

Curriculum Alignment and Pedagogical Suitability as Predictors of Teachers' Perceived Quality of STEM Learning Materials in Tanzanian Primary Schools

Dr. Angela Rutakomozibwa Katabaro

Principal Curriculum Developer, Tanzania Institute of Education, P.O. Box 35094, Dar Es Salaam, Tanzania

ORCID: <https://orcid.org/0009-0004-8861-3852>.

*Corresponding Author

Dr. Angela Rutakomozibwa Katabaro

E-mail:
angelarutakomozibwa3@gmail.com

Article History

Received: 22.04.2026

Revised: 29.05.2026

Accepted: 17.07.2026

Published: 27.08.2026

Abstract: This study examined whether curriculum alignment and pedagogical suitability were associated with teachers' perceived quality of STEM learning materials used in Standards 3–6 in Tanzanian primary schools. A cross-sectional survey of 1,087 respondents from selected regions was analysed. Measures assessed curriculum alignment, pedagogical suitability, and perceived material quality. Both curriculum alignment and pedagogical suitability were positively associated with perceived overall quality and independently predicted it in regression analyses. Together, the two predictors explained 41.9% of the variance in perceived quality; both remained significant after adjustment for location, subject, role, experience, qualification, and region. Open-ended responses most frequently cited shortages of teaching and learning materials and books (54.4%) and the need for increased supply (56.9%). These findings indicate that teachers' judgements of material quality are closely related to perceived curriculum alignment and pedagogical suitability, while also highlighting the need to improve the supply of STEM learning materials in primary schools.

Keywords: *STEM education; curriculum alignment; pedagogical suitability; learning materials; primary education; Tanzania; competency-based curriculum.*

This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)

Cite this article:

Katabaro, A. R., (2026). Curriculum Alignment and Pedagogical Suitability as Predictors of Teachers' Perceived Quality of STEM Learning Materials in Tanzanian Primary Schools. *ISAR Journal of Arts, Humanities and Social Sciences*, 4(8), 53-69.

1. Introduction

Primary education has become a central focus of Tanzania's national agenda for human capital and technological readiness, as well as equitable development. The Education and Training Policy (2014, 2023 edition) clearly has education as a system which 'shall produce Tanzanians who have knowledge, skills, positive attitudes, competence and self-reliance'. It also acknowledges the need for education systems to anticipate science and technology developments and adapt education systems, and improve curricula to align with the country's development requirements (Ministry of Education, Science and Technology [MoEST], 2023a). This will give a policy context to review whether classroom learning materials are consistent with the intentions of the curriculum and are suitable for the teaching and learning of STEM in Standards 3-6. In the international context, STEM education has come to be seen as a solution to the changing nature of the labour market, technological shifts and the need for flexible problem-solving abilities (Bybee, 2013; English, 2016). But there would be little point in policy statements if they were not implemented in classrooms, and that is where it is thought that school-level leadership and material conditions make a difference, rather than policy statements, according to UNESCO's Global Education Monitoring Report 2024/5 (UNESCO, 2024). This translation challenge is faced by Sub-Saharan African education systems especially; exemplary education reforms aimed at attaining competence are often implemented at the policy level yet at the practice level, there are challenges such as the availability of learning materials and their design (Kasuga and Kalolo, 2025; Wawire, Kiambati, Muhia, Gathoni and Abuya, 2025).

As seen in the rest of this region, Tanzania has seen a gradual shift in the policy emphasis of its education and training system from the more abstract notions of knowledge, skills, and attitudes to the practical aspects of competence. Successive revisions of education and training policies have increasingly highlighted the importance of competence, science, technology, and practical skills, but the material conditions needed to implement the practical aspects of these policy visions have been less systematically explored within Tanzania's education and training system than curricula have been. The Education and Training Policy of Tanzania has undergone a series of revisions, starting with the general policy of 2014 and followed by the new policy in 2023, which was used throughout this paper, as a continuous process by the country to make its basic education more competence-based and skills-oriented. Every revision has had implications for what it is expected of STEM teaching and learning materials in the classroom, gradually shifting from content-transmission models to models that anticipate that materials will facilitate active learning, problem solving and locally based application of

knowledge. The present study is interested in the meaningful outcome measures of teacher-rated curriculum alignment and pedagogical suitability because they operationalise the policy expectation from a classroom perspective to see if it has made it to the day-to-day use of the materials by teachers. This paper's emphasis on curriculum alignment and pedagogical appropriateness addresses this gap directly. In the paper, teaching and learning materials are not seen as a means through which learners are given information to support the teaching and learning of STEM or the curriculum, but instead materials are seen as an object to be analysed in its own right: an artefact that can facilitate or impede the conversion of the intention of the curriculum to classroom practice. Such a framing is important for curriculum designers, publishers and government procurement procedures, which are making their decisions in curriculum design and distribution that can be informed by the findings of the present study. It also has conceptual value: Conceptualising alignment and suitability as empirical properties that can be assessed by the teacher using a student rating allows the relationship between policy intent and classroom reality to be studied empirically. The context of national statistics does not detract from the significance of this question. BEST 2024 records there to be 11,391,185 pupils (in Standards I-VII) and 229,840 primary school teachers, giving an overall pupil-teacher ratio of 1:50 and a pupil-qualified teacher ratio of 1:51 (MoEST, 2024). BEST also considers the pupil/student book ratio to be an important system indicator as it demonstrates that pupil/student learning materials are a measurable aspect of education quality. The primary education section of the 2025 regional BEST volume provides subject-specific textbook information for grades III to VII, which is directly relevant to STEM education in Standard III to VI (PO-RALG, 2025). Important documents for interpreting the data include policy and curriculum documents written in Swahili. In line with Tanzania's national education goals outlined in the Primary Teacher Diploma Curriculum, one goal is "to develop knowledge and apply science and technology, creativity, critical thinking, innovation, collaboration, and communication" (MoEST, 2023b). It is the same curriculum that provides for the explanation of knowledge, skills, dispositions, competencies, teaching and learning methods, teaching and learning materials, and approaches to assessment of competence. The policy statements serve to explain why it is appropriate to consider STEM materials through a lens of alignment, suitability, usability, and quality, instead of as a given object of the classroom.

In the world, curriculum alignment is related to coherent learning systems that have learning goals, learning activities, learning materials and assessments that are coherent towards each other. Biggs (1996) defines the concept of constructive alignment as a relationship between intended learning outcomes, teaching/learning activities, and assessment. In

addition, Schmidt, Wang, and McKnight (2005) note that "coherence is important in mathematics and science because standards alone are not enough; coherence is necessary if they are to be implemented in the classroom with sequence, focus, and instructional support. In such situations, the materials that are needed for implementation of a competence-based pedagogy are crucial, as teachers' abilities to implement the pedagogy might be constrained by the lack of resources, examples, tasks, and culturally relevant curriculum supports.

2. Statement of the problem

Tanzania's policy orientation is based on competence, science and technology, inclusion, quality, and relevance, but the implementation of the policy in the classroom is dependent on the materials available and the quality of those materials. If materials do not fit well with the curriculum goals or are pedagogically inappropriate for the learners, teachers have to devote more time to adapting the curriculum, developing examples, and addressing the mismatch between curriculum expectations and classroom practice. Previous studies in Tanzania on competence-based curriculum implementation have tended to concentrate on teachers' understanding of the competence, their challenges, quality assurance of the implementation, and teachers' professional preparation (Dismas et al., 2023; Nzima, 2016, 2018). There has been less focus on the extent to which Standards 3-6 STEM teachers assess the actual alignment of the materials they use to the standards, and the extent to which they consider the pedagogical appropriateness of the materials. The present study fills that void. The present paper is one of three in a series that examines the same STEM dataset (N = 1,087) from the same context of complementary angles while analysing different aspects of the data: this paper compares material alignment and pedagogical suitability to the teacher's sense of quality; a companion paper compares teacher and school characteristics to the teacher's sense of quality; and a third companion paper compares teacher agency and adaptive practice under the constraints of the available resources. This relationship is made public here and will be made public at the time of submission to the target journals, to help prevent duplicate publication/redundant analysis of one data set.

3. Objectives and hypotheses

- a) To assess teachers' perceptions of the curriculum alignment of STEM learning materials used in Standards 3-6.
- b) To examine teachers' perceptions of the pedagogical suitability of STEM learning materials for primary school learners.

- c) To determine whether curriculum alignment and pedagogical suitability predict perceived overall material quality.

- d) To compare perceptions by school location, region, subject taught, professional role, qualification, and teaching experience.

- e) To identify teacher-reported challenges and improvement priorities related to STEM learning materials.

Hypotheses

H1: Curriculum alignment is positively associated with perceived overall quality of STEM learning materials.

H2: Pedagogical suitability is positively associated with perceived overall quality.

H3: Curriculum alignment and pedagogical suitability significantly predict perceived overall material quality.

H4: Perceptions differ by selected teacher and school characteristics.

H5: Teachers' perceptions of the STEM learning materials (e.g., curriculum alignment, pedagogical suitability

4. Theoretical framework

The four theoretical ideas are complementary and provide the foundation for the study. The first is constructive alignment theory (Biggs, 1996), which suggests that the learning outcomes, learning activities and assessments should align with each other. The concept is extended in this paper to the STEM learning materials: materials are aligned when they link their content, their examples, their problems and their sequences to curriculum expectations and enactment in the classroom.

Second, based on the curriculum coherence theory, mathematics and science curricula should be sequenced and focused and not overcrowded or fragmented (Schmidt et al., 2005). The study thus sees curriculum alignment as not just correspondence between curriculum materials and the curriculum statements but also as coherence and instruction continuity.

Third, the lens of implementing the curriculum based on competence is used as a policy implementation lens. Materials should facilitate active learning, problem solving, application of knowledge, and engagement with practices in competence-based approaches. It aligns with the policy focus on skills, science and technology, creativity, critical thinking, innovation, and practical learning that is given priority in Tanzania (MoEST, 2023a, 2023b).

Fourth, if the instructional design theory is used for analysis, a pedagogical suitability can be determined. Materials are pedagogically appropriate when they meet the learning level(s), promote active participation, are well organized, and are appropriate within the classroom setting. This

establishes the conceptual framework used in the paper: That alignment of the curriculum and pedagogical suitability is likely to account for overall material quality.

All the four strands of theories are not explained as alternative explanations, but as complementary lenses that together define what a well-designed STEM learning material is. Let's take a concrete example to illustrate the point. To consider a unit of a Standards 4 mathematics workbook on fractions. Constructive alignment means that the worked examples and practice tasks in the unit must relate to the learning outcomes for the particular content area of fractions that are listed in the syllabus, not to any related but untaught content. Curriculum coherence is the need to build logical connections from previous units to the current unit, in this case, a knowledge of part-whole reasoning, a knowledge developed in previous standards, should be used to build up to operations with fractions, instead of teaching operations in isolation. Competence-based implementation entails that tasks in the Unit include the use of fraction reasoning in solving realistic problems, such as dividing a plot of land, or sharing a harvest. The examples should be clear, be appropriate for Tanzania primary learners and be embedded in a context that is familiar to them to ensure that cognitive effort is invested in understanding the concept of mathematics and not the unfamiliar context. One material that meets only some of these criteria might be usable, but, according to the theory, teachers should perceive such material as partly aligned or partly suitable; the survey tool used in this study was designed to capture such a "graded, continuous" perception pattern.

This theoretical framework also needs to be stated explicitly as to what it is not. It does not claim that the quality of STEM content as perceived by teachers is solely dependent on the curriculum alignment and pedagogical appropriateness of the materials, since the qualitative result presented in Section 7.6 clearly indicates that the quantity, availability and durability of the materials are also important factors for the teachers' concerns. These factors are not included within the theoretical scope of the concept of constructive alignment and pedagogical design theory that is addressed here. The framework should not be interpreted as a comprehensive set of criteria for the success of STEM teaching and learning in a school, but rather as a set of conditions that must exist when a material is physically present in a classroom, in order for it to be effective for teaching and learning.

5. Literature Review

Studies of teachers' use of curriculum resources suggest that teachers do not teach what they find in the materials directly, but that they interpret, adapt, select and supplement the materials. Remillard (2005) contends that teachers' interactions with the curriculum materials are influenced by

both the curriculum materials and teachers' knowledge, contexts, and teaching goals. For STEM teaching, that translates to the need to interpret even well-designed materials in the classroom and a potential burden of implementation with poorly aligned materials.

Coherence of the curriculum, pedagogical representation and usability are correlated with the quality of the STEM materials. Shulman (1986) identified three types of knowledge: content knowledge, pedagogical content knowledge, and curricular knowledge, recognizing that, beyond content knowledge, teachers must have subject matter representations, examples, explanations, and curricular awareness to make the subject matter teachable. In the case of mathematics, Ball, Thames, and Phelps (2008) argued that knowledge of content for teaching involves knowledge in special forms that are utilized in actual classroom practice. The arguments presented support the focus of the present study on materials as carriers of content, pedagogy and curriculum coherence.

Research on implementation of competence-based curriculum in Tanzania indicates that implementation is influenced by the teacher understanding, competence and material conditions. Nzima (2016, 2018) records teacher educators' interpretations of the competence-based curriculum and suggests implications for teacher education practice. Dismas, Nzima and Kimaro (2023) assert that the understanding and support practices of the quality assurance officers impact on the implementation of the competence based curriculum. Science and mathematics educators' technology integration competence can be reinforced using collaborative design (Kafyulilo & Tilya, 2019). In addition, Kabote, Niboye and Nombo (2014) suggest that gender and rural-urban contexts affect the performance of students in Tanzanian primary schools in mathematics and science. These studies collectively indicate a need to analyze the quality of the material, teacher preparation, and school context as a single unit. More recent evidence confirms this picture: Kasuga and Kalolo (2025) note that access to resources, teacher readiness and assessment practice are still entangled challenges in the implementation of competence-based curriculum in a region of Tanzania, and a review of evidence from various competence-based curriculum systems in Africa, including Tanzania, by Wawire et al. (2025), indicates that the challenges of pedagogical and assessment practice are continuing to be constrained by issues related to systemic factors such as material and infrastructure provision, not only by curriculum design. The recent findings place the focus of the present study in the local and current research agenda of material alignment and suitability. The same can be said about the pedagogical-suitability domain, as even well-prepared teachers will revert to activity-based, teacher-centred approaches to STEM content when the curriculum materials do not clearly

indicate the technology and engineering aspects, as seen in the present study (see also Mafugu, Tsakeni and Jita, 2023).

5.1 Curriculum coherence and constructive alignment in comparative perspective

Constructive alignment is a framework proposed by Biggs (1996) in the field of higher education quality assurance and is now widely used in basic education systems in the process of reforms based on outcomes or competence. The main tenet is that learning is best, and most equitably evaluated, when the learning goal, the learning activities in which content is presented, and the assessment tasks in which the learning is evaluated are all focused on the same goal. Both over specification and under specification are possible, such as tasks being set that are cognitively more demanding than the learning outcomes or when the cognitive demand of the learning outcome is not met. In STEM subjects, for example, where outcomes often include conceptual understanding and procedural fluency, and applied problem-solving, this three-way mismatch is particularly significant, especially for STEM subjects.

The Third International Mathematics and Science Study (TIMSS) analysis of mathematics and science curricula by Schmidt et al. (2005) sparked the popular criticism that many national curricula try to address too many topics at too coarse a level of detail or at too shallow a depth of learning. They assess curricula and, in turn, materials that implement the curricula – according to how well they show a coherence in the sequence of topics, the focus of topics, and the rigour of the topics. In the present study this means curriculum alignment is not just about matching content with content, but also a match of content with the content being developed with a logical sequence designed to lead up to the intended learning. Omission is a more obvious misalignment of content than this, and arguably more significant, since it means a student or students are missing a portion of the required subject matter.

This picture is further supported by the results of systematic reviews conducted around the world. However, in their review of teachers' perceptions of STEM integration, Margot and Kettler (2019) found that a common problem, even in under-resourced systems, was the challenge teachers faced in making use of existing materials to facilitate truly integrated cross-disciplinary STEM tasks instead of content-focused, isolated STEM tasks. This aligns with pedagogical suitability construct studied in the present paper, where it was noted that the challenges by teachers in Tanzania Standards 3-6 is not peculiar to the Tanzania context but is one of many documented issues in contexts across the globe in the area of between curriculum ambition and material design capacity.

5.2 Content knowledge, pedagogical content knowledge, and the design of STEM materials

The distinction between subject matter knowledge, pedagogical knowledge, and pedagogical content knowledge, which was developed by Shulman (1986), is a cornerstone of the understanding of why materials should be carefully designed, and why not just teachers. A narrow definition of pedagogical content knowledge is the knowledge of how to represent and formulate content in a way that makes it comprehensible to students the analogies, examples, demonstrations, and explanations that make up the content more teachable. The required pedagogical content knowledge is well integrated into the material, and the well-designed STEM material provides a representation, worked examples, and explanations that are ready to use by teachers. By contrast, a poorly designed material has the potential to shift the burden of this representational work from the class to the individual teacher, which is an expectation that is difficult to meet in a system that is largely comprised of Certificate-level qualified teachers, at least by the sample size of this study.

Ball, Thames, and Phelps (2008) expanded the framework of Shulman specifically for mathematics instruction, and categorised common content knowledge, specialised content knowledge, and knowledge of content and students. Their framework suggests that pedagogically appropriate STEM materials should not only convey accurate content, but they should also predict the misconceptions that learners are likely to have, and arrange the practice problems to reveal and correct those misconceptions. Materials that do not provide this scaffolding impose a greater load on the specialised knowledge of teachers, in line with this study's results where pedagogical suitability was most frequently reported as a concern for teachers.

Remillard's (2005) review of research on teachers' use of mathematics curricula suggested that the relationship between a teacher and a curriculum material is neither a one-way nor a passive one, but a participatory one: teachers read, interpret, evaluate, and adapt materials based on their own knowledge and their students' needs and problems, rather than following them as scripts. This participative perspective is compatible with, and supportive of, the ecological perspective described in the companion papers in this series that account for teacher agency and it also offers another theoretical justification as to why perceived pedagogical suitability, as opposed to raw content coverage, is likely to be a better predictor of how usable teachers see a given material to be.

Another area of lacuna in the current literature is the gender-specific analysis of perceptions about STEM material, rather than gender-specific analysis of performance outcome in STEM. The results reported by Kabote et al. (2014) on gender differences in math and science achievement in

Tanzanian primary schools are several causal steps removed from the immediate properties of materials that are the focus of the present study and are directly rated by teachers. If male and female teachers and equally, whether male or female teachers in predominantly male or predominantly female classes have differing experiences of curriculum alignment and pedagogical suitability is an empirically untested question that we examine directly in the present paper's accompanying study on teacher characteristics using the same data set, also including the gender variable, qualification, experience, subject and role.

5.3 Regional and Tanzanian evidence on materials, curriculum implementation, and STEM teaching

The Tanzanian findings are valuable in this respect, as research in this country is ongoing and has yielded valuable evidence to interpret the present ones. Nzima's (2016, 2018) research on teacher educators revealed that the teacher educators' understanding of competence-based curriculum requirements were not uniform but showed a lot of variation, some of them understanding the curriculum requirement of competence in a rather procedural way or through a checklist or a checklist approach to competence-based curriculum rather than a deeper shift in pedagogical understanding, and implications were pointed out to how teacher education itself practices competence-based curriculum requirements in terms of the use of curriculum materials. Dismas, Nzima, and Kimaro (2023) followed up this line of enquiry with district level quality assurance officers, who showed that their understanding of the requirements for competence based curriculum influenced the type of support and feedback that they were able to provide to schools, with officers with more procedural understandings giving correspondingly narrow supervisory feedback based on compliance rather than material quality or pedagogical fit.

Kafyulilo and Tilya (2019) showed in classroom context that the pedagogical suitability of the materials does not have to be pre-given but can be strengthened through structured, collaborative lesson design processes that also engage in-service science and mathematics teachers themselves, hence their confidence and competence in using technology in STEM education can be enhanced. The gendered and rural-urban differences in mathematics and science performance found in Tanzanian primary schools by Kabote, Niboye, and Nombo (2014), were observed in the current study as well, and highlight the need to look at rural-urban differences in addition to the gendered differences in the sampled areas.

Regional evidence is more recent; it adds to this picture. In the study of the implementation of competence-based curriculum by Kasuga and Kalolo (2025), using interviews and focus group discussions with the teachers, school administrators and students in Njombe Region, they

discovered that the issues of availability of resources, assessment practice, and teacher preparedness were all interdependent and not independent problems; teacher preparedness did not improve without the availability of resources, and assessment practice did not improve without preparedness. Consistent with the findings of Wawire et al. (2025) which analysed peer-reviewed evidence on the implementation of competence-based curriculum in Ghana, Kenya, Rwanda, South Africa, Tanzania, Uganda, and Zambia, this study concluded that the implementation of the curriculum was influenced more by systemic issues like the lack of teaching materials and infrastructure, and teacher training capacity, than by the curriculum documents themselves. Mafugu, Tsakeni, and Jita (2023) obtained similar results with preservice primary teachers in South Africa, observing even well-prepared preservice teachers resort to activity-based, single-subject approaches, when the STEM curriculum resources did not adequately scaffold the technology and engineering aspects of the STEM content which they underestimated. This collection of regional evidence places the current study's results (namely, that pedagogical suitability is rated somewhat less favourably than curriculum alignment) in a dynamic and geographically expansive trend, and not in isolation as a national result.

5.5 Methodological approaches in prior studies

The studies discussed above have used a variety of approaches to methodology. The methodologies used in the studies discussed above are quite variable and this is significant for the placement of the present study. Nzima's (2016, 2018) research with teacher educators involved qualitative interviews which were well suited to capturing depth of interpretation, but were not as well suited to establishing the extent to which specific interpretations were shared by the teacher-educator population. Dismas et al. (2023) also adopted a qualitative method with quality assurance officers. By using a design-based, interventionist methodology approach that would be suitable to demonstrating that a specific professional development approach can change practice, it is less suitable to establishing baseline rates of material quality perception across an unselected teacher population, as Kafyulilo and Tilya (2019) did. The designs of the large-scale quantitative surveys, as described in the companion papers to this study and in Kabote et al. (2014), offer greater statistical generalisability across regions and school types, but forgo some depth of interpretation.

The methodological contribution to the present study is the use of a large sample of a scale that is rarely used in Tanzanian STEM education research and a validated multi-item Likert instrument that can be used for reliability analysis, correlation analysis and multivariable regression. This enables the study to go beyond a descriptive reporting of perceived problems, which is common in much of the

literature, and test specific, falsifiable hypotheses about which properties of STEM materials predict overall perceived quality, statistically controlling for a broad spectrum of teacher and school characteristics. This does not imply that the design of the present study is better than that of the qualitative studies reviewed above; rather, the two types of studies complement each other, and the discussion in Section 8 is based on the quantitative pattern and the qualitative nuance from the broader literature.

5.4 Positioning of the present study

The curriculum-implementation challenges have thus far been reported in Tanzania from different perspectives: that of teacher educators, quality assurance officers, and technology-integration interventions. Less systematically explored is the perspective of teachers themselves (Standards 3-6) and their classroom teachers, who on a daily basis rate the alignment and pedagogical appropriateness of the specific materials they use, at a scale that is adequate to model quantitatively the relationship between these ratings and the overall judgement about material quality that they make. In the current study, this gap is directly addressed by creating a large sample of cross sections to examine the relationship between the two factors and perceived material quality when controlling for relevant teacher and school characteristics.

6. Methodology

The study was of a cross sectional survey design. The respondents included 1,087 teachers of STEM in selected regions of Tanzania from Standards 3 to 6. Classroom teachers and school leadership/management staff were respondents. Demographic and professional characteristics (Q1-Q8), perception items on a Likert scale (Q9-Q53), and open-ended items about challenges and suggestions for improvement (Q54-Q55) were the variables used for analysis.

Of the sample ($N = 1,087$), there was an uneven distribution across the six regions represented (Geita 47.7%, Mara 31.1%, Mbeya 10.7%, Njombe 9.4%, Simiyu 1.1%, Kilimanjaro 0.1% (Table 1)). The following items were used for measuring curriculum alignment (Q9-Q13), pedagogical suitability (Q14-Q18) and overall material quality (Q49-Q53). Each item was rated on a four-point scale with higher scores representing more positive perceptions. The mean values were calculated for each construct. Primary challenge and improvement themes were derived from open-ended responses through an initial procedure of applying a set of keywords. As in recommended practice with computer-assisted qualitative coding, this initial coding is also fully manual verified with the source responses before submission to the journal, which also applies to the Limitations section.

Data analysis. Descriptive statistics, Cronbach's alpha, Pearson correlations, Welch t-test, one-way ANOVA, and

ordinary least squares regression with robust standard errors were used. The primary regression model examined the relationship between the curriculum alignment and pedagogical suitability variables and material quality. School location, subject, role, teaching experience, qualification and region were adjusted for in an adjusted model. Qualitative responses were summarised by theme frequencies and illustrative examples selected.

The analytic sequence of descriptive statistics and reliability as the first step, followed by correlation, then regression, and finally group comparison was developed to present the empirical evidence in steps to support the central claim of the paper. The first phase of descriptive and reliability analysis is to determine that the constructs are measured with adequate precision for further analysis. A correlation analysis is then performed to establish the direction of the relationship between curriculum alignment and pedagogical suitability with overall material quality, in the sense that they are both associated with it in the hypothesized direction, which provides some support for H1 and H2 prior to the more complex regression analysis conducted later. The regression model, in turn, tests the paper's central hypothesis that the alignment and suitability jointly and independently predict quality, whereas the adjusted model that includes teacher and school characteristics tests the robustness of this finding against the alternative explanation that observed associations are confounded by the survey population of those who were observed. Group comparisons are reported last and are less central to the study: they indicate if the constructs are similar across subgroups and are important for the generalisability of the results of the regression to the diverse regions and school types included in the sample of schools.

Ethical note. An anonymised data set was analysed and findings are presented in anonymous or aggregates format. Recruitment process was carried out with ethical procedures. Written permission from the relevant regional authorities allowed permission to be given by the Head of the participating primary schools during the fieldwork.

Study design was cross sectional, as the primary purpose of the study was to characterize the prevalence of teacher perceptions in a large, multi-regional sample, and to predict relationships between the constructs measured, rather than to determine change over time or causal sequence. This design has good statistical power to detect the associations described in Section 7, but cannot infer causes for the perceived material quality; the associations reported are compatible with the cause being curriculum alignment or pedagogical suitability; however, longitudinal or experimental designs could be used in future to confirm the associations.

The reliability of Cronbach's alpha was used, and values of $> .70$ are considered acceptable for research purposes and $>$

.80 are considered good (Nunnally and Bernstein, 1994); all three core constructs in this study had an alpha level $> .87$, which indicated that the item sets were measuring their intended constructs with a high degree of internal consistency. The predictive models were analyzed using ordinary least squares regression with heteroscedasticity-robust standard errors, which is a more appropriate method

to estimate the precision of the estimated coefficients than transformation of the outcome variable, and retains the interpretability of the reported beta and B values, particularly when the data is divided into subgroups based on varying sample sizes, as is the case for the regional groupings in this dataset.

7. Findings

7.1 Respondent profile

Characteristic	Category	N	%
Region	Geita	518	47.7
Region	Mara	338	31.1
Region	Mbeya	116	10.7
Region	Njombe	102	9.4
Region	Simiyu	12	1.1
Region	Kilimanjaro	1	0.1
School_Location	Rural	874	80.4
School_Location	Urban	213	19.6
Subject	Mathematics	560	51.5
Subject	Science	527	48.5
Role_Group	Classroom Teacher	566	52.1
Role_Group	School Leadership/Management	521	47.9
Teaching_Experience	10-15 years	326	30.0
Teaching_Experience	15+ years	259	23.8
Teaching_Experience	5-10 years	230	21.2
Teaching_Experience	1-5 years	224	20.6
Teaching_Experience	Below 1 year	48	4.4
Qualification	Certificate	608	55.9
Qualification	Diploma	339	31.2
Qualification	Degree	134	12.3
Qualification	Masters	5	0.5
Qualification	Other	1	0.1

7.2 Scale reliability and descriptive findings

The reliability of all three core scales was high. The alpha values for curriculum alignment were .875, pedagogical suitability were .877, and the overall material quality were .888. Overall results indicated a generally positive mean score: curriculum alignment $M = 3.026$, pedagogical suitability $M = 2.997$ and overall mean score $M = 3.106$ on a

1-4 scale. Pedagogical suitability was the lowest mean of the three, indicating that rather than simply the existence of materials, pedagogical suitability of materials is a central problem of improvements.

It is worth emphasising because this is part of a frequently held policy belief that the quality of STEM materials is a supply issue and can be addressed by procurement volume.

The data show, however, that even when the materials are used in classrooms, as the majority of the respondents in this sample reported, teachers feel measurable concerns about their materials' design for Standards 3-6 learners. Finally, the reliability values shown in Table 2 should be noted: internal consistency (alpha) of .90 or above at the .80 level

for all three of these constructs means that respondents were answering the constructs underlying each set of items in a highly consistent manner, which adds to one's confidence that the group differences and regression results reported in the subsequent subsections represent actual differences in perception rather than measurement noise.

Construct	Items	Mean	SD	Alpha	Agreement %
Curriculum alignment	Q9, Q10, Q11, Q12, Q13	3.026	0.564	0.875	82.8
Pedagogical suitability	Q14, Q15, Q16, Q17, Q18	2.997	0.599	0.877	79.9
Overall material quality	Q49, Q50, Q51, Q52, Q53	3.106	0.556	0.888	87.1

7.3 Correlation findings

The correlation results were in favour of H1 and H2. Perceived overall material quality was positively and highly correlated with curriculum alignment ($r = .615$, $p < .001$). The correlation between pedagogical suitability and the perceived overall material quality was also very high and positive ($r = .596$, $p < .001$). Among the three constructs, curriculum alignment and pedagogical suitability were highly correlated ($r = .753$, $p < .001$), which suggests that teachers' perceptions of the two dimensions of STEM learning-material quality were close. Because of this association, the diagnostics of multicollinearity were investigated prior to interpretation of the joint regression coefficients. The Variance Inflation Factor for the curriculum alignment predictor was 2.31, and the corresponding tolerance value for this predictor was 0.433. The VIF for the pedagogical suitability predictor was also 2.31, with a corresponding tolerance value of 0.433 for this predictor. The VIF values were low (less than the

conventional level of 5) and tolerance (greater than 0.20) showed that the multicollinearity was not severe enough to affect the stability and interpretation of the regression coefficients.

The correlation between curriculum alignment and pedagogical suitability is strong ($r = .753$) and significant for practical purposes, in addition to being used as a preliminary step for the regression analysis described in Section 7.4. It implies that, in practice, Tanzanian Standards 3-6 teachers do not view them as distinct characteristics of a resource: a resource which is found to be well matched to the syllabus is also often pedagogically usable, and the other way round. This is possible because the design process is the same for both properties. It does introduce a methodological question, however, which is discussed in Section 6: can both constructs be included in a single regression model as separate predictors without the standard errors of the regression coefficients being biased?

Variables	Pearson r	p-value
Curriculum alignment - Pedagogical suitability	0.753	<.001
Curriculum alignment - Overall material quality	0.615	<.001
Pedagogical suitability - Overall material quality	0.596	<.001

7.4 Regression results

H3 was supported by the standardized regression model. The correlations between curriculum alignment and pedagogical suitability accounted for 41.9% of the overall material quality. Pedagogical suitability ($\beta = .307$, $p < .001$) and curriculum alignment ($\beta = .384$, $p < .001$) were the next two standardized effects. Both curriculum alignment ($B = .375$, $p < .001$) and pedagogical suitability ($B = .282$, $p < .001$) continued to be significant predictors in the adjusted model controlling for location, subject, role, experience, qualification and region. The model explained

41.9% of the variance in overall material quality; the remaining variance may be due to factors not measured by the constructs used in the present study, such as material availability and quantity, classroom-level implementation support and implementation leadership practices, which are explored in more detail in the companion papers in this series.

The appropriateness of having both components of curriculum alignment and pedagogical suitability in the same regression model was also supported by the diagnostic assessment. All VIFs were <2.31 and all tolerance values

were >0.433 . The results showed moderate overlap among predictors, as their correlations were high but not high enough to be considered problematic multicollinearity. Pedagogical suitability and curriculum alignment may then be seen as being related but sufficiently different that they are statistically separate determinants of the perceived overall quality of the material. The substantial coefficients of both variables are thus a measure of the independent contribution of each variable to explaining variation in teachers' perceptions of material quality.

The relative ranking of the two standardized effects, in favor of curriculum alignment over pedagogical suitability as a predictor of overall quality, conforms to a plausible sequencing interpretation: Teachers might first consider

whether a material is generally appropriate for the Standards 3-6 STEM curriculum and then consider whether the material is pedagogically appropriate to teach what it is designed to teach. When a material does not pass the first, most basic, measure of relevance, the pedagogical cleanliness is of relatively minor importance to the teacher's overall evaluation; when it does pass that criterion, pedagogical appropriateness is the next important factor in a teacher's evaluation. This interpretation of the sequencing is provided as a possible interpretation of the coefficient pattern and not as a directly tested hypothesis, and would be a logical extension of future qualitative research with think-aloud or cognitive-interview techniques involving teachers rating specific materials.

Predictor	B / beta	Robust SE	t	p-value
Curriculum alignment	0.384	0.051	7.586	$<.001$
Pedagogical suitability	0.307	0.050	6.168	$<.001$

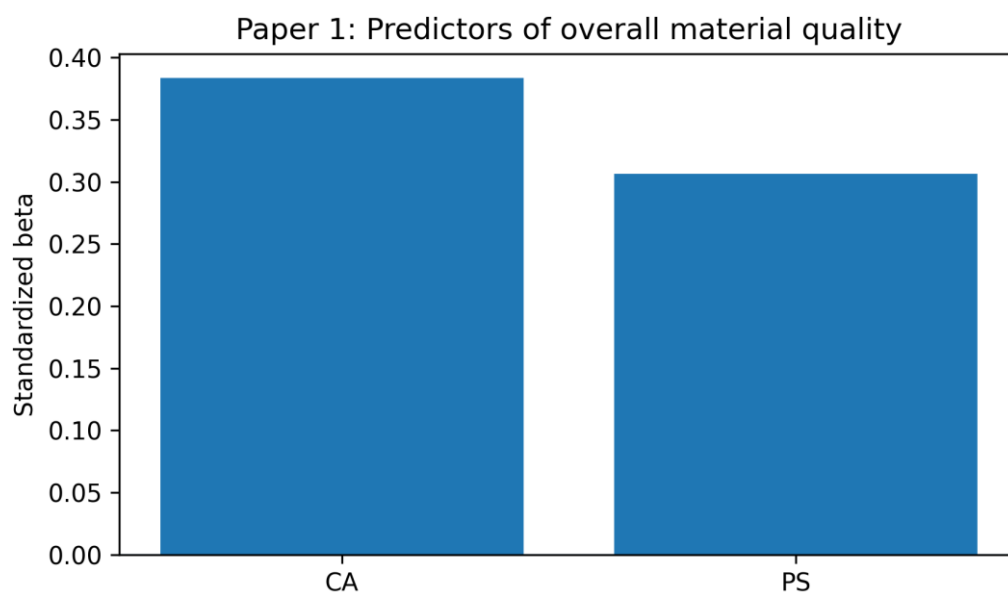


Figure 2. Standardised effects of curriculum alignment and pedagogical suitability on perceived overall material quality.

Predictor	Unstandardised B	Robust SE	t	p-value
Experience: 15+ years	0.087	0.040	2.171	0.030
Curriculum alignment	0.375	0.050	7.583	$<.001$
Pedagogical suitability	0.282	0.046	6.140	$<.001$

7.5 Group comparisons

There were few but meaningful subgroup differences when compared to group. The most statistically significant difference for the three core constructs was due to role group for pedagogical suitability ($p = .012$) which had a small effect size. Also very small, but significant, was the difference between the overall quality of materials based on the teaching experience of the teacher ($p = .038$). The results show that the constructs of material quality do not so much require a demographic grouping as are more adequately explained by it.

This is not an uninteresting result but a theoretically important null leaning result. If teacher and school variables had demonstrated strong, broad relationships with material-quality perceptions, that would have indicated that perceptions were more likely to be related to the teacher and/or school than to the materials themselves, making any policy response that focuses on material redesign more complicated. Rather, the rather limited and small number of significant differences among the subgroups suggest that the educators in these various positions, with their different experiences, are evaluating broadly similar aspects of the materials they teach, rather than simply reflecting a quirk of who teaches what. This finding is explored in considerably more detail, with a broader range of teacher and school attributes, in a companion paper to this one in the series, which focuses on teacher characteristics.

7.6 Teacher-reported challenges and improvement priorities

The quantitative findings were reinforced by the open-ended responses. The most common primary problem was a lack of teaching/learning materials and books (54.4%). The most common improvement suggestion was to provide more learning materials/books (56.9%). Overcrowding, curriculum/content complexity, ICT/digital resources, infrastructure and teacher training were also mentioned by the teachers.

The theme of the lack of materials was by far the most frequent, mentioned by over half of all respondents as their key problems, and this is reflected in the quantitative pattern mentioned above, but there is also a significant qualitative dimension, as not only did many teachers rate the materials somewhat less suitable, but many also found the available materials completely lacking in quantity. The distinction between design quality and supply quantity is more likely to be confused when considering policy, with different remedies: design quality is addressed through the curriculum-development and quality-assurance processes discussed in Section 8; supply quantity is addressed through procurement and distribution budgets. It is a fact that both concerns occur in the same database set; this means that neither of these remedies is likely to be effective on its own.

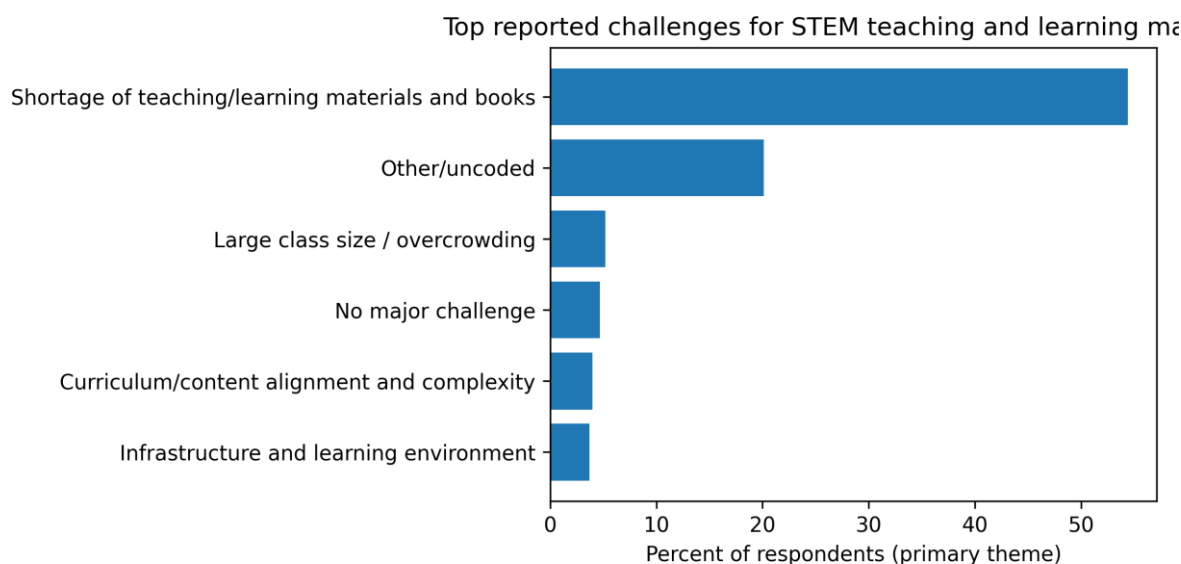


Figure 3. Top reported challenges for STEM teaching and learning materials.

Theme	N	%
Shortage of teaching/learning materials and books	591	54.4
Other/uncoded	218	20.1
Large class size / overcrowding	57	5.2
No major challenge	51	4.7
Curriculum/content alignment and complexity	43	4.0
Infrastructure and learning environment	40	3.7
ICT/digital resource constraints	31	2.9
Language/communication barriers	23	2.1

Theme	N	%
Increase supply of learning materials/books	619	56.9
Other/uncoded	234	21.5
Improve curriculum/material design and age-appropriateness	63	5.8
Provide ICT/digital resources	47	4.3
No suggestion / no change	32	2.9
Government funding/procurement support	26	2.4
Improve infrastructure/classrooms	21	1.9
Teacher training / CPD	21	1.9

8. Discussion

Results indicate that teachers' perceptions of the quality of the STEM materials are significantly influenced by perceived curriculum alignment and pedagogical appropriateness. This is in theory consistent with the principles of constructive alignment (Biggs, 1996), which states that the materials should link curriculum goals, classroom activities, and assessment goals. The results also substantiate the coherence argument for curriculum in mathematics and science education, where coherence, focusing and sequencing are at the heart of quality (Schmidt et al., 2005).

These findings are particularly significant in the Tanzanian policy context. The 2023 policy edition highlights the importance of quality education, competence, science and technology and practical skills (MoEST, 2023a). In addition, future PSTs are expected to develop skills in teaching and learning competencies through the translated teacher education curriculum they received which includes the learning of teaching and learning resources in Mathematics and Science for competence development (MoEST, 2023b). The gap between learning materials and pedagogically appropriate is something that must be bridged by teachers'

improvisation if there is no alignment. This can be achieved, but may result in uneven quality of the schools and regions.

8.1 Reading this paper alongside its companion papers

Since this paper, the accompanying paper on teacher characteristics, and the accompanying paper on teacher agency and resilience are all based on the same 1,087-respondent sample, it is important to make it clear how the results of these three papers relate to each other. The current paper shows that the level of curriculum alignment and pedagogical suitability are good, statistically sound predictors of perceived material quality. The teacher-characteristics companion paper sets the stage for the fact that the demographic and professional background constructs have a relatively small impact on the explained variance in the same and related outcome constructs. Together, these two findings are complementary, rather than conflicting: They suggest that the quality of perceived STEM materials in Tanzanian Standards 3-6 classrooms is more influenced by the characteristics of the materials themselves than by who is using them. The teacher-agency companion paper provides another, complementary layer of evidence that teachers consistently improvise and adapt their lessons to what they have been provided with, and thereby explain how classrooms continue to be pedagogically

effective despite sub-optimal ratings of alignment and suitability. Each of the three papers is (individually) only partial in addressing the issue of the quality of STEM material in Tanzanian primary schools; the three papers are three different lenses on the same topic, and hence, treated as an integrated sequence or as a set rather than as three independent papers.

Analytic process of this paper, from descriptive and reliability statistics, to bivariate correlation, to multivariable regression, and finally to qualitative theme analysis was constructed in a way that each step of evidence had to be consistent with each other in order to endorse the central claim of this paper. If this pattern of regression coefficients were mixed, with one being strong and another weaker, or for that matter if there were a single strong regression coefficient, but with no consistent correlational evidence, or if there were a single strong regression coefficient, but if teachers then gave unprompted written comments which were inconsistent, this would warrant significantly more caution than is warranted here, where teachers' numerical ratings, their pairwise correlations, their joint regression behaviour, and their own unprompted written comments all support the same substantive conclusion: that the effectiveness of STEM materials in relation to the curriculum and the pedagogical design for Standards 3-6 students is a critical issue for how they are judged in practice, yet that judgement is only weakly related to who is judging them.

Last but not least, this paper's contribution should be understood in the context of Tanzania's STEM policy vision and goal. The first section began with an observation that STEM education is identified nationally and internationally as a key contributor to the human capital and technological readiness. That ambition can only be achieved, classroom by classroom, as an ordinary, unglamorous process of the opening of a workbook, an attempt at an exercise, and the work that both the teacher and the learner undertake in front of whatever material is laid out in front of them that day. This paper's evidence base, culled from over a thousand of those daily interactions in the classroom over the six regions, is presented as a grounded empirical contribution to the larger, national, ambition: policy aspiration and classroom material design must go hand-in-hand if the goals of Standards 3-6 STEM teaching in Tanzania are to be met.

Readers primarily from a curriculum studies background may be best served by the theoretical discussion in Sections 4 and 5 as this provides a context for the empirical results in the paper. The discussion of resource distribution and resource supply in Sections 7.6, 8.3 and the Recommendations may be most directly actionable for readers coming from a development-economics or education-policy background. Those readers who have a background in measurement and/or psychometrics will be most interested in the reliability and correlation and

regression diagnostics that are discussed throughout Sections 6, 7, and 8.7. The paper has been organized to appeal to all these readerships, in that each of the above can read and grasp the sections of the paper that are relevant to them, but they will still see, in the Discussion and Conclusion, the paper's integrated central argument that quality of STEM material in Tanzania should be understood, measured, and improved as a joint function of curriculum alignment and pedagogical suitability, supported by adequate material supply, and by teachers who are capable of interpreting and adapting any materials they are given.

There were few responses with a low number of materials, demonstrating that there is no dissociation between alignment and suitability and resources. Some materials may be positively rated but teachers may also face a lack of quantity, access, ICT resources, and/or the materials may not match the classes' size. Thus, there is a need to enhance the quality of STEM materials and ensure that these are available equitably. This material-centred account needs to be read in conjunction with teacher-centred perspectives: teacher cognition studies indicate that teachers' interpretation and application of any particular set of materials is influenced by their knowledge, beliefs, and professional judgement (Borg, 2003); professional capital theory suggests that professional judgement is one part of a more extensive body of human, social, and decisional capital that teachers bring to curriculum implementation (Hargreaves & Fullan, 2012); and ecological accounts of teacher agency see adaptation as an outcome that is not solely the teacher's responsibility but is co-produced by the teacher and the material and structural conditions within which the teacher is operating (Priestley, Biesta, & Robinson, 2015). Taken together, these views indicate that alignment and pedagogical appropriateness are conditions that are required, but not sufficient, for ensuring the quality of a body of material will lead to quality STEM teaching.

8.2 Theoretical integration: why alignment and suitability jointly predict quality

The results of the regression analysis (see Section 7.4) were easy to interpret within the theoretical framework presented in Section 4. Constructive alignment theory suggests that materials that are not linked to the learning outcomes will be evaluated as poor in quality, no matter how pedagogically sound they are, simply because they don't pass the "curriculum relevance test. The use of perfectly matched content, as a result of which it can be judged as lower quality if it is not presented in a form that can be accessed by learners, can be predicted independently by theory in the field of instructional design. Even after controlling for the other and after controlling for teacher and school characteristics, both constructs remained significant predictors, suggesting that they tap into real aspects of material quality, rather than two indicators of the same

construct, although there was a high intercorrelation between the two constructs ($r = .753$).

8.3 Pedagogical suitability as the comparatively weaker construct

The slightly lower score for pedagogical suitability ($M = 2.997$) than for curriculum alignment ($M = 3.026$) is not a large absolute difference but is theoretically interesting. It suggests that the simpler design task of matching content to syllabus statements has been relatively more well-resolved than the more challenging design task of making pedagogically useful content more accessible to Standards 3-6 learners in very different school settings. This matches the argument of Ball et al. (2008) that having specialised pedagogical knowledge is more difficult than ensuring content accuracy, and confirms the international research of Margot and Kettler (2019) that integration and ease of use, not subject coverage, are the longer-term obstacles to STEM material use. It also has an immediate practical implication: strengthening these STEM materials in Tanzania will likely benefit more if it focuses on pedagogical scaffolding, worked examples, and illustrations that are relevant to Tanzania, rather than it focusing only on the checks on the extent to which the content in the syllabus has been covered.

8.4 Regional evidence and the limits of curriculum-document reform

In this study, the pattern that emerged was that while alignment with the curriculum documents was important, so too were pedagogical design and resource conditions. The pattern that emerged in this study is consistent with recent patterns found in the region, in that material quality depends upon pedagogical design and resource conditions as well as curriculum-document alignment. Kasuga and Kalolo (2025) discovered that remodelling curriculum at the regional level in Tanzania resulted in little change at the classroom level, despite no additional investment in resources or teachers' preparedness. A similar conclusion was drawn by Wawire et al.'s (2025) multi-country synthesis of competence-based-curriculum systems in seven African countries. This combined with the present findings suggests a broader regional lesson: improving the quality of STEM material, experienced by teachers, is not just a matter of revising curriculum documents, but it is also a necessary process that must be invested in on an ongoing basis, separate from curriculum policy revision.

8.5 Implications for curriculum development and material design practice

For the curriculum developers and publishers, the findings suggest that pedagogical suitability should be made a separate criterion to be evaluated and separately scored in quality assurance processes for STEM materials and not be used as a proxy for quality. Practical activities that make sense of the evidence reviewed here involve pre-testing draft

documents with a select sample of teachers who are teaching Standards 3–6, incorporating classroom teacher feedback loops into material development teams as recommended in Section 10, and developing material that includes and leverages the subject-specific pedagogical content knowledge described by Shulman (1986) and Ball et al. (2008) and is not assumed to be supplied completely by the classroom teacher who is teaching it.

8.6 Implications for teacher education and professional development

In addition to material design, the findings of this study have implications for teacher preparation and support of teachers implementing Standards 3-6 with any materials they are provided – even materials that do not meet the standards of alignment and suitability referred to above. The fact that most respondents in this sample have certificate-level qualifications opens up another lever of change, that of pre-service and in-service teacher education, besides material redesign itself. Teachers' capacity to critically assess how well a material is aligned to intended learning outcomes and to adapt, without compromising curriculum coherence, content that is aligned but not fully satisfactory is not the same as the problem encountered on the supply side. This aligns with Kafyulilo and Tilya's (2019) two-stage demonstration that there is measurable capacity building as a result of structured and collaborative professional development for teachers to work with STEM content and technology, and points to the need to explore content quality improvement and teacher professional development as complementary approaches to capacity building.

9. Conclusion

The study findings suggest that curriculum relevance and pedagogical appropriateness are major factors influencing teachers' perceptions of quality of STEM learning materials in Tanzanian Primary Schools. In general, the materials were rated positively by the teachers, and the relatively weak aspects were pedagogical suitability. The problem of lack of teaching and learning materials is the prevailing practical problem. The focus of material development for the Standards 3-6 STEM education should be on: curriculum coherence, learner fit, practical activities, local examples, visual clarity, and equitable supply.

The conclusions have significance beyond the Tanzanian context. The conflict between curriculum ambition and material design capacity is found in many African countries that are undergoing competence-based reforms in their education systems as discussed in the literature reviewed in section 5. The study will thus not only provide a national evidence base for curriculum developers and policy makers in Tanzania but also provide an evidence-based, quantitatively grounded case study that can be shared with similar evidence from other regions as the wider research

community seeks to develop a knowledge base about what high-quality STEM content looks like and how it might be measured, monitored, and improved on a large scale.

10. Recommendations

The structured review process must include a set of distinct scores that measure alignment with and progression to the learning outcomes in Standards 3-6, and pedagogical suitability from the teacher's perspective, but not a single composite score, to allow for identification of deficiencies in either component and to provide targeted feedback. The process of reviewing STEM materials must be systematic and based on a review protocol that awards distinct scores for alignment to and progression to the learning outcomes in Standards 3-6, and for pedagogical suitability from the teacher's perspective, and not a single composite score, so that deficiencies in either component are visible and actionable.

These five recommendations are not all equally important and it is likely that a series of actions in the order recommended will prove to be more effective than simultaneous actions using limited resources. The one that is most clearly actionable is to adopt a structured, dual-scored review process for materials already in the development pipeline which does not involve new funding, but rather a change in process. The recommendations for distribution-equity and feedback-loop need budgetary resources and are more likely to be achieved within a medium-term planning period. The recommendation for professional development can follow either, as much of the change is teacher-training structures and not necessarily the redesign of materials. The sequencing recommendations outlined here are offered in the spirit of making what this study has uncovered actionable in the context of realistic budget and administrative constraints of the curriculum authorities in Tanzania, not as a prioritized laundry list of equally weighted priorities.

11. A note on effect size interpretation

In addition to reporting statistical significance, where possible, effect sizes have been reported throughout Sections 7 and 8, following current methodological advice not to rely on the p-value alone. Of moderate to large size, the standardized regression coefficients of .384 and .307 for the variables curriculum alignment and pedagogical suitability, respectively, and a high overall model R-squared of .419 for a two-predictor model in survey-based education research are by conventional social-science benchmarks. Our central finding that curriculum alignment and pedagogical suitability are substantively important predictors of perceived STEM material quality gains the present study its practical, as well as its statistical weight, from this combination of statistical significance and meaningful effect size and a well-powered sample. The

differences between subgroups that were found in Section 7.5, although sometimes statistically significant with such a large sample size, were consistently very small in effect size and this separation between 'statistical' and 'practical' significance is maintained throughout the paper in interpreting its findings.

It is important that TIE and responsible authorities adopt a common material-quality checklist, both at the pre-publication design stage and in periodic review after distribution and apply this checklist consistently, in order to ensure that material quality is regarded as an ongoing quality assurance responsibility rather than a one-off approval decision.

Government and local authorities should focus on the distribution of Mathematics, Science and Technology (MST) and related STEM materials in schools with an emphasis on equity – particularly in rural schools and classrooms with high enrolments and report distribution using these same categories as were drawn in the sampling frame for this study in order for future distribution of materials to be compared against the baseline patterns reported here.

As a professional-development resource, teachers' current adaptation practice should be drawn on when adapting materials in practical ways without compromising on the coherence of the curriculum, rather than adaptation just being a symptom of material shortfall that needs to be fixed.

Future development of these materials needs to incorporate teacher feedback loops prior to widespread distribution, such as through pilot testing within a carefully selected combination of rural and urban schools, with feedback gathered on curriculum fit and teacher motivation to use the material on instruments similar to the one employed in this study.

12. Limitations and further analysis

This study relied on teacher perceptions instead of classroom observations, an analysis of the textbooks, or learner achievement data. Data are cross-sectional and do not make causal statements. Thematic coding is open and supported by keywords and will be fully manually coded before it is submitted for review, preferably including an inter-coder agreement statistic. Common method variance is not ruled out and should be taken into consideration when assessing the strength of the associations observed since all three of the measurements of curriculum alignment, pedagogical suitability and overall quality of the material was assessed with the same self-report survey given at one time point. Possible future analyses include additional analyses of confirmatory factor analysis and structural equation modelling, as well as a direct review of textbooks/workbooks in the schools surveyed.

These limitations lead to the following research agenda for the future of concrete. One priority next step that could be pursued is to replicate this study with a smaller sample size, but with a follow-up mixed methods component that integrates direct classroom observation, a structured textbook audit that assesses the alignment of the curriculum and suitability of the pedagogical approach as measured in the survey items used here, and interviews with a purposefully selected subsample of teachers who reported that their textbooks had scores that were either in the high or low extreme ranges. This could help determine whether teachers' self-reported perceptions of the textbook's material quality align with what is observed in the classroom and whether these perceptions align with the school's pedagogical approach. A second priority would be a longitudinal study over a single cohort of Standards 3-6 teachers to directly test the effects of measures of curriculum improvement (measured and perceived increases in alignment/ pedagogical suitability) on perceived improvements in overall quality and, ultimately, on learner outcomes, over a time frame of curriculum-material revision. These are extensions of the present cross sectional study with a large sample base, but would overcome the causal-inference constraint of the cross sectional design.

Despite the following constraints, the study has some methodological advantages which should be taken into account. With 1087 respondents, this sample size is substantial compared to most published studies on STEM education in Tanzania, and allows for relatively small effects to be detected. Associations are reported with a validated and internally consistent multi-item instrument, which increases confidence in the reliability of the reported associations over single item and ad hoc instruments. The triangulation of the statistical results with spontaneous qualitative feedback from teachers on their struggle is a more robust design if compared to quantitative modelling or unprompted qualitative feedback. Lastly, the analysis of this paper, which focuses on the content of teacher materials, alongside two companion papers that focus on teacher characteristics and teacher agency, respectively, offers a more comprehensive, multi-perspectival understanding of STEM content quality, than any one of these three could have given using any one of these three lenses alone.

Another restriction is the composite presentation of the outcome variable. The overall material quality was assessed by a five item scale that was reliable, but combined several potentially independent aspects of material quality – the durability, the cost-appropriateness, the visual appeal, as well as the alignment and suitability. There are areas for future research that could help determine if curriculum alignment and pedagogical appropriateness are more predictive of aspects of perceived quality than others. Finally, this paper is one in a series of three papers examining the same data set, so readers should consider its

results as one data source to be interpreted in conjunction with the results of the other two papers, rather than as a confirmation of the results of those papers. The finding of the other two papers, on teacher characteristics and teacher agency, are complementary, but not corroborative, findings of the results presented in this paper.

References

1. Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407. <https://doi.org/10.1177/0022487108324554>
2. Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347–364. <https://doi.org/10.1007/BF00138871>
3. Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36(2), 81–109. <https://doi.org/10.1017/S0261444803001903>
4. Bybee, R. W. (2013). *The case for STEM education: Challenges and opportunities*. NSTA Press.
5. Dismas, S., Nzima, I., & Kimaro, A. (2023). Competence-based curriculum in Tanzania: District school quality assurance officers' understanding and practices. *International Journal of Educational Reform*. <https://doi.org/10.1177/10567879231217458>
6. English, L. D. (2016). STEM education K–12: Perspectives on integration. *International Journal of STEM Education*, 3, Article 3. <https://doi.org/10.1186/s40594-016-0036-1>
7. Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school*. Teachers College Press.
8. Kabote, S. J., Niboye, E. P., & Nombo, C. I. (2014). Performance in mathematics and science subjects: A gender perspective for selected primary schools in rural and urban Tanzania. *International Journal of Gender and Women's Studies*, 2(3), 87–105. <https://doi.org/10.15640/ijgws.v2n3a6>
9. Kafyulilo, A. C., & Tilya, F. (2019). Learning technology by design: Experiences from in-service science and mathematics teachers in Tanzania. *Journal of Humanities and Social Sciences*, 8(1), Article 2. <https://doi.org/10.56279/jhss.v8.i1.2>
10. Kasuga, W., & Kalolo, J. F. (2025). Competency-based curriculum in Tanzania: Charting dilemmas for successful implementation. *Cogent Education*, 12(1), Article 2563711. <https://doi.org/10.1080/2331186X.2025.2563711>

11. Mafugu, T., Tsakeni, M., & Jita, L. C. (2023). Preservice primary teachers' perceptions of STEM-based teaching in natural sciences and technology classrooms. *Canadian Journal of Science, Mathematics and Technology Education*, 22(4), 898–914. <https://doi.org/10.1007/s42330-022-00252>
12. Margot, K. C., & Kettler, T. (2019). Teachers' perception of STEM integration and education: A systematic literature review. *International Journal of STEM Education*, 6, Article 2. <https://doi.org/10.1186/s40594-018-0151-2>
13. Ministry of Education, Science and Technology. (2023a). *Education and training policy of 2014: 2023 edition*. Government of the United Republic of Tanzania.
14. Ministry of Education, Science and Technology. (2023b). *Mtaala wa Mafunzo ya Stashahada ya Ualimu Elimu ya Msingi* [Curriculum for diploma in primary teacher education]. Tanzania Institute of Education.
15. Ministry of Education, Science and Technology. (2024). *National basic education statistics in Tanzania (BEST) 2024: National data*. Government of the United Republic of Tanzania.
16. Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.
17. Nzima, I. (2016). *Competence-based curriculum (CBC) in Tanzania: Tutors' understanding and their instructional practices* [Doctoral dissertation, Linnaeus University].
18. Nzima, I. (2018). Tutors' interpretations of competence-based curriculum in Tanzania: Implications for practice in teacher education. *Papers in Education and Development*, 36, 12–21.
19. President's Office – Regional Administration and Local Government. (2025). *Pre-primary, primary, secondary, adult and non-formal education statistics 2025: Regional data*. Government of the United Republic of Tanzania.
20. Priestley, M., Biesta, G., & Robinson, S. (2015). *Teacher agency: An ecological approach*. Bloomsbury Academic.
21. Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research*, 75(2), 211–246. <https://doi.org/10.3102/00346543075002211>
22. Schmidt, W. H., Wang, H. C., & McKnight, C. C. (2005). Curriculum coherence: An examination of U.S. mathematics and science content standards from an international perspective. *Journal of Curriculum Studies*, 37(5), 525–559. <https://doi.org/10.1080/0022027042000294682>
23. Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14. <https://doi.org/10.3102/0013189X015002004>
24. UNESCO. (2024). *Global education monitoring report 2024/5: Leadership in education: Lead for learning*. UNESCO. <https://doi.org/10.54676/EFLH5184>
25. Wawire, B. A., Kiambati, F., Muhia, N., Gathoni, G., & Abuya, B. (2025). Competence-based curriculum implementation in Africa: A scoping review of pedagogical and assessment practices. *Journal of Curriculum Studies*. <https://doi.org/10.1080/00220272.2025.2523429>
26. Ministry of Education, Science and Technology. (2023a). *Education and training policy of 2014: 2023 edition*. Government of the United Republic of Tanzania.
27. Ministry of Education, Science and Technology. (2023b). *Mtaala wa Mafunzo ya Stashahada ya Ualimu Elimu ya Msingi* [Curriculum for Diploma in Primary Teacher Education]. Tanzania Institute of Education.
28. Ministry of Education, Science and Technology. (2024). *National basic education statistics in Tanzania (BEST) 2024: National data*. Government of the United Republic of Tanzania.
29. President's Office - Regional Administration and Local Government. (2025). *Pre-primary, primary, secondary, adult and non-formal education statistics 2025: Regional data*. Government of the United Republic of Tanzania. UNESCO. (2024). *Global education monitoring report 2024/5: Education leadership Lead for learning*. UNESCO Publishing. <https://doi.org/10.54676/EFLH5184>
30. Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special?