have the opportunity to gain.

Figure 15: Coherence between curriculum standards, national exams, and teacher instruction in Tanzania and Uganda: Coarse grain (topic level)



Note: 0.50 on the alignment scale is considered reasonably well-aligned. For standards vs. exams the standards are aggregated across all grades and exams are aggregated across the three years analyzed. For standards vs. teacher instruction, the standards and instruction are aggregated across the grades for which there is teacher instructional data. For exams vs. teacher instruction the exams are aggregated across the three years analyzed and teacher instruction is aggregated across the grades for which there is teacher instructional data. In Uganda the teacher instruction sample did not allow analysis of alignment with exams.

Third, although ostensibly intended as summative measures of achieved learning, the end-of-primary examinations in many countries serve many purposes including accountability for schools/teachers, and certification, selection, and sorting for students (Burdett, 2016). These additional roles, if not held in check, may assume more importance than measuring learning as prescribed in the standards. We find the primary leaving exams in both Uganda and Tanzania indeed are poorly aligned with the curriculum standards (left panel of Figure 15). Their typically narrower focus on fewer topics, and emphasis on lower-order levels of cognitive demand suggest they are not primarily aimed at assessing children's mastery of the curriculum standards. In the extreme case, Tanzania's English exam covers only three of the fourteen topics in the curriculum standards, and only assesses them at the recall/memorize cognitive level.

Finally, putting together the findings on teachers' instructional content and exams in a standards-based education system, low alignment of teachers' instruction and national examinations to curriculum standards leaves open the possibility that teachers have narrowly aligned their instruction to the exam content. Perhaps they feel pressure to prepare children for the exam and

prioritize this over the content in the standards. We find, however, that teacher's instruction is also poorly aligned with examination content for the grades/subjects where this alignment analysis is possible (right panel of Figures 18 and 19). For Tanzanian English instruction the alignment with the national exam is only 0.11. This suggests that neither are teachers targeting their teaching to the expected content on the examinations nor are the examinations driven by the content that is taught in the classrooms.

Building on these findings, the current study suggests multiple possible extensions of this work and research.

First, the systematic analysis of curriculum standards can inform a review and restructuring of curriculum standards to better align with children's levels and pace of learning. The structured analytical approach that SEC provides can help policy makers map the path they expect children to take through content, from students' learning levels when they enter school to the aspirations that education systems have for their students, ensuring children have the opportunity to learn all necessary and appropriate content along the way. The same approach can then be used to ensure support provided to teachers, such as textbooks, instructional materials, and training, are coherent with the standards.

Second, without a systematic approach for analyzing alignment, teachers are often in the dark about how well their instruction truly aligns with prescribed standards. Results like those in this study have been used in other contexts to provide structured insight to individual teachers on their instructional content and its alignment with curriculum standards. By training teachers on how to interpret the content maps, SEC has been used to support teachers' professional development, providing a structured, systematic way for teachers to better understand their own instructional practices, and enabling them to reflect on and deliberately adapt their instructional content and bring it into alignment with the standards (Blank, 2004; Smithson & Blank, 2006). At a school level, the methodology has also been used to enable peer learning, in which a group of teachers work together to compare content maps, identify areas for improvement, and learn from each other (CCSSO, 2000). Future research could examine the effectiveness of such approaches in developing country contexts.

Third, larger scale surveys of teachers' instructional content, beyond the piloting of these types of surveys in the current study, would provide valuable information to policy makers and curriculum developers. By comparing teacher instructional content maps to those of the curriculum standards, these actors could identify areas in which the standards and teacher instruction are consistently misaligned. Such misalignment could indicate that standards are overambitious or otherwise inappropriate for the students or environment in the classroom or infeasible for teachers to cover in the time allotted, that teachers have not been adequately trained on the standards they are meant to cover, that teachers have not been supported in implementing the standards in a consistent and meaningful way, or any number of other scenarios preventing teachers from following prescribed content. Identifying the areas of misalignment (in terms of topics, cognitive demand, and progression of content) is a first step for bringing instruction and standards into alignment in a way that promotes learning.

Fourth, results such as those in this study can inform efforts by the examination and curriculum bodies in the two countries to establish examination content validity. The common language, provided in the SEC approach, for coding content can facilitate discussion across these two bodies, and the content maps enable pinpointing of exam content that could be brought into better alignment with standards. In both Uganda and Tanzania, for example, the analysis reveals a need to expand the topics and levels of cognitive demand covered by primary leaving exams to better reflect what is prescribed in the standards and ensure children are being tested on the content the standards prescribe that they learn.

A fifth possible extension for future work could assess children against the same taxonomy used to code curriculum standards, examinations, and teacher instruction. This type of analysis has been conducted with useful results in the United States (Smithson, 2017). Such analysis would show which parts of the curriculum standards are well aligned with children's learning levels, and which may be over- (or under-) ambitious (Pritchett & Beatty, 2012). Comparing student learning levels to teachers' instructional content would similarly show how aligned instruction is to children's learning levels and could also identify patterns in teacher instruction, for example if instruction tends to favor some students (such as high performers) over others.

Finally, adopting a common SEC analysis framework across both Uganda and Tanzania (and potentially other countries) could further contribute to improved learning. For example, we found significant location-based instructional disparities in Uganda and none in Tanzania, presenting opportunities for cross-country learning. Another critical focus area that might greatly benefit from this cross-country learning relates to the country-level differences in focus on cognitive skills' development.

An important contribution of this study is to demonstrate the use of the SEC methodology for analysis of instructional components and diagnosis of their alignment with each other. In addition to providing empirical results for Uganda and Tanzania, it serves as a proof of concept for an approach that can be used in other contexts to diagnose areas of instructional incoherence in education systems. With its structured, systematic approach to quantifying content across multiple instructional components, and its adaptability to other uses such as teacher professional development, this type of analysis can provide valuable insight to education systems and guide reform efforts towards achieving coherence for learning.

References

Allen, R., Elks, P., Outhred, R., & Varly, P. (2016). *Uganda's assessment system: A road-map for enhancing assessment in education* . DFID Health & Education Advice & Resource Team (HEART). Retrieved June 28, 2020, from https://assets.publishing.service.gov.uk/media/585a8c7740f0b60e4a0000dc/Final-report-Enhancing-Ugandas-Assessment-System-September-2016.pdf

Altinyelken, H. K. (2010). Curriculum change in Uganda: Teacher perspectives on the new thematic curriculum. *International Journal of Educational Development, 30*, 151-161.

Anderson, L. W. (2002). Curricular alignment: A re-examination. *Theory into practice, 41*(4), 255-260.

Asim, S., Chugunov, D., & Gera, R. (2019). *Student Learning Outcomes in Tanzania's Primary Schools : Implications for Secondary School Readiness (English).* Washington, D.C.: World Bank Group. Retrieved June 26, 2020, from http://documents1.worldbank.org/curated/en/637761553580119324/pdf/Student-Learning-Outcomes-in-Tanzania-s-Primary-Schools-Implications-for-Secondary-School-Readiness.pdf

Atuhurra, J. (2016). Does community involvement affect teacher effort? Assessing learning impacts of Free Primary Education in Kenya. *International Journal of Educational Development, 49*, 234–246.

Atuhurra, J., & Alinda, V. (2017). *Basic education curriculum effectiveness analysis in East Africa: Using the 'Surveys of Enacted Curriculum' framework to describe primary mathematics and English content in Uganda.* Munich Personal RePEc Archive - MPRA. Retrieved June 26, 2020, from https://mpra.ub.uni-muenchen.de/79017/

Atuhurra, J., & Alinda, V. (2018). *Basic education curriculum effectiveness in East Africa: A descriptive analysis of primary mathematics in Uganda using the 'Surveys of Enacted Curriculum.* Oxford: RISE programme. Retrieved June 26, 2020, from https://www.riseprogramme.org/sites/www.riseprogramme.org/files/inline-files/Atuhurra_1.pdf

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., . . . Walton, M. (2016). Mainstreaming an Effective Intervention: Evidence from Randomized Evaluations of 'Teaching at the Right Level' in India. *MIT Department of Economics Working Paper, No. 16-08*. doi:http://dx.doi.org/10.2139/ssrn.2846971

Bashir, S., Lockheed, M., Ninan, E., & Tan, J.-P. (2018). *Facing Forward. Schooling for Learning in Africa.* Washington DC: The World Bank.

Blank, K. R., Porter, A., & Smithson, J. (2001). *New tools for analyzing teaching, curriculum and standards in mathematics and science. Results from Survey of Enacted Curriculum project. Final report.* Washington DC: Council of chief state school officers. Retrieved June 26, 2020, from https://files.eric.ed.gov/fulltext/ED458275.pdf

Burdett, N. (2016). *The good, the bad and the ugly - testing as a part of the education ecosystem.* RISE.

Burdett, N. (2017). *Review of high stakes examination instruments in primary and secondary school in developing countries.* RISE.

Byamugisha, A., & Ssenabulya, F. (2005). *The SACMEQ II Project in Uganda: A Study of the Conditions of Schooling and the Quality of Education.* Harare: SACMEQ.

Case, B. J., Jorgensen, M. A., & Zucker, S. (2004). *Assessment report: Alignment in educational assessment* . London: Pearson Education Inc.

CCSSO, C. o. (2000). *Using Data on Enacted Curriculum in mathematics and science.* Washington D.C: Council of Chief State School Officers. Retrieved August 7, 2020, from https://curriculumanalysis.org/Reference/DECStudy2000.pdf

City, E., Elmore, R., Fiarman, S., Teitel, L., & Lachman, A. (2009). *Instructional Rounds in Education : A Network Approach to Improving Teaching and Learning.*

Conn, K. (2014). *Identifying Effective Education Interventions in Sub-Saharan Africa: A meta-analysis of rigorous impact evaluations.* New York: Columbia University.

Crouch, L., & DeStefano, J. (2017). *Doing reform differently: Combining rigor and practicality in implementation and evaluation of system reforms.* North Carolina: RTI International.

Ezati, A. B. (2016). *The recent history and political economy of basic education curriculum reforms in Uganda.* Kampala: Not published.

Freudenberger, E., & Davis, J. (2017). *Tusome external evaluation—Midline report.* Prepared for the Ministry of Education of Kenya under contract no. AID-615-TO-16–00012. Washington, DC: USAID/Kenya and the UK DFID. Retrieved June 26, 2020, from https://shared.rti.org/content/tusome-external-evaluation-midline-report

Gamoran, A., Porter, C. A., Smithson, J., & Paula, W. A. (1997). Upgrading high school mathematics instruction: Improving learning opportunities for low-achieving, low-income youth. *Educational Evaluation and Policy Analysis, 19*(4), 325-338.

GoT. (1973). *The national examinations council of Tanzania act, 1973.* Dar es salaam: Government of Tanzania.

GoT. (1975). *The institute of education act, 1975.* Dar es salaam: Government of Tanzania.

GoT. (2019). *the national examinations council of Tanzania act, 2019.* Dodoma: Government of Tanzania.

GoU. (1973). *The National Curriculum Development Centre Act 1973.* Kampala: Government of Uganda.

GoU. (1983). *Uganda National Examinations Board Act 1983.* Kampala: Government of Uganda.

GoU. (1992). *Government white paper on the education policy review commission report entitled 'education for national integration and development'.* Kampala: Ministry of Education and Sports (MoES).

GPE. (2015). *Results for learning report 2014/15 .* Washington DC: Global Partnership for Education. Retrieved March 23, 2017, from http://www.globalpartnership.org/content/results-learning-report-2014-15

Jones, S. (2015). How does classroom composition affect learning outcomes in Ugandan primary schools? . *International Journal of Educational Development, 48*, 66–78.

Kitta, S., & Afeli, S. (2017). *Examinations and competitive exams in Tanzania.* Dakar: education-2030africa.org. Retrieved from http://education2030-africa.org/images/talent/atelier061217/PPT_Tanzania.pdf

Krathwohl, D. R. (2002). A revision of Bloom's Taxonomy: An overview. *Theory into Practice, 41*(4), 212-218.

Kurz, A., Elliott, N. S., Wehby, H. J., & Smithson, L. J. (2010). Alignment of the intended, planned, and enacted curriculum in General and Special Education and its relation to student achievement. *The Journal of Special Education, 44*(3), 131-145. doi:https://doi.org/10.1177/0022466909341196

Luigi Giussani. (2017). *National symposium on Uganda's examinations and assessment system. Present and future prospects.* Kampala: Luigi Giussani Institute of higher education.

Matete, R. (2016). Challenges Facing Primary Education under Decentralization of Primary School Management in Tanzania. *International Journal of Humanities and Social Science, 6*(1), 175-184. Retrieved June 26, 2020, from http://www.ijhssnet.com/journals/Vol_6_No_1_January_2016/22.pdf

Mbiti, I., & Rodriguez-Segura, D. (Forthcoming). *Evaluating Curriculum Reforms in Developing Countries: Evidence from Tanzania.*

MoEST. (2018). *Education sector development plan (2016/17 - 2020/21).* Government of tanzania.

MoEST. (2019). *National Curriculum framework for Basic and Teacher Education.* Dar es Salaam: Tanzania Institute of Education.

MoESTS. (2014). *Education statistical abstract: Findings from the annual school census, 2014.* Kampala: Education planning and policy analysis department, MoESTS.

Munene, C. J. (2017). *Refocusing and revitalizing Uganda's national assessment system. Peeping into emerging patterns of a sprouting institution.* Kampala: Keynote presentation at the National Symposium on Evolution of Uganda's assessment and examinations system.

NCDC. (2006a). *The national primary school curriculum for Uganda. Primary 1.* Kampala: National Curriculum Development Centre.

NCDC. (2006b). *The national primary school curriculum for Uganda. Teacher's guide Primary one.* Kampala: National Curriculum Development Centre.

NECTA. (n.d). *Exams Design and Development.* NECTA. Retrieved June 28, 2020, from https://www.necta.go.tz/exams_design_development

Nicholas, M. C., & Raider-Roth, M. (2016). A hopeful pedagogy to critical thinking. *International Journal for the Scholarship of Teaching and Learning, 10*(2), 1-10.

Nzima, I., & Mkumbo, K. (2017). *Review of Curriculum reform processes and outcomes in Tanzania. Implications for learning outcomes.* Dar es salaam: Prepared for Twaweza East Africa - Not Published.

Piper, B., DeStefano, J., Kinyanjui, E. M., & Ong'ele, S. (2018). Scaling up successfully: Lessons from Kenya's Tusome national literacy program. *Journal of Educational Change, 19*, 293-321.

Polikoff, S. M., Porter, C. A., & Smithson, J. (2011). How well aligned are state assessments of student achievement with state content standards? *American Educational Research Journal, 48*(4), 965-995.

Porter, C. A. (2002). *Measuring the content of instruction: Uses in research and practice. 2002 presidential address.* Madison: University of Wisconsin.

Pritchett, L. (2015). *creating education systems coherent for learningoutcomes: making the transition from schooling to learning.* Research on Improving Systems of Education (RISE).

Pritchett, L., & Beatty, A. (2012). *The negative consequences of overambitiuos curricula in developing countries.* Washington, DC: Centre for Global Development.

Rowan, B. (1998). The task characteristics of teaching: Implications for the organizational design of schools. In R. Bernhardt, C. N. Hedley, G. Cattaro, & V. Svolopoulos, *Curriculum leadership: Re-thinking schools for the 21st century.* Hampton Press.

SACMEQ. (2020). *SACMEQ Indicators.* Retrieved from http://www.sacmeq.org/ReadingMathScores

Smithson, J. (2013). *Measuring opportunity to learning (OTL): Benefits for research and practices .* Madison: Wisconsin University.

Smithson, J. (2015). *North Carolina 2015 alignment study report.* Madison: Wisconsin.

Smithson, J. (2017). The science of standards-based education. *The journal of the national association of state boards of education, 17*(2), 15-21.

Smithson, J., & Blank, R. (2006). *Longitudinal study of the effects of MSP-supported professional development on improving mathematics and science instruction.* Washington D.C.: Council of Chief State School Officers (CCSSO). Retrieved August 7, 2020, from https://curriculumanalysis.org/Reference/MSP-PDStudyYr2Report05.pdf

Smithson, J., & Collares, C. (2007). *Alignment as a predictor of student achievement gains.* Chicago: 2007 annual meeting of the American Education Research Association (AERA) .

Solomon, S., & Zeitlin, A. (2019). *What do Tanzanian parents want from primary schools - and what can be done about it?* RISE.

Twaweza. (2015). *What makes a curriculum effective in promoting children's learning? A concept note for analyzing basic education curriculum effectiveness in East Africa.* Dar es salaam: Twaweza East Africa. Retrieved June 26, 2020, from https://www.twaweza.org/uploads/files/ConceptNoteDraft08102015mc+rc.pdf

UNICEF. (2017). *Education Factsheet: Tanzania.* Dar es salaam: UNICEF. Retrieved June 26, 2020, from https://www.unicef.org/tanzania/reports/education-fact-sheet

USAID. (2015). *Testing and Reading, Writing, and Arithmetic (3Rs) Reform in Tanzania.* Presentation prepared for 2015 Global Education Summit.

USAID/RTI. (2014). *The status of early grade reading and teaching reading in primary school: Cluster 2 baseline report.* Kampala: USAID.

Uwezo. (2016). *Are our children learning? Uwezo Uganda 6th learning assessment report.* Kampala: Twaweza East Africa.

Vuzo, M. (2018). Towards achieving the Sustainable Development Goals: Revisiting language of instruction in Tanzania secondary schools. *International Review of Education, 64*, 803-822.

Ward, M., Penny, A., & Read, T. (2006). *Education reform in Uganda 1997 – 2004: reflections on policy, partnership, strategy and implementation.* London: Department for International Development (DfID).

World Bank. (2012). *Student assessment, Uganda – SABER country report 80071 - Uganda.* Washington D.C: The World Bank.

Appendix A: Achieved samples of primary Mathematics and English teachers

Table A.1: Achieved samples of primary Mathematics and English teachers surveyed in Tanzania

Grade	Primary Mathematics			Primary English		
	All	*Mkuranga*	*Ilala*	*All*	*Mkuranga*	*Ilala*
Standard 3	20	2	18	15	3	12
Standard 4	11	3	8	21	7	14
Standard 5	15	8	7	10	2	8
Standard 6	14	6	8	16	4	12
Standard 7	14	9	5	15	8	7
Total	**74**	**28**	**46**	**77**	**24**	**53**

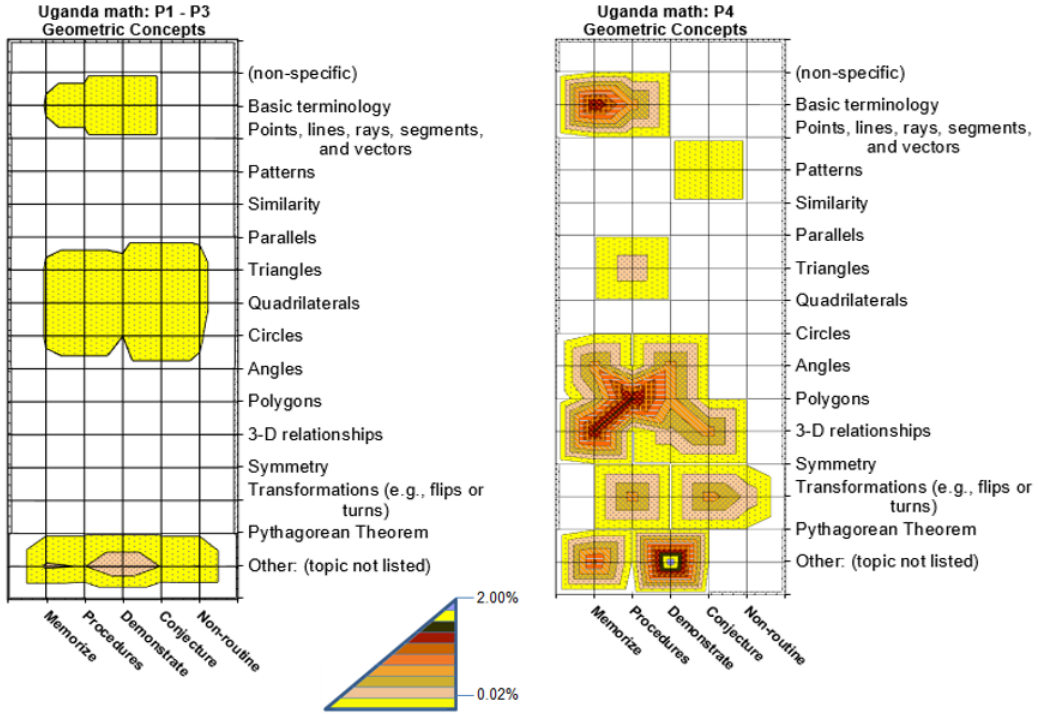
Table A.2: Achieved samples of primary Mathematics and English teachers surveyed in Uganda

Grade	Primary Mathematics			Primary English		
	All	*Iganga*	*Wakiso*	*All*	*Iganga*	*Wakiso*
Primary 3	73	46	27	99	44	55
Primary 5	101	43	58	81	42	39
Total	**174**	**89**	**85**	**180**	**86**	**94**

Appendix B. Illustrative fine grain content map

Each topic contains a set of subtopics, based on the taxonomy developed for each country. Figure B.1. provides an example of a fine grain content map. It is interpreted the same way as the coarse grain maps throughout the paper. Each topic in a coarse grain map represents an aggregation of fine grain maps like the one in B.1.

Figure B.1: Ugandan fine grain (subtopic) curriculum standards for geometric concepts in P1-P3 and P4.



Note: The geometric concepts topic included 14 subtopics. While some overlap of subtopics is observed between the early primary years (P1-P3) and P4, there are substantial differences in subtopic coverage.