

Folding Imperfections: Mixed Methods Insight into the Limits of Curriculum Change for the Trinidad and Tobago Primary School Curriculum Rewrite

Jerome De Lisle Sharmila Harry

The University of the West Indies, St. Augustine

Abstract

We theorize on the limits of curriculum change in countries of the Caribbean and Global South using integrated evidence derived from a qualitatively driven, mixed methods evaluation (MME) of the Primary School Curriculum Rewrite (PCR) in Trinidad and Tobago. We use a folding metaphor and implementation science to explain how an imperfect system might produce a flawed curriculum product that misdirects sensemaking and generates resistance during curriculum change. The formative process evaluation consisted of two complete research projects, with the second qualitative project prioritized.

Project 1 was a quantitative survey of 454 teachers in 58 schools. Project 2 was a multiple case study of subjectivities-in-action in 9 schools of varied implementation outcomes. Findings from Project 1 showed mostly acceptable implementation outcomes and small associations between antecedents and outcomes. However, findings from Project 2 suggested that 6 of 9 schools were classified as nonusers, defined as implementing institutions with most teachers still exploring or installing the new curriculum or reverting to the old curriculum. In the thematic, cross-case analysis, participants at the various sites identified curriculum imperfections, inadequate training, and work intensification as core barriers to change. To achieve curriculum change, Caribbean countries such as Trinidad and Tobago will require improved systems for curriculum engineering.

Keywords: curriculum change, implementation gap, sensemaking, formative process evaluation, mixed methods

Introduction

Education systems in the Caribbean and the rest of the Global South have highly centralized bureaucracies meant to manage curriculum-making (Montero-Sieburth, 1992). The installation of new curricula is often seen as some sort of ground-breaking and transformative societal act (Jansen, 1998). High expectations may easily arise from the perceived power of being able to mandate national policy. However, political power and technocratic mandates are

usually insufficient for uniform and effective implementation (Bantwini & Letseka, 2016). The limits of curriculum change are clearly seen across the Global South, from the Caribbean to sub-Saharan Africa (Altinyelken, 2010; Bellino, 2016; Jennings, 1993, 1994, 2017).

Curriculum change includes alterations to the design, development, and implementation of existing curricula. As a process, it is often explicitly political, meant to re-shape "relationships

between individuals and institutions of the nation-state through the selection and organization of school knowledge" (Rosenmund, 2006, p. 177). This perspective on change focuses attention not just upon the final curriculum documents, but also upon implicit and explicit purposes (Fullan, 1983; Nakabugo & Siebörger, 2001). Change will make manifest differences in the intended, enacted, assessed, and learned curriculum. These differences are the very essence of the implementation gap (Bantwini, 2012; Porter & Smithson, 2001; Yan, 2012).

Very often, then, well-intentioned curriculum policies do not translate into classroom reality (Ball et al., 2012). Large-scale curriculum change is problematic everywhere (Grundy, 2002). It could be, however, that implementation failures are either more common or have greater impact in countries of the Global South (Altinyelken, 2010). Considering the possible impact of implementation failure, the Inter-American Bank (IDB) has estimated that it might take 29 years for a country like Trinidad and Tobago to reach the average science score in the Programme for International Student Assessment (PISA) (Bos et al., 2016). Therefore, Caribbean countries can ill afford repeated implementation failure. This paper explores implementation failure for the IDB funded Primary School Curriculum Rewrite (PCR), a revision of the national primary school curriculum meant to foster integrated, seamless learning.

Theoretical Framework

Implementation Theory Matters

Curriculum implementation is the planned introduction of a curriculum innovation into an education space that is meant to achieve promised outcomes (Gouédard et al., 2020). An education space has physical, social, and cultural dimensions (Cook & Hemming, 2011). The implementation process can be explained by different models and factors (Century & Cassata, 2016; Durlak & DuPre, 2008). Saunders (2016) has argued that some factors might be less amenable to influence. Considering the importance of the teachers' role,

older models of curriculum implementation have often emphasized teacher concerns (Barrow & De Lisle, 2009). Notably, however, teachers have been excluded from key stages of curriculum development in Trinidad and Tobago (Gift, 1987).

Nevertheless, the complexity of context and practice in the Caribbean is not easily captured by older models of change, such as the Concerns Based Adoption Model (CBAM), which might see schools and teachers as locked into mechanical implementation stages progressing towards some plateau phase (Harry, 2007). It might be unwise, then, to neglect organizational and societal influences when studying curriculum implementation in the Caribbean (Harry, 2019).

Arguably, by using both a systems perspective and complexity theory, we can provide a more robust explanatory framework (Alhadeff-Jones, 2008; Bantwini, 2010). This is the essence of curriculum change from a complex adaptive systems perspective (Doll, 2008). A systems perspective will examine the whole, with subsystems and interconnections made explicit (Kenny & Cirkony, 2022; Maxwell et al., 2022). Therefore, although teachers might be considered critical to implementation, so too are intersecting and interacting institutional and contextual factors (Peurach, 2011).

Adopting a complex adaptive systems perspective for curriculum change in the Caribbean will focus attention upon building the infrastructure and processes needed for change (Beauchamp, 1972). These elements constitute a system for curriculum engineering designed to produce high-quality curricula (Gift, 1987). According to Gift (1984), a curriculum engineering system has all the elements to effectively perform 4 curriculum functions: (1) design, (2) development, (3) implementation, and (4) evaluation.

The complexity construct is also applicable to both innovation and context. For example, many curricula are multidimensional, with interdependent components and enactment requiring several emergent behaviours by stakeholders. This complexity makes change unpredictable (Morrison, 2003).

Additionally, in postcolonial contexts such as Trinidad and Tobago, intertwined historical and cultural factors also influence curriculum change (Hickling-Hudson, 2010; Lemke & Sabelli, 2008).

Newer implementation models might better explain change within complex adaptive systems. For example, Nilsen (2020) notes that implementation science (IS) frameworks are aligned to a systems approach and envisage interacting factors at different levels in the system. IS builds upon earlier theories of implementation failure such as diffusion theory, and considers the fidelity demands when implementing evidence-based interventions (Bertram et al., 2015; Nilsen, 2015). IS also provides well-defined constructs and tenets to guide successful implementation practice, paying attention to organizational factors such as implementation readiness and capacity (Wallace et al., 2008).

IS envisions change as a process with stages and core components, requiring drivers, teams, and the improvement process. The stages of exploration, installation, initial implementation, full implementation, innovation, and sustainability are iterative, recursive, and non-linear (Fixsen et al., 2009). This means that an individual can move back and forth between stages (Braithwaite et al., 2018). Implementation drivers are integrated and compensatory and provide the infrastructure of change, with three key leverage points: (a) competency, (b) leadership, and (c) enabling organization factors. Therefore, in exploring context as well as policy actors, IS is also congruent with a systems perspective (Huybrechts et al., 2024; Nilsen, 2020).

Beyond Traditional Implementation Theory

Three useful conceptual tools to guide implementation in the Caribbean are: (1) the folding metaphor for curriculum change (Webb & Gulson, 2013); (2) the theory on policy enactment in schools (Ball et al., 2012); and (3) sensemaking theory (Spillane et al., 2002). Webb and Gulson (2013) applied the Deleuzian folding metaphor to policy implementation (Conley, 2005). They identified policy intensions and policy enactments as the key processes.

Both policy intensions (the way the policy is sensed) and enactments fold onto and within policy subjects. However, Deleuze's original writings conceived of folding, unfolding, as well as refolding; with refolding being more than an opposite of folding (Deleuze & Strauss, 1991; Deleuze, 1993). Applied to curriculum change, folding, unfolding, and refolding is neither uncomplicated nor rational and can result in twisted knots when policy contests occur.

Applying this metaphor further, we argue that curriculum policy can fold onto policy actors and policy subjects with its many inherent contradictions, indeterminacies, and contingencies. Unfolding will reveal these imperfections, which are filtered through the self. Refolding over time will add and retain additional and different types of imperfections. Folding and refolding makes curriculum change subjective, intransient, and varied. Change becomes complex, hidden, and unpredictable, dependent upon the folding processes and perceived imperfections. All of this suggests that we cannot understand folding processes completely through quantitative cross-sectional data. To gain insight into variation across context and individuals, we must also explore subjective experiences and stories.

Ball et al.'s (2012) explanation of policy enactment also captures the complexity and unpredictability of implementation. Ball et al. (2012) reported on the difference between interpretations and translations of policy. Meanings are subjective and influenced by the different policy positions adopted, which in turn partly depends upon an individual's tenure and experiences. The point is that policy enactment is dependent upon filtering and selective focusing and there are likely to be different versions of policy enactment within schools.

We find the folding metaphor of Webb and Gulson (2013) central to understanding policy enactment by emphasizing the role of cognition and context, and the intimate connection between policy and individual purpose. For successful change, teachers must judge themselves in relation to some declared "better" practice. This makes identity and positionality critical to

successful change (Harré et al., 2009). Moreover, teachers are also required to actively decode and recode the intentions and desires of the policy. This, according to Spillane et al. (2002), is the essence of sensemaking—a process with both individual and situational elements that is imbued with emotion. We reason that deficient systems for curriculum engineering often produce imperfect curricula (intended, enacted, and assessed), which heighten experiences of ambiguity and conflict; signalling and redirecting sensemaking (Choi & Walker, 2018; Gift, 1984).

Why Models and Methods Matter when Evaluating Curriculum Change

In a system for curriculum engineering, the role of evaluation is to produce credible evidence for effective implementation planning (Harachi et al., 1999; O'Donnell, 2008). Evaluation evidence is needed to improve development and planning of curricula (Durlak & DuPre, 2008). For better planning, a key purpose is to better understand the process of implementation (Stake, 1976). However, evaluation has multiple, additional purposes. For example, another purpose is to discover whether a curriculum is fulfilling its original objectives. Glatthorn et al. (2018) considered evaluation to be necessary for judging merit and worth. However, worth can only be judged in relation to the context and will vary across different constituents (Lincoln & Guba, 1980).

There are several types of evaluation with different roles. For example, formative evaluation will help in program refinement, summative evaluation provides a comprehensive overview, process evaluation is focused upon the fidelity of implementation, outcome evaluation seeks actual results, and impact evaluation captures long-term effects. Based on our desire for better understanding of a complex setting and process, we used a formative process evaluation (Brand et al., 2019).

A formative process evaluation can provide insight into the black box of implementation, directly answering questions on whether

the curriculum works, why, and under what circumstances (Fullan, 1983; Saunders, 2016). The information derived can then be used for planning and redirecting curriculum improvement. Dyer (1999) has reminded us that it is the lack of thoughtful planning that leads to stronger resistance from subjects and results in unintended outcomes. Short-term strategies to address deficits will do little to prevent further dilution of policy efficiency.

In evaluation, the folding metaphor will direct attention to the need to measure implementation fidelity. The trueness of curriculum change remains a global concern, but is magnified in the Caribbean because teachers are often required to adopt global policy that sometimes conflicts with strongly held traditional beliefs and practices (Jennings, 2001; Rogan, 2007). The folding process expresses the difficulty of changing persistent traditional practice and highlights the need to investigate change using naturalistic methods to reveal hearts, minds, and hands (Thirsk & Clark, 2017).

Evaluation models are theories, approaches, and methods for achieving the intended evaluation purpose (Mertens & Wilson, 2012). In this study, we argue for using a *responsive evaluation model* to capture the nuances and subjectivities theorized in the folding metaphor. Stake (2004) described responsive evaluation as focused upon “orienting to the experience of personally being there, feeling the activity, the tension, knowing the people and their values” (p. 86). The model is therefore active, reactive, and dialogic, with a focus on values rather than on measurement (Freeman & Hall, 2012; Stake, 1976). Responsive evaluation explicitly emphasizes the uniqueness and social plurality of the curriculum, with different judgements and contradictory standards often held by various stakeholders and different participant groups (Abma & Stake, 2001).

Responsive evaluation, then, by privileging subjectivity and pluralism (on the surface), places greater value on qualitative approaches and is more congruent with the folding metaphor. At

the same time, Stake (2004) also acknowledged clients' demands for accountability in standards-based evaluation approaches, calling for a "combination of responsive and standards-based" designs (p. 94). Thus, Stake (2004) explicitly identified mixed methods evaluation (MME) designs as the preferred solution for managing subjectivity and accountability as competing demands. Although there is the possibility of divergent findings, these can lead to new insights and creative leaps in understanding (Greene et al., 1989).

The Present Study

The nation-state of Trinidad and Tobago is a small twin-island developing state situated in the southern Caribbean Sea, just off the coast of Venezuela. This Anglophone Caribbean country is classified as high-income by the World Bank, but collective capacity for change is very low. Scores on the Progress in International Reading Literacy Study (PIRLS) from 2011–2016 suggest that the Trinidad and Tobago education system is low performing by Organization for Economic Co-operation and Development (OECD) standards. However, there is evidence of significant improvement over the period (Mullis et al., 2017). Within the low performing Latin America and Caribbean region in PISA, Trinidad and Tobago might be regarded as one of the better performers (Bos et al., 2016).

The programme being evaluated is called the Primary School Curriculum Rewrite (PCR) (Trinidad and Tobago Ministry of Education [TTMoE], 2013a, 2013b). The PCR is part of a series of innovations in the seamless education project for Trinidad and Tobago, funded by the Inter-American Development Bank (IDB). The TTMoE developed this new curriculum with support from a Canadian-based consultancy firm (Erskine, 2013). The expectation was to implement this new curriculum in all 477 public primary schools in Trinidad and Tobago distributed across the seven education districts and one division. These geographical areas have very different characteristics and outcomes (De Lisle et al., 2010).

The new curriculum consisted of nine subject areas with eight focal areas: ICT infusion, learning styles, assessment modalities, 21st century skills, international standards, literacy and numeracy, alignment, and systemic support. Published documentation included curriculum guides, teacher guides, an overall framework, instructional toolkits, planning and recording documents, and resources such as manipulatives. Multiple components, integration, and student-centred learning contributed to innovation complexity (Morrison, 2003).

Table 1 shows a Theory of Action (ToA) for the new curriculum developed by the evaluation team. Listed are the core components and principles, linked to inputs, interactions, outputs, and outcomes. A primary focus of the ToA was curriculum integration. Although this innovation has been introduced in past Caribbean reforms, the PCR makes this integration central to enactment (Jennings-Wray, 1984; TTMoE, 2013a). As a key document, the implementation toolkit promoted primarily cross-disciplinary (interdisciplinary) themes, but also implicitly promoted other forms of integration (Drake & Burns, 2004).

In the implementation toolkit, teaching thematic lessons was associated with student-centred classes and formative classroom assessment (Brough, 2012; TTMoE, 2013a). The various curriculum innovations were expected to enhance love of learning, greater engagement, and increased teacher capacity; resulting in outcomes such as critical thinking, creativity, and citizenship (Trinidad and Tobago Parliament, 2014). Curriculum enactment requires pedagogical changes that might prove difficult in a postcolonial setting, where traditional teaching approaches dominate (Lam et al., 2013). Well-designed site-based professional learning might be the key to successful adoption (Polly & Hannafin, 2011).

The original curriculum development plan included a laboratory and field try-out. However, the political directorate insisted on immediate implementation in all schools directly after the laboratory try-out, circumventing the field try-out process (Lewy, 1990). The intention, then, was to scale up the curriculum by year-groups

Table 1*Theory of Action for the Trinidad and Tobago Primary Curriculum Rewrite (PCR)*

Stated core component	Input	Interaction	Output	Short term Outcome	Long Term Outcome
<i>Curriculum aligned to the MOE value outcomes</i>					
<i>Equity focused</i>		<i>Student centred learning</i>	<i>Greater inclusion and differentiation</i>		<i>Improved achievement in aligned summative tests</i>
<i>Fostering 21st century skills</i>	<i>Professional development for teachers</i>	<i>Differentiated learning</i>	<i>Teachers committed to and trained for new curriculum</i>		<i>Greater physical and emotional health</i>
<i>Emphasizing critical thinking/ problem-solving</i>	<i>New curriculum guides for each of the year levels</i>	Thematic, integrated teaching and learning	More active and student-centred classes	Enhanced love of learning	Greater creativity
Promoting integration (thematic)	A comprehensive teachers' guide and instructional toolkits	Active learning	Greater formative assessment practice	Greater engagement in lessons	Enhanced citizenship
<i>Making use of ICT</i>	<i>Institutional supports and building site leadership</i>	Problem solving activities	Thematic lessons	Enhanced capacity for integrated/ thematic learning	Greater critical thinking
<i>Ensuring cross-curricular development of literacy and numeracy</i>	<i>MOE supports and central administration and district leadership</i>	<i>Inquiry approach</i>		<i>Decreased achievement gap</i>	<i>Improved achievement in international large-scale assessments (ILSA)</i>
<i>Encouraging VAPA and citizenship</i>		<i>Assessment as and for learning</i>			<i>Improved economic competitiveness</i>
<i>Focusing on values and socio-emotional learning</i>					
<i>Aligned to ECCE and secondary curricula</i>					

Note. Core focus on integration highlighted.

across sites. This approach possibly risked low organizational readiness (schools could not be primed for change) and negative organizational cultures (as higher grades learnt of the difficulty of curriculum change). It might be that the political directorate was more concerned with the

utility and "value" of the new curriculum, with stakeholders having expressed some concern over the choice of consultancy firm ("Local graduates ignored", 2013). In 2015, with IDB funding, the Government of the Republic of Trinidad and Tobago (GORTT) used a competitive

public tender to hire an independent consultant firm. The firm was required to (a) evaluate the current status of implementation, and (b) inform further progress on implementation.

Research Design

Purpose and Evaluation Questions

The intent of this study was to (a) understand the relationships between factors, enactment, and outcomes in implementation, and (b) develop theory for curriculum change in the context (Century & Cassata, 2016). The initial list of evaluation questions was prioritized based upon suggestions from the client, the ToA, and a literature review. The three main research questions considered in this paper were:

1. What is the current status of the primary school revision project? (Mixed Methods)
2. What factors explained the degree of implementation of the curriculum change? (Mixed Methods)
3. What progress have teachers and leaders made in implementing school-wide curriculum change? (Qualitative)

The Mixed Methods Evaluation Design

The paradigmatic framework adopted for mixing was critical realism (CR) based upon: (a) understanding reality as complex, systemic, and contextual, (b) adopting a process approach to causality; and (c) using methodological tools such as triangulation (Gerrits & Verweij, 2013; Maxwell, 2024; Modell, 2009). Critical realism is also an especially useful paradigmatic framework for qualitative multiple case studies (Kessler & Bach, 2014; Maxwell, 2022; Parra et al., 2021). One benefit of CR is the adoption of a process approach to causality, which is different to the regularity approach conceived in post-positivism (Maxwell, 2021).

Sayer (2000) distinguished between intensive and extensive research and argued for the value of contextual causal explanations in the intensive (rather than the mainly descriptive) generalizations in the extensive. Hurrell (2014)

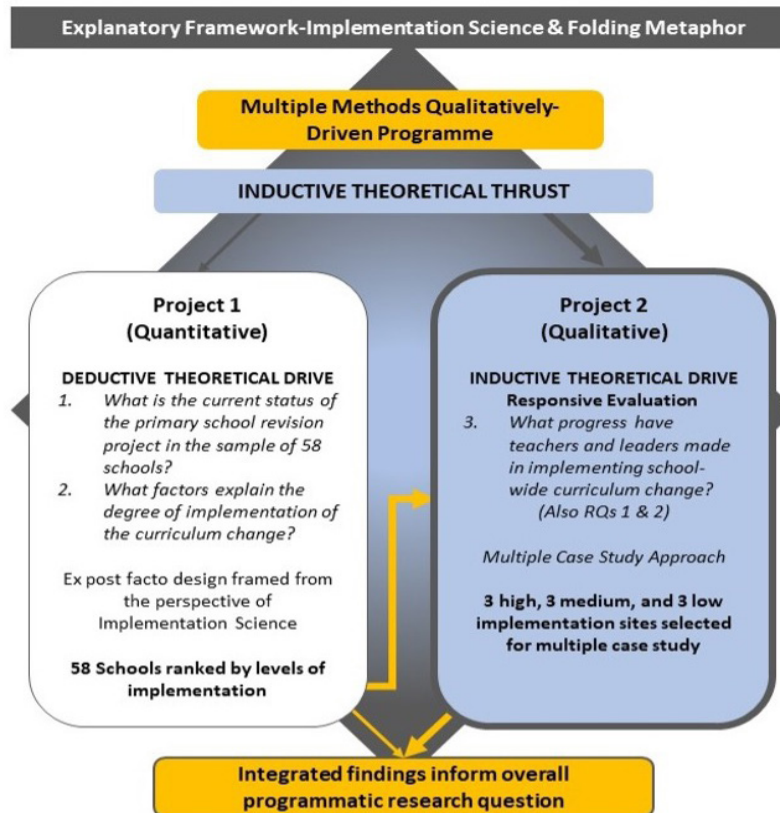
interpreted Sayer's (2000) recommendations to mean the rejection of inferential statistics and multivariate statistical modelling. However, both descriptive and inferential statistical methods can elucidate patterns of associations (demi-regularities) to complement analytic generalizations (Eastwood et al., 2016). Quantitative inferences may be used to generate tentative hypotheses which are then explored and reconstructed in the qualitative project. This process requires both abduction and retroduction (Meyer & Lunnay, 2013; Mukumbang, 2023).

Based upon the typology of Morse and Niehaus (2009), the evaluation design is called a *multiple methods programme*. Morse and Cheek (2014) define multiple methods as two or more complete projects attached to an overall aim. Although these authors conceived of separate quantitative and qualitative research questions, this study puts forward mixed methods and qualitative research questions. For this study, an overall inductive theoretical thrust was chosen to privilege the qualitative findings (Cheek & Morse, 2022). Morse and Niehaus (2009) defined theoretical thrust as the overall inductive or deductive intent of the programme; similar to, but larger than, the theoretical drive of a project.

The overall design is qualitatively-driven in that the qualitative component is given priority and emphasis (De Lisle, 2011; Hesse-Biber, 2022). There is increased attention to qualitatively-driven designs in mixed methods (Hesse-Biber, 2022; Morgan & Hoffman, 2021). Figure 1 shows the qualitatively-driven mixed methods programme, with separate and complete (but linked) quantitative and qualitative projects (Morse & Niehaus, 2009). The quantitative Project 1 conducted first was governed by a deductive theoretical drive and fulfilled the requirements of a standards-based evaluation. Project 2 was governed by an inductive theoretical drive. Following Stake (2004), Project 2 was meant to capture curriculum subjectivities-in-action. Mixed methods designs that prioritise the qualitative project even when conducted second are increasingly common in organizational science (Gibson, 2017).

Figure 1

The Evaluation Model and Mixed Methods Design for Studying the 2013 Trinidad and Tobago Primary School Curriculum



Quantitative data from Project 1 were used to purposefully select nine sites for the multiple case study of Project 2 (Stake, 2006). Creswell and Plano Clark (2018) called this quan→QUAL design *a case selection variant of the explanatory sequential design*. However, Morse and Niehaus (2009) better capture the intent of the design used in this study because it makes explicit the overall inductive intent and the complete but linked quantitative and qualitative projects. Integration was achieved through (a) linking of the projects, and (b) construction of meta-inferences combining findings from both projects.

Sample and Sampling Strategy

A sequential MME sampling strategy was implemented (Teddlie & Yu, 2007). The target population was all 477 public primary schools in Trinidad and Tobago. To obtain a representative

stratified sample, two strata were used: education district and school performance. In each district, schools were classified as high, medium, and low performing using the NLA school performance index. Sixty schools were selected for the quantitative study, with 58 participating. One school in Trinidad declined to participate and data from another school in Tobago was not collected within the required timeframe. From the 58 participating schools, 454 teachers responded to the survey. Field assistants were required to have completed a master's degree in Curriculum and were further trained in field methods and data collection. For the qualitative project, nine sites were selected based upon the level of implementation of the core curriculum components. Three sites were randomly selected from each of the high, medium, and low implementation categories (See Table 3).

Variables and measures

For quantitative project 1, a comprehensive survey instrument measuring 14 variables was developed and administered (see supplemental Table 1). These variables were (a) teacher individual characteristics, (b) use of key components, (c) use of curriculum materials, (d) use of subject areas, (e) levels of use, (f) use of integration, (g) curriculum orientations, (h) perceptions of the curriculum, (i) attitude towards the curriculum, (j) teacher professional learning, (k) measure of uncertainty, (l) organizational readiness, (m) perception of system training, and (n) curriculum leadership. These antecedent factors were organized using the IS framework of competency (g to k), organizational (l and m), and leadership (n) drivers.

We measured the degree of implementation using the self-reported frequency of (a) eight core activities, (b) nine subject areas taught, (c) six curriculum documents, and (d) type and level of curriculum integration approach. Fidelity, the extent to which the curriculum is implemented as planned, was measured both by (a) performance of core activities and (b) use of integration. For the latter, we calculated an overall integration score as well as the frequency of the specific type of integration employed (by discipline, theme, and student interest).

For the qualitative Project 2, the focus was on stakeholder perceptions of the antecedents, transactions, and outcomes of curriculum change. Qualified field assistants collected data over a period of 2–3 weeks. At each school, field assistants observed, documented practice, and conducted interviews with teachers and students. Supplemental Table 1 shows data from event logs, observations of classroom lessons, and classroom processes. Whereas the questionnaire survey measured only the structural dimension of fidelity, a process dimension measure was obtained in Project 2 by reviewing (a) a checklist of core activities, (b) artefacts related to integration, and (c) field notes and anecdotal records (Harn et al., 2017; O'Donnell, 2008).

The Mixed Methods Analysis Strategy

The MME purposes were triangulation and complementarity. Triangulation seeks convergence and corroboration, whereas complementarity captures elaboration and clarification (Greene et al., 1989). Elaboration and clarification were achieved primarily through Project 2. For these mixed methods purposes, both parallel and concurrent mixed analyses were used. Additionally, for complementarity, sequential approaches were employed (Onwuegbuzie et al., 2007). Variable oriented analyses were used in both Projects 1 and 2 (Onwuegbuzie et al., 2009).

In quantitative Project 1, the variables related to the degree of implementation and factors influencing change. We conducted descriptive (means and correlational analysis) and multivariate (MANOVA) statistical analyses. MANOVA evaluates differences across multiple continuous dependent variables simultaneously. The dependent variates consisted of (1) the outcome measures for use of materials, subjects, activities and (2) approaches to integration. Each independent variable was categorized using percentile scores, creating high, medium, and low groups. Each factor was examined for the two dependent variates. MANOVA explains how dependent variables vary across these groups. Wilks' Lambda is the key test statistic. We report both statistical and practical significance (Tabachnick & Fidell, 2019).

For Project 2, thematic and cross-case analyses¹ were the main strategies (Ayres et al., 2003; Herriott & Firestone, 1983). We conducted thematic analysis across all nine sites. There are several approaches to thematic analysis, including coding-focussed strategies and reflexive approaches (Byrne, 2022). Our preference was for an inductive process with the emphasis on generating latent codes, consistent with our understanding of critical realism. We followed the 3-step process described in Terry et al. (2017). This process involved familiarization and initial coding, development of preliminary

¹ Within-case analysis was also conducted, but is not reported in this paper.

themes, and reviewing and defining themes. We also used constant comparison as a subsidiary tool. Constant comparison involves 4 stages: (1) comparing incidents applicable to each category, (2) integrating categories and their properties, (3) delimiting the theory, and (4) writing the theory (Glaser, 1965).

We conducted thematic analyses on the text and documentary across the 9 sites. As explained by Herriott and Firestone (1983), cross-case themes from multisite studies will allow a greater level of generalizability and an explication of cause-and-effect relationships. Stake (2006) argued for using multisite cases to build theory that can inform programme improvement. To further enhance analytic generalization, we employed three strategies: (a) ideal typologizing, (b) category zooming, and (c) positioning (Halkier, 2011; Polit & Beck, 2010).

Ideal typologizing means building a description from a relatively limited number of descriptions that highlight particular characteristics at the expense of others. Category zooming involved focusing upon single aspects in the qualitative data. This allowed for an in-depth analysis, even with multiple cases with influential elements highlighted. Positioning and negotiation are used to draw inferences from the stories and discourses of participants to understand situational dynamics.

Integrated or meta-inferences were considered the key to generating robust and variegated theory (Teddlie & Tashakkori, 2009). To create meta-inferences from Projects 1 and 2 (integrated inferences), we utilised primarily complementary analyses (Bazeley, 2017). De Lisle (2013) described *linking, comparison, correlation, and merging* as possible integration approaches. Linking is illustrated in joint displays. Correlation means merging similar inferences from Projects 1 and 2. Comparison involves placing the inferences in parallel, even though they may be discordant. Merging or interweaving occurs when a single coherent statement is constructed from the inferences of Projects 1 and 2.

These integration approaches follow the qualitative principle of integrated (similar inferences) and dendritic crystallization (multiple,

dispersed inferences) (Ellingson, 2009). We classified meta-inferences as convergent, discrepant, or elaborative. Convergent meta-inferences combined similar findings, discrepant inferences merged contrasting findings, and elaborative findings offered a different or unique perspective from the findings of a single method (Younas et al., 2023). Discrepant and elaborative meta-inferences are consistent with an extended definition of triangulation, which acknowledges likely differences in qualitative and quantitative findings (Denzin, 2012).

Findings for the Quantitative Project

What is the current status of the primary school revision project? (MME)²

Quantitative Project-Descriptive Data

Table 2 shows the mean scores, standard deviations, and maximum scores for eight implementation outcomes (including a combined integration score) and the antecedent variables. The mean scores were very high for activities ($M = 4.74$; $SD=0.91$); subjects ($M = 4.50$; $SD = 0.57$); and materials ($M = 4.01$; $SD = 1.10$). Also notable is the low standard deviation for teaching different subjects and the high standard deviation for using curriculum materials. This suggests that not all individuals are using the full set of curriculum materials. The self-report scores for the three types of integration are also comparatively high; Multidisciplinary ($M = 3.62$), Interdisciplinary ($M = 3.92$), and Transdisciplinary ($M = 3.71$). On the surface, this provides support for the practice of different forms of integration.

The mean score on the Levels of Use³ (LoU) scale was surprisingly high at 5.88. This perhaps suggested that many respondents believed they had already progressed beyond a routine pattern of use and were now making changes to the curriculum to increase impact.

² This research question is also addressed by the thematic analysis.
³ The Level of Use Instrument is one of the tools used in CBAM (Hall et al., 2006). It is a self-report evaluation instrument asking teachers to rate Levels of Use in the innovation, including nonuse, orientation, preparation, mechanical use, routine use, refinement, integration, and renewal.

Table 2

Descriptive Statistics for Key Antecedent and Outcome Variables in Project 1

Measure	N	Min.	Max.	Mean	SD
<i>Curriculum implementation outcome measures</i>					
Activities	448	0.00	6.00	4.74	0.91
Subject	448	1.40	6.00	4.50	0.57
Materials	446	0.83	6.00	4.01	1.10
Multidisciplinary integration by subject	438	0.00	6.00	3.62	1.28
Interdisciplinary integration by theme	433	0.00	6.00	3.92	1.29
Transdisciplinary integration by student interest	425	0.00	6.00	3.71	1.48
Combined integration score	438	0.00	5.89	3.70	1.18
Levels of use	426	1.00	8.00	5.88	1.99
<i>Antecedent measures</i>					
<i>Organizational scores</i>					
Organizational leadership	430	0.63	5.00	3.06	1.03
Organizational readiness	428	1.00	5.00	3.08	0.93
Perception of system training	419	0.16	5.00	3.17	0.79
<i>Teacher curriculum orientation scores</i>					
Academic	448	0.67	5.00	3.76	0.73
Cognitive	448	0.67	5.00	3.86	0.72
Social reconstructionist	448	1.00	5.00	3.68	0.77
Technological	445	0.83	5.00	4.04	0.72
Humanistic	445	0.50	5.00	4.19	0.73
<i>Teacher attitude and perception scores</i>					
Teacher professional learning	443	2.50	5.00	4.29	0.57
Teacher certainty	443	0.20	5.00	3.42	0.83
Teacher perception of curriculum	433	1.27	5.00	3.03	0.53

Note. Most outcome measures were based on a scale of 1 (Never) to 6 (Every day).
 The Levels of Use was based on a scale from 1–8. Measures were based on a scale from 1–5.

Table 3

Joint Display and Classification of Sites in Project 2 Showing Ideal Typologizing and Positioning

Site	Typology of institutional use	Site description from ethnographic field notes/Project 1 activity & integration scores	Sample individual positioning and negotiation
-----High implementation sites-----			
A	True user	A small denominational school situated within a rural community. The physical structure consists of a two-storey building with seven classrooms, the audio-visual room, and administration. There is one class at each year level. Each class has approximately 12 students. Mean Activity Score=5.3 Thematic Integration Score=4.7	I took my personal cash and purchased an aquarium because that was our theme for them [the students] and later down at the end of the term... I saw the aquarium when we [were] finished with the theme and [we were] ready to move on. It came in late (Teacher 1). The children love some areas a bit. For example, the Spanish and so on and they excel at it. The dance/drama, they love it, they are very appreciative of it (Teacher 2).

Site	Typology of institutional use	Site description from ethnographic field notes/Project 1 activity & integration scores	Sample individual positioning and negotiation
B	True user	<p>A medium-sized denominational school of about 475 students and 21 educators. The site is located in a highly commercialized urban community. The site has three classes for each grade level. Specialist teachers exist for Physical Education and Library/Reading.</p> <p>This library arrangement facilitates a longstanding, functional lending system that greatly supports the school's reading programme.</p> <p>Mean Activity Score=5.3 Thematic Integration Score=3.3</p>	<p>Teacher collaboration has increased. It has increased significantly for this to work. I allow them time to meet on a weekly basis sometimes, but it is difficult sometimes to leave classes unattended (Principal).</p> <p>Our principal is very much determined to be the best at National Test. We focus a lot on Mathematics and Science. How can I teach PCR, but we must have standardized testing? Because at the end of the year is National Test and CAC business. He is only concerned about the results of my National Test (Teacher).</p>
C	Reluctant user	<p>A medium-sized government co-educational school situated in a suburban area. The school is a two-storey building housing the principal's office, the library, the computer lab, and staff room. These connecting structures house the 21 classrooms and the toilets. The staff comprises a female principal, 22 teachers – (1 male and 21 females). There are 487 students.</p> <p>Mean Activity Score=5.3 Thematic Integration Score=4.7</p>	<p>We worked as a team in 2nd Year just as we were trained. We would help each other out (Teacher 1).</p> <p>So, like you are fishing and by the time I was getting familiar with what I had to do, I was no longer in Infants but in Standard 2. So, now I am fishing my way through Standard 2 (Teacher 2).</p> <p>I remember when I was doing the PCR with Second Year, the parents had problems with the [large number] of projects. Everything was a project... (Teacher 3).</p>
-----Medium implementation site-----			
D	General nonuser	<p>A small denominational school located in an urban community. There is an acting female principal, 12 teachers (6 males and 6 females) including an Acting Head of Department, in addition to one Business Office Assistant (BOA) and two On-the-Job Trainees (OJTs). The two-storey structure air-conditioned building houses two classes at each level.</p> <p>Mean Activity Score=4.5 Thematic Integration Score=4.0</p>	<p>I asked the principal to put my class upstairs because I can't take that noise. In the beginning when we did this training and we went to learn about it, the theory of it looked perfect - the answer; but with all that we learnt, we forgot about the imperfect world we live in. (Teacher 1).</p> <p>I'm satisfied with the contents and that they give the children a little bit of everything. However, the infrastructure, the lack of resources, ICTs ... all of these wonderful things – they did not come with the curriculum. (Teacher 2).</p>
E	General nonuser	<p>A large government primary school located in a town but catering for a rural community. The school has 550 students and 26 educators, including the vice-principal and the principal.</p> <p>Mean Activity Score=4.7 Thematic Integration Score=3.3</p>	<p>We used to cover about 3 topics with just 1 lesson. We naturally do that because we were trained to teach like that, but this syllabus and the toolkit is confusing. And I tried my best to get the toolkit out... That is when I [went] back to be the teacher that I was before – when I go back to teaching my core. I cannot integrate everything. I am trying to do it and it is pulling me apart.</p>
F	General nonuser	<p>A medium sized Government school comprising of 478 students and 28 educators. This school serves a mixed, suburban community. The ageing physical plant is relatively large with self-contained classrooms. The site has a three-class structure at each level where the PCR has been implemented.</p> <p>Mean Activity Score=4.6 Thematic Integration Score=3.3</p>	<p>I had a major problem with the vagueness. You have to use both the books and some topics may not be in the scheme. So, teachers [are] not sure where they [are] supposed to go. I want them to just give me a scheme (Principal).</p> <p>They are not putting as much emphasis on it as we would have done in the past. For instance, in Reading, with the amount and the volume of work to be done I find we are not able find the time to do individual reading. Reading is suffering. (Teacher 1).</p>
-----Low implementation site-----			
G	Aversive nonuser	<p>A small denominational school situated in a changing suburban community. With a mixed population of boys and girls, along with an all-female staff. There is a female principal, 15 teachers, including an Acting Vice-Principal and Acting Head of Department. There are two classes at each level except for Standard One which has three classes. Each class has approximately 20-plus students with a total population of 283 students.</p> <p>Mean Activity Score=3.7 Thematic Integration Score=2.7</p>	<p>Working with the old syllabus, the old curriculum for so long now you have to come and change your whole mind-set at the beginning of a term, and they want things done in a certain way. (Teacher 1).</p> <p>Now if you have mild to moderate disabilities and also gifted children, then you have to differentiate every single lesson, every single day. I am not seeing it anywhere in the curriculum where we will cater for a child who has disabilities. (Teacher 2).</p>
H	Confused nonuser	<p>A small denominational school situated in a rural community. The physical structure consists of an aged two-story building which houses the principal's office, 13 classrooms, and the computer room. The staff comprises a female principal, 13 teachers and one voluntary administrative assistant. There are two classes at each level except second year. Each class has approximately 25 students.</p> <p>Mean Activity Score=4.1 Thematic Integration Score=3.6</p>	<p>They did not understand the curriculum; they did not understand what was required. It was just too much to actually sit and put into perspective at such a short time (Principal).</p> <p>There were inconsistencies in content and methodologies, what they might have been exposed to and what was required of them. In some cases, there was real confusion with even the execution, particularly how the children were to record and write in the books. There was one practice with the themes they had to explain to me exactly what was required (Principal).</p>
I	Aversive nonuser	<p>A small government school in a rural community serving children from marginalized communities. There is one class at each level. Most students are from disadvantaged families. The staff comprise a female principal, seven teachers – one male and six females – and one administrative assistant. There is one class at each level. Each class has approximately 25 students.</p> <p>Mean Activity Score=2.6 Thematic Integration Score=2.2</p>	<p>As simple as reading the documents, we are still having problems with how to understand the toolkit, how to mesh it with the actual syllabus in terms of subject area, how to pool it together. (Teacher 1).</p> <p>Daily plans are demanding, it takes a lot of time, and teachers cannot do it during the class time. (Teacher 2).</p>

Table 4
Inter-Correlation Matrix for Antecedent and Outcome Variables in the Study (n = 454)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
Implementation outcomes																			
1. Materials																			
2. Activities	.412**																		
3. Subjects	.417**	.435**																	
4. Integration by subject	.202**	.193**	.215**																
5. Integration by theme	.319**	.294**	.213**	.557**															
6. Integration by student interest	.299**	.244**	.177**	.418**	.631**														
7. Overall integration combined	.339**	.303**	.262**	.781**	.872**	.844**													
8. Levels of use	.134**	.182**	.034	-.018	.048	.100**	.052												
Curriculum orientation scores																			
9. CO (academic)	.190**	.063	.045	.072	.107*	.151**	.158**	.035											
10. CO (cognitive)	.165**	.101*	.083	.119*	.148**	.158**	.213**	.053	.733**										
11. CO (social recon.)	.149**	.111*	.143**	.123**	.146**	.189**	.214**	.097*	.628**	.669**									
12. CO (technological)	.075	.038	.099*	.037	.078	.08	.123*	.075	.686**	.672**	.567**								
13. CO (humanistic)	.141**	.150**	.125**	.094	.091	.135**	.152**	.058	.624**	.684**	.670**	.642**							
Attitude and perception scores																			
14. Teacher professional learning	.222**	.339**	.208**	.074	.218**	.215**	.204**	.210**	.282**	.291**	.321**	.297**	.353**						
15. Teacher certainty	.258**	.263**	.094*	.078	.187**	.127**	.128**	.074	-.02	.074	-.028	-.109*	.045	.066					
16. Perceptions of curriculum	.227**	.193**	.172**	.084	.188**	.196**	.192**	.064	.253**	.287**	.280**	.267**	.333**	.144**	.171**				
Organizational/system variables																			
17. Organizational leadership	.364**	.225**	.206**	.146**	.232**	.258**	.246**	.024	.196**	.200**	.109*	.102*	.100**	.161**	.176**	.270**			
18. Organizational readiness	.315**	.202**	.219**	.157**	.201**	.190**	.198**	.031	.194**	.206**	.117*	.155**	.160**	.165**	.226**	.316**	.811**		
19. System training perception	.239**	.115*	.122*	.122*	.212**	.182**	.208**	0.088	.197**	.258**	.139**	.158**	.219**	.139**	.255**	.525**	.367**	.407**	

Note. *p < .05. **p < .01. ***p < .001.

Table 5

MANOVA Results for Dependent Variate-Implementation Outcomes and Integration

Factor	Dependent variate-implementation outcome							Dependent variate-integration (fidelity)						
	MANOVA index				Wilks' Lambda			MANOVA index				Wilks' Lambda		
	Wilks'	Pillai's	Hotelling	Roy	Eta ²	NC	OP	Wilks'	Pillai's	Hotelling	Roy	Eta ²	NC	OP
Organizational Leadership	.846***	.154***	.181***	.177***	.080	70.233	1	.917***	.085***	.090***	.071***	.043	35.49	.998
Organizational Readiness	.901***	.101***	.109***	.094***	.051	43.913	1	.946***	.054***	.056***	.053***	.027	22.606	.964
Perception of System Training	.912***	.089***	.095***	.079***	.045	36.329	.997	.953**	.047**	.049**	.048***	.024	18.239	.912
Academic CO	.956*	.044*	.046*	.044***	.022	19.49	.904	.975	.025	.026	.016	.013	10.718	.676
Cognitive CO	.938***	.063***	.065***	.045***	.031	27.571	.981	.973	.027	.028	.024*	.014	11.505	.713
Social Reconstruction CO	.964*	.036*	.037*	.037**	.018	15.682	.816	.975	.025	.026	.025*	.013	10.821	.681
<i>Technological CO</i>	.974	.026	.027	.023*	.013	11.218	.644	.978	.022	.022	.020*	.011	9.248	.599
<i>Humanistic CO</i>	.973	.027	.028	.024	.014	10.534	.61	.979	.021	.021	.015	.011	8.088	.532
Professional Learning	.924***	.076***	.082***	.080***	.039	34.274	.996	.941***	.060***	.062***	.054***	.030	25.824	.982
Teacher Uncertainty	.924***	.077***	.082***	.079***	.073	33.537	.999	.976	.024	.024	.020*	.012	10.111	.646
Perception of the Curriculum	.947**	.053**	.056**	.054***	.027	23.143	.952	.952**	.048**	.050**	.047***	.024	20.404	.943

Note. Antecedents with significant associations on one or more dependent variable are highlighted.

CO = Curriculum Orientation

NC = non-centrality parameter; OP = observed power.

* $p < .05$. ** $p < .01$. *** $p < .001$. Eta Squared Benchmarks: small ($\eta^2 = 0.01$), medium ($\eta^2 = 0.06$), and large ($\eta^2 = 0.14$).

However, the qualitative data from Project 2 did not support these high self-report ratings. This divergence was especially evident with findings from the **ideal typologizing and positioning** analyses. Table 3 shows the ideal types for the implementing institutions. Two user and three nonuser variants were identified. User schools have most teachers at the stage of initial implementation, whereas nonuser schools have most teachers still exploring or installing the new curriculum or reverting to the old curriculum.

For the user variants, a distinction was made between true and reluctant user institutions. Reluctant user institutions were schools in which the majority of implementers demonstrated

a mechanical approach to adoption and early implementation, compared to a true user institution in which the process for most implementers was both engaging and adaptive. Among nonuser institutions, three identified variants were categorized as general nonusers, aversive nonusers, and confused nonusers. General nonusers were institutions with many teachers making little progress in installing the new curriculum. For confused nonusers, multiple factors at the site congealed on non-use, resulting in high levels of implementer uncertainty. For the aversive nonuser, negative attitudes among teachers predominated, with most implementers outwardly opposed to the new curriculum.

Table 3 also provides sample statements from individuals at the nine sites. The statements reveal how particular individuals positioned themselves and others in discourse on practice within the institution. Although these are individual responses, they are symptomatic of teachers in the user institution type. For example, at the two aversive nonuser sites, teachers spoke of continued difficulties and the requirement to change their worldviews of teaching and learning. At nonuser sites, respondents described their reactions and that of others to the curriculum, focusing on ambiguity and failed expectations. In the two true user sites and even in the reluctant user sites, the teachers defined their work collaboratively, despite curriculum imperfections and inadequacies of earlier training.

What factors explain the degree of implementation of the curriculum change? (MME)⁴

Quantitative Project-Intercorrelation Matrix

The pattern of associations between antecedent and outcome variables across all 58 schools was deciphered from the intercorrelation matrix and the MANOVA results. Table 4 shows the intercorrelation matrix for implementation outcomes, curriculum orientation scores, teacher attitude, perception scores, and organizational and system variables. As shown, most of the bivariate relationships between implementation outcomes are significant, but the sizes of the correlations are low to moderate.

The Pearson correlation coefficients for materials with activities and subjects were $r = .412, p < .01$ and $r = .417, p < .01$ respectively. The correlation between subject and activities was $r = .435, p < .01$. The bivariate correlations between curriculum orientation components and outcome variables were significant, but very low. The bivariate correlations between teacher attitude and perception were higher, especially for teacher professional learning and implementation activities ($r = .339, p < .01$). For the bivariate

relationships between organizational and system outcomes, the highest correlations were between leadership and use of materials ($r = .364, p < .01$), and organizational readiness and use of materials ($r = .315, p < .01$).

Quantitative Project-Inferential Data

Table 5 was a MANOVA table showing the relationship between antecedent factors and the dependent variate of implementation outcomes. The value of Wilks' Lambda indicates that nine of the 11 organizational and teacher variables were statistically significant. The factors without significant Wilks' Lambdas were technological and humanistic curriculum orientations. However, most effect sizes for the statistically significant factors were small. Of these, **organizational readiness** (Wilks' $\Lambda = .90, p < .001, \eta^2 = .05$) and **perception of system training** (Wilks' $\Lambda = .91, p < .001, \eta^2 = .05$) were the largest. Additionally, medium sized effect sizes were found for **organizational leadership** (Wilks' $\Lambda = .85, p < .001, \eta^2 = .08$) and **teacher uncertainty** (Wilks' $\Lambda = .92, p < .001, \eta^2 = .07$). For the integration variate, only five antecedents were significant. All effect sizes for the antecedents and variates were small, with the largest for **organizational leadership** (Wilks' $\Lambda = .92, p < .001, \eta^2 = .04$) and **teacher professional learning** (Wilks' $\Lambda = .94, p < .001, \eta^2 = .03$).

Findings for the Qualitative Project

1. ***What is the current status of the primary school revision project? (MME)***
2. ***What factors explain the degree of implementation of the curriculum change? (MME)***
3. ***What progress have teachers and leaders made in implementing school-wide curriculum change? (Qual)***

Thematic analysis of interview data, observations, and artefacts across the nine sites resulted in three themes:

⁴ This research question is also addressed by the thematic analysis.

1. Perceived curriculum imperfections folded onto subjects and thereby hindered adoption.
2. Lack of quality training severely limited competence and collective capacity.
3. Curriculum change led to the intensification of teachers' workload.

We report on each theme using in-vivo expressions constructed from participant interviews (Bazeley, 2013). Themes 1 and 3 were most common in low and medium implementation sites, with all six of such sites reporting these themes. Theme 2 was particularly reported in medium or high implementation sites, with all high implementation sites reporting the theme when describing their protocol for success. Theme 1 was reported at 8 of 9 sites. This theme focused upon perceived and real imperfections of the intended curriculum and the influence on curriculum change. Perceived weaknesses were folded onto teachers and interacted with their perceptions and concerns. This folding and unfolding process hindered adoption, with teachers in some schools hesitant to explore or install the new curriculum. In some cases, this hesitancy morphed into covert or open resistance. In other cases, there appeared to be philosophical conflicts between the intended approaches and the understanding of what was necessary in schools.

Theme 2 was the least frequently reported; observed or mentioned at only five sites. This theme focused on the inadequacy of training opportunities and the perceived lack of preparedness for enactment. Capacity and competence were common subthemes, but teachers also recognized the absence of site-based professional learning as critical to successful implementation. Theme 3 was reported at all 9 sites and focused on the intensification of teachers' work created by the demands of the new curriculum and the limits to practice created by the persistent absence of resources, inattention of the innovation to contextual peculiarities and student needs, and lack of organizational readiness and support. A frequent subtheme was the perceived incongruence between the physical

environment and the space requirements for enacting integrated, active learning within classrooms. In Theme 3, increased workloads and lack of resources were also repeated subthemes.

Theme 1: Perceived curriculum imperfections folded onto subjects and thereby hindered adoption...*so, it comes like we are guinea pigs?*

At two high implementation sites, the deficiency implied by Theme 1 was directly overcome through (a) teacher teams in site A or (b) teachers having a "just do it" attitude in site B. At site A, teacher collaboration was promoted by the school principal who established structures and time for joint planning. In such contexts, teachers still perceived curriculum imperfections as obstacles, but were often able to work around any deficiencies through collaborative planning and experimentation. However, at low and medium implementation sites, such collaborative structures and positive attitudes were not as readily apparent, and therefore the imperfections of the draft curriculum documents became constant and overwhelming sources of irritation, creating ambiguity, fear, and resistance.

Another notable issue was the ongoing tension between traditional teaching practices and the required emphasis on integration. The toolkit included details of integrated lessons, but was seen as fully aligned to the standard curriculum document with curriculum objectives. Interviews conducted with administrative staff at the Curriculum Planning and Development Division (CPDD) revealed that the idea of thematic integration as a valuable innovation had been developed over time. However, during development, guidance by the consultants was often vague and curriculum writers had been given too much latitude by curriculum officers to explore this feature. Indeed, as one curriculum officer said on reflection, "we all need to question where the decision to make it a thematic integrated approach came from." Likewise, from the perspective of individual users, the intention to prioritize integration was by no means clear, as evident in the view of one teacher at Site G:

The syllabus [materials] are confusing. I guess it's the methods that we are accustomed using. This throws a different spin on everything and the themes [they chose] are not related to everything on the syllabus. It's just a few things they actually picked out and they make their themes around that. [There are] so many standalone lessons still and they are fretting so much with this thematic [approach].

There was a tendency, especially in the low and medium implementation schools, for teachers to quickly abandon installation and return to traditional teaching practices focused upon individual subjects and on reading. Some teachers blamed this quick reversal on the perceived imperfections of the curriculum. As one teacher at Site E admitted:

Even with the ones that started, there is a tendency to revert to the old way, even with the use of the copy books. They were suggesting that you use one book per theme, but that was confusing to the children. So, what have teachers done? They have gone back to separate copy books.

To some extent, the folding and unfolding of these perceived curriculum inadequacies onto users occurred because of the absence of a field try-out. As one of the developers noted, "When our government changed, and the former Minister of Education got into the mix in November 2011, he changed the implementation plan and instructed us to go national immediately." Nevertheless, as a teacher at Site D observed, "No company that is bringing out a new product will make 10,000 copies of a product if they're not sure."

Therefore, when teachers identified imperfections, they were actually being critical of that circumvented process. However, other teachers felt that curriculum documents were inherently unclear and frequently ambiguous in purpose. This perceived lack of clarity was enhanced further by the large number of released documents, as observed by one teacher from Site B:

[This] curriculum is overloaded. I have a problem with the language [of the documents]. If you are [going] to standardize the curriculum, it would be easier for the teachers to spend more time making the lesson supportive of the

child's level. That means that the lesson must be child-centred. So, the time frame is the problem. When you have this big bulk of curriculum, you're not catering for [the] reality.

Several teachers expressed concern with the lack of alignment between student assessment and the curriculum. Most of these concerns referred to the national learning assessments, but some teachers also considered current classroom assessment practice. The assessment aspect, both classroom and large-scale, was mostly left in abeyance by the curriculum documents. Consequently, participants at different sites were quick to observe that the content taught was not being assessed. As one principal (Site G) noted:

What is more frustrating for them (and a big concern for me), is that the new curriculum is not assessed at [the] National Test [level]. Here you are teaching a new curriculum and [then] you have a standardized test at the end of the year where you are going to be judged as a school on your students' overall performance at something that is not aligned.

Theme 2: Lack of quality training severely limited competence and collective capacity

"I have no training in using the PCR."

From the perspective of CPDD staff, training for the PCR was extensive compared with past efforts. However, the workshops were centralized rather than site-based. This exhaustive nationwide training exercise had greatly expended the energies of administrative staff. Nevertheless, both divisional staff and teachers described the training as inadequate and rushed. One officer of the CPDD described the situation as "...horrific. The Minister wanted curriculum implementation by September, so we had to figure out a way to train teachers by September."

Teachers were not enthusiastic about the quality of centralized training; describing it as too short, somewhat hurried, too theoretical, and mostly didactic. For example, one teacher from School B observed, "It was a very short course – just 4 days – and I found that it wasn't

elaborate enough. I found myself to be a little bit lost." She continued, "This is my second year in First Year, and I am all over the place. I would suggest a month or two weeks [of the] workshop again."

This perception of inadequate training may have directly influenced teachers' self-efficacy and competence. As one teacher at Site B admitted, "My problem is that I don't feel adequate enough. I have no training in using the PCR. I find myself more or less scrambling. I am not the only one in this position." Notably, the teacher referred both to her own position and to that of others in the school. In many cases, new issues arose as implementation proceeded. One teacher from School B noted that, "in the beginning when we did this training and we went to learn about it, the theory of it looked perfect, but with all that we learnt, we forgot about the imperfect world we live in." Thus, teachers felt that the training offered was insufficient for installing the curriculum in the poorly resourced classrooms of local schools.

Most schools did not have formal collaborative structures such as teams to facilitate on-site curriculum planning and implementation. Several variables, then, may have interacted to generate and magnify the perceived lack of competence. The most notable was a strong negative attitude towards the curriculum, which was evident at most of the low and medium implementation sites. Although most teachers were trained, they still struggled with the new content to be taught. Sometimes, contextual issues further exacerbated the training-competence issue. For example, at School C, a high achieving urban school, a policy of teacher rotation limited the application of what was taught in the workshops because teachers had been trained at a specific level. Further, principals themselves had received only a basic orientation (rather than extensive training), which limited their capacity for instructional leadership.

Theme 3: Curriculum change led to the intensification of teachers' workload

"Daily plans are demanding, it takes a lot of time, and teachers cannot do it during the class time."

Planners of the new curriculum had not anticipated work intensification as an unintended outcome, and therefore few schools had established structures or routines to reduce the workload. This omission was impractical, considering the volume of curriculum materials. Indeed, some teachers, especially those in low performing and high poverty schools, felt "bombarded" with information. Additionally, there were persistent inadequacies of physical space and resources at most sites (imperfections of context). This limited teacher experimentation and use of the curriculum.

The lack of physical space and resources also made it difficult to teach specific subjects (such as Social Studies and Art) in an authentic and integrated fashion. These inadequacies further intensified teachers' work by demanding creative solutions for what should be routine tasks. In some cases (for example in School G), the situation was one of accessibility and management of resources. Physical space was an issue in many schools. Some buildings were poorly maintained or in need of repair. For example, in School I, teachers complained bitterly of overcrowded classrooms, in addition to the lack of resources in key areas. In this school, only one teacher continued onto installation. She said:

I like the curriculum. I see how it can help my children and it does cater for all students. You just have to take the time to plan. I spend a lot of my time planning. I have the weakest class and the most students, but I try to do a lot of remedial reading and mathematics during my lunch time with them, and it is helping to prepare the children for the National Test.

The majority of teachers at this and other low and medium implementation sites, however, believed that the new curriculum did not fully mesh with the heterogeneous realities of schools in the local contexts. More specifically, they believed that the curriculum did not cater for the learning needs of all students, especially struggling readers and students with special

needs. Teachers observed that the curriculum was mostly silent on these matters, although one of the stated tenets of the PCR was “differentiated learning.” One teacher at Site H gave this deficiency as a reason for abandoning further exploration at an early stage. She said:

To be honest, I have gone back to the old way. Look at this. When you have children who can't write. I have to do what is top priority. I need to teach them the skills to be ready for second year, so I really don't teach that way.

Cross-case Variable Analysis and Category Zooming

We used cross-case analysis to identify positive and negative antecedent variables (see Supplemental Table 2). At the high implementation sites, positive teacher attitudes and collaboration appeared to be important for successful implementation. It is likely that the relationship between these two variables was recursive, with collaboration leading to positive attitudes and these positive attitudes in turn fuelling collaboration.

After conducting the cross-case analysis, we used category zooming to identify three variables: *collaboration*, *resources*, and *training* (see Supplemental Table 3). These variables relate to Themes 2 and 3, which were identified at the majority of sites. Collaboration, whether formal or informal, was noted only at the high implementation sites. The lack of resources, physical space, and time were issues at most of the medium and high implementation sites. Concerns over alignment were present at one high implementation site, compared to two of the three low implementation sites. Inadequate training and lack of support by the system and school were issues at eight of the nine sites. It is possible that high quality collaboration compensated for inadequate training at high implementation sites.

Generation of Integrated Inferences

Ten meta-inferences were generated through comparison, correlation, and merging of findings from the two projects. The integrated findings are presented in Table 6. Five meta-inferences

were convergent, with similar conclusions from both projects. Three were elaborative, with the qualitative project adding additional insight and perspective to findings from Project 1. Two were discordant, with contradictory findings from both projects. The convergent meta-inferences showed that teachers had expanded the repertoire of subjects taught, although in some cases the constraints of resources and space limited teaching of Visual and Performing Arts (VAPA) and Information Technology (IT). Data from both Projects 1 and 2 confirmed widespread negative attitudes towards the curriculum, lack of understanding, and an absence of commitment; such that the folding of the curriculum onto self was often incomplete. Additionally, there was little pressure from colleagues or parents to install the curriculum, and training systems proved universally weak.

The qualitative project revealed that some sites lacked readiness for change and pointed to teachers quickly reverting to prior traditional practices. Reported levels of using curriculum materials appeared low. The interview data suggested that this was because of perceived imperfections and the lack of high-quality support at the sites. Although the quantitative findings suggested that outcomes were moderate to high, implementation at the majority of sites often appeared to be in the very early stages, with little progress or reversals at some sites. The lack of clarity on integration type was evident in the variety of approaches used in the self-report measure. However, observation of practices at school sites suggested that approaches to integration were limited and often lacked fidelity. Field assistants often questioned the accuracy of teachers' understanding of the processes involved in integrating subject areas.

Discussion and Conclusion

Utility of Evaluation Evidence

This paper explores evidence from a formative process evaluation of a newly installed curriculum in Trinidad and Tobago, a small island developing state in the Caribbean. The responsive

Table 6*Meta-Inferences, Types, Theoretical Perspectives, and Implementation Strategy Recommended*

Meta-inference on implementation status	Type	Theoretical perspective	Implementation strategy recommended
1. Teachers were using the nine subject areas, although the frequency of practice for some subject areas was low.	Convergent	IS	Design and implement a monitoring & evaluation system
2. Many teachers had a neutral or negative view of the curriculum, which may impact on sustainability. The multiple case study suggested that at high implementation sites, many teachers demonstrated a "just do it" attitude and were willing to problem-solve issues. However, negative teacher attitudes were overwhelming at low or medium implementation sites.	Convergent	TRC	Develop an enhanced professional learning strategy
3. Overall, teachers were not strongly committed to the new curriculum, and some appeared not to fully understand the theory of action.	Convergent	DoI	Develop & implement a comprehensive marketing strategy
4. There were insufficient pressures and benefits in the social system to implement the new curriculum.	Convergent	DoI	Develop & implement a comprehensive marketing strategy
5. Implementation drivers related to the perception of uncertainty were important, but training systems were considered insufficient or weak. In the qualitative multiple case study, there was universal agreement that training systems were insufficient.	Convergent	IS	Employ multi-functional implementation teams at all levels
6. Some schools were not ready for change, and this will be a barrier to vertical scale-up. The multiple case study suggested that the physical environment and resources could be both a <i>facilitator</i> and a barrier to successful initial implementation.	Elaborative	IS	Promote leadership of curriculum change throughout the system
7. Use of the curriculum materials was much lower and more variable than the activity outcome measure. Findings from the multiple case study suggested that teachers at some sites believed the materials to be difficult to use. However, these issues were reduced with high quality support, such as instructional coaching.	Elaborative	IS	Employ multi-functional implementation teams at all levels
8. It is possible that some teachers might have begun experimenting with alternative approaches, perhaps due to less than positive perceptions of the innovation. Findings from the multiple case study also suggested that it was more likely that these teachers had reverted to more traditional practice.	Elaborative	IS/CBAM	Develop an enhanced professional learning strategy
9. Self-reported implementation rates as measured by the eight core activities were relatively high; but on-site observation suggests that practice was not high quality or high intensity, except at higher implementation sites.	Discrepant	IS	Employ multi-functional implementation teams at all levels
10. The quantitative study suggested that integration was practiced, most commonly by themes (interdisciplinary), but also in some cases with the child as the centre of integration (transdisciplinary). However, observations in the multiple case study suggested that high quality integration was rare, except at two sites.	Discrepant	IS	Design and implement a monitoring & evaluation system Employ multi-functional implementation teams at all levels

Note. IS = implementation science; TRC = teacher receptivity to change; DoI = diffusion of innovation; CBAM = concerns-based adoption movement.

evaluation model and a qualitatively-driven MME design provided insight into the complex nature of implementation and the factors influencing adoption in this country context. Quantitative data was gathered from a representative

sample of 58 schools and qualitative data from a multiple case study of nine sites. Evidence from quantitative Project 1 suggested that implementation outcomes were mostly adequate and moderately associated with uncertainty

and school leadership. Evidence of recurring processes and generative mechanisms in the qualitative Project 2 generally supported the findings from the statistical associations in the multiple case study.

The integrated inferences offered further insight into possible mechanisms and recurring processes of implementation failure within these contexts. Perceptions of system training and organizational readiness were identified as the critical variables in the multiple case study. Teachers at different sites often questioned the imperfections, fit, and relevance of the curriculum. The poor quality of system-level training also led to higher levels of uncertainty. At the high implementation sites, however, leaders were able to foster conditions that facilitated teacher learning during early implementation. Most commonly, these were structures and processes related to collaboration.

The folding metaphor of Webb and Gulson (2013) was a valuable heuristic for understanding curriculum change in this under-resourced context. Self was central, so that folding and unfolding were imbued with emotions, identities, and positionalities. The convergent meta-inferences suggested that feelings about the curriculum were part of the enactment process and were critical in either rejecting or facilitating change. Teachers often adopted strong emotional positions about the curriculum, and this influenced their intentions and actions. Such positioning determined whether they continued to implement or returned to their original practices. Negative attitudes towards curriculum change do not bode well for the sustainability of the process.

The folding metaphor could also be used to explain the unintended consequences of the training schedule. This schedule focused upon initial, didactic training and excluded work-based training at the site. This approach could not foster sufficient insight into the rationale for the change or the practices needed for enactment. The initial training also failed to foster fidelity of practice. Although teachers reported practicing some forms of integration; the observations, artefacts, and interviews in the multiple case study suggested

that integration practice was inconsistent or of poor quality. In this study, structural measures of fidelity were limited as sources of evidence, but process measures provided a better judgment on the trueness of curriculum change (Mowbray et al., 2003).

The folding metaphor highlighted the complexity of change in this context. Like the metaphor of quilt-making, folding emphasizes the complex and personal nature of curriculum change (Hizli Alkan, 2022). Variables that constitute barriers and facilitators to change were best understood as emerging and interacting. Although similar barriers were also present at low and medium implementation sites, at high implementation sites, leadership and collaboration could be used to overcome resistance to change. For work intensification, lack of resources was intertwined with space, both interacting to limit implementation. This network of interacting and emerging barriers included variables such as poor-quality system training, curriculum imperfections, and the absence of site-based professional learning.

Although these findings on integration might not be new for the Caribbean (Jennings, 1994), they do point to complexity as a core issue. Such unpredictable, emerging, and contingent factors during implementation demand better monitoring systems to respond to challenges. The findings also suggest that the imperfections folded in came from the intended curriculum as well as other system aspects such as inadequate training, institutional readiness, resources, and unexpected demands. As the curriculum folded (unfolded and refolded) on subjects and actors, these imperfections interacted, accumulated, and became magnified. They were not easily forgotten and became persistent irritants, amplifying the negative emotional response to curriculum change.

Implications for Curriculum Change in the Caribbean

Gift (1987) identified several deficiencies in the curriculum engineering system of Trinidad and Tobago; including responsiveness to societal issues, dissemination strategies, evaluation

processes, and limited involvement of and ownership by teachers. In this study, the lack of pilot testing, monitoring, and evaluation in the curriculum engineering system appeared to magnify the curriculum imperfections and altered the perceptions of actors and subjects in sensemaking. Reaction to curriculum imperfections often unfolded during enactment. Evidence from this study highlights the need, then, for installing a comprehensive and robust curriculum engineering system, with improved data-driven planning and monitoring and evaluation frameworks guided by IS theory. Such a system will ensure credible information on barriers and facilitators, instead of adopting a cookbook listing of methods and approaches.

For the evaluation function of curriculum engineering systems, there is value in generating data using a qualitatively-driven MME to inform curriculum change. From a critical realist perspective, intensive and extensive approaches provide both complementary and contradictory evidence, which, when integrated, can generate insight. Contradictory or divergent evidence will occur in a responsive evaluation where different perspectives from a variety of stakeholders are actively sought. It might have been useful in this study to also gather data from parents and community members to understand the social pressures leading to adoption. The evidence from teachers in this study suggests that implementers did not perceive such pressures, pointing to the need to better market the PCR.

Key levers of change (as specified in the IS model) were neglected in the original curriculum implementation plan. The lack of attention to levers of change such as principal leadership and organizational support was surprising. Indeed, most of the principals were unprepared because of their lack of training, and therefore remained at the periphery of the change process. This finding shows the value of using IS as a rubric for planning change. Training of principals should not only focus on curriculum content and instructional practice, but also on the key processes of instructional and distributive leadership (Newton, 1993).

In the IS model, team-based collaborative planning is encouraged at multiple levels. Collaborative teams are useful for flexible planning, and to generate monitoring and evaluation data needed for a quick response to evolving factors in these complex contexts. One essential inclusion of an evidence-informed implementation plan is site-based learning using experimentation. Plans might also emphasize training strategies such as coaching and mentoring fitted into an overall collaborative structure. Evidence informed by IS theory will help build the stronger curriculum engineering systems required to reduce the implementation gap.

This work pays tribute to the founders of curriculum and curriculum change and evaluation in the Caribbean. These include Edric Gift from Trinidad and Tobago, Earle Newton from Barbados, and Zellynne Jennings from Guyana and Jamaica. Edric Gift worked in the field of curriculum; Zellynne Jennings in teacher education, curriculum and reading; and Earle Newton in educational administration and evaluation. We can learn much from their published ideas as we build collective capacity and work to install robust curriculum engineering systems. Insight into the implications of integration for teachers and teaching in a postcolonial Caribbean setting will help us improve implementation fidelity. Current training systems must build collective capacity for modern day curriculum designs and strategies such as backward design, co-curricular themes, integration, new pedagogies, and formative classroom assessment. Innovative evaluation designs in mixed methods are now available to provide added insight into curriculum implementation (Palinkas et al., 2019).

Acknowledgements

Tracey Lucas, Gwendolyn Ahyoung, and Victoria Cunningham were research assistants on this project.

Jerome De Lisle, PhD is Professor of Educational Leadership at the School of Education at The University of the West Indies, St. Augustine. He was technical lead and evaluation specialist for the University of the West Indies project team evaluating the Primary School Rewrite for the Government of the Republic of Trinidad & Tobago. He is a formally trained evaluator and a member of the American Evaluation Association (AEA).

Email: jerome.de_lisle@sta.uwi.edu

Sharmila Harry, PhD is a Lecturer in Curriculum Studies and Teacher Education at The University of the West Indies, St. Augustine. Her PhD thesis was on curriculum implementation. She was a Curriculum Consultant for The University of the West Indies project team evaluating the Primary School Rewrite for the Government of the Republic of Trinidad and Tobago.

Email: sharmila.harry@sta.uwi.edu

References

- Abma, T. A., & Stake, R. E. (2001). Stake's responsive evaluation: Core ideas and evolution. In J. Green & T. A. Abma (Eds.), *Responsive evaluation: New Directions for Evaluation*, 92, (pp. 7–23). <https://doi.org/10.1002/ev.31>
- Alhadeff-Jones, M. (2008). Three generations of complexity theories: Nuances and ambiguities. *Educational Philosophy and Theory*, 40(1), 66–82. <https://doi.org/10.1111/j.1469-5812.2007.00411.x>
- Altinyelken, H. K. (2010). Curriculum change in Uganda: Teacher perspectives on the new thematic curriculum. *International Journal of Educational Development*, 30(2), 151–161. <https://doi.org/10.1016/j.ijedudev.2009.03.004>
- Ayres, L., Kavanaugh, K., & Knafl, K. A. (2003). Within-case and across-case approaches to qualitative data analysis. *Qualitative Health Research*, 13(6), 871–883. <https://doi.org/10.1177/1049732303013006008>
- Ball, S., Maguire, M., & Braun, A. (2012). *How schools do policy: Policy enactments in secondary schools*. Routledge.
- Bantwini, B. D. (2010). How teachers perceive the new curriculum reform: Lessons from a school district in the Eastern Cape Province, South Africa. *International Journal of Educational Development*, 30(1), 83–90. <https://doi.org/10.1016/j.ijedudev.2009.06.002>
- Bantwini, B. D. (2012). Primary school science teachers' perspectives regarding their professional development: Implications for school districts in South Africa. *Professional Development in Education*, 38(4), 517–532. <https://doi.org/10.1080/19415257.2011.637224>
- Bantwini, B. D., & Letseka, M. (2016). South African teachers caught between nation building and global demands: Is there a way out/forward? *Educational Studies*, 52(4), 329–345. <https://doi.org/10.1080/00131946.2016.1190366>
- Barrow, D., & De Lisle, J. (2009). A qualitative evaluation of the lower secondary SEMP Science curriculum of Trinidad and Tobago. *Caribbean Curriculum*, 16(2), 73–97.
- Bazeley, P. (2013). *Qualitative data analysis. Practical strategies*. Sage.
- Bazeley, P. (2017). *Integrating analyses in mixed methods research*. Sage.
- Beauchamp, G. A. (1972). Basic components of a curriculum theory. *Curriculum Theory Network*, 3(10), 16–22. <https://doi.org/10.2307/1179213>
- Bellino, M. J. (2016). So that we do not fall again: History education and citizenship in "postwar" Guatemala. *Comparative Education Review*, 60(1), 58–79. <https://doi.org/10.1086/684361>
- Bertram, R. M., Blase, K. A., & Fixsen, D. L. (2015). Improving programs and outcomes: Implementation frameworks and organization change. *Research on Social Work Practice*, 25(4), 477–487. <https://doi.org/10.1177/1049731514537687>

- Bos, M. S., Elías, A., Vegas, E., & Zoido, P. (2016). Latin America and the Caribbean in PISA 2015: How Much Did the Region Improve? IDB. Retrieved from <https://publications.iadb.org/bitstream/handle/11319/7992/Latin-America-and-the-Caribbean-in-PISA-2015-How-Much-Did-the-Region-Improve.PDF?sequence=5&isAllowed=y>
- Braithwaite, J., Churrua, K., Long, J. C., Ellis, L. A., & Herkes, J. (2018). When complexity science meets implementation science: a theoretical and empirical analysis of systems change. *BMC Medicine*, *16*, 1–14.
- Brand, S. L., Quinn, C., Pearson, M., Lennox, C., Owens, C., Kirkpatrick, T., Callaghan, L., Stirzaker, A., Michie, S., Maguire, M., Shaw, J., & Byng, R. (2019). Building programme theory to develop more adaptable and scalable complex interventions: realist formative process evaluation prior to full trial. *Evaluation*, *25*(2), 149–170.
- Brough, C. J. (2012). Implementing the democratic principles and practices of student-centred curriculum integration in primary schools. *Curriculum Journal*, *23*(3), 345–369. <https://doi.org/10.1080/09585176.2012.703498>
- Byrne, D. (2022). A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Quality & Quantity*, *56*(3), 1391–1412. <https://doi.org/10.1007/s11135-021-01182-y>
- Century, J., & Cassata, A. (2016). Implementation research: Finding common ground on what, how, why, where, and who. *Review of Research in Education*, *40*(1), 169–215. <https://doi.org/10.3102/0091732X16665332>
- Cheek, J., & Morse, J. M. (2022). The power of qualitative research in mixed methods research designs. In U. Flick (Ed.), *The SAGE Handbook of Qualitative Research Design* (pp. 636–51). SAGE.
- Choi, T. H., & Walker, A. D. (2018). A heuristic model for tailoring teacher development to educational reforms: Focusing on ambiguity and conflict generation. *Teaching and Teacher Education*, *74*, 72–84. <https://doi.org/10.1016/j.tate.2018.04.013>
- Conley, T. (2005). *Folds and folding*. In C. Stivale (Ed.), *Gilles Deleuze: Key concepts* (pp. 170–181). McGill-Queen's University Press. <https://doi.org/10.1515/9780773584884-018>
- Cook, V. A., & Hemming, P. J. (2011). *Education spaces: Embodied dimensions and dynamics*. *Social & Cultural Geography*, *12*(1), 1–8. <https://doi.org/10.1080/14649365.2011.542483>
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research* (3rd ed.). SAGE.
- Deleuze, G. (1993). *The fold: Leibniz and the baroque* (original work published 1988). Foreword and translation by T. Conley. University of Minnesota Press.
- Deleuze, G., & Strauss, J. (1991). The fold. *Yale French Studies*, *80*, 227–247. <https://doi.org/10.2307/2930269>
- De Lisle, J. (2011). The benefits and challenges of mixing methods and methodologies: Lessons learnt from implementing qualitatively led mixed methods research designs in Trinidad and Tobago. *Caribbean Curriculum*, *18*, 87–120. <https://journals.sta.uwi.edu/ojs/index.php/cc/article/view/580>
- De Lisle, J. (2013). Exploring the value of integrated findings in a multiphase mixed methods evaluation of the continuous assessment program in the Republic of Trinidad and Tobago. *International Journal of Multiple Research Approaches*, *7*(1), 27–49. <https://doi.org/10.5172/mra.2013.7.1.27>

- De Lisle, J., Smith, P., & Jules, V. (2010). Evaluating the geography of gendered achievement using large-scale assessment data from the primary school system of the Republic of Trinidad and Tobago. *International Journal of Educational Development, 30*(4), 405–417. <https://doi.org/10.1016/j.ijedudev.2010.01.001>
- Denzin, N. K. (2012). Triangulation 2.0. *Journal of Mixed Methods Research, 6*(2), 80–88. <https://doi.org/10.1177/1558689812437186>
- Doll, W. E. (2008). Complexity and the culture of curriculum. *Educational Philosophy and Theory, 40*(1), 190–212. <https://doi.org/10.1111/j.1469-5812.2007.00404.x>
- Drake, S. M., & Burns, R. (2004). *Meeting standards through integrated curriculum*. Association for Supervision and Curriculum Development.
- Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*(3–4), 327–350. <https://doi.org/10.1007/s10464-008-9165-0>
- Dyer, C. (1999). Researching the implementation of educational policy: A backward mapping approach. *Comparative Education, 35*(1), 45–61. <https://doi.org/10.1080/03050069928062>
- Eastwood, J. G., Kemp, L. A., & Jalaludin, B. B. (2016). Realist theory construction for a mixed method multilevel study of neighbourhood context and postnatal depression. *SpringerPlus, 5*, 1–26.
- Ellingson, L. L. (2009). *Engaging crystallization in qualitative research: An introduction*. Sage.
- Erskine, B. (2013, January 2). Education group wins curriculum contract. *Chronicle Herald online [Halifax]*. <http://thechronicleherald.ca/business/360319-education-group-wins-curriculum-contract>
- Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on Social Work Practice, 19*(5), 531–540. <https://doi.org/10.1177/1049731509335549>
- Freeman, M., & Hall, J. N. (2012). The complexity of practice participant observation and values engagement in a responsive evaluation of a professional development school partnership. *American Journal of Evaluation, 33*(4), 483–495. <https://doi.org/10.1177/1098214012443728>
- Fullan, M. (1983). Evaluating program implementation: What can be learned from follow through. *Curriculum Inquiry, 13*(2), 215–227. <https://doi.org/10.1080/03626784.1983.11075881>
- Gerrits, L., & Verweij, S. (2013). Critical realism as a meta-framework for understanding the relationships between complexity and qualitative comparative analysis. *Journal of Critical Realism, 12*(2), 166–182. <https://doi.org/10.1179/rea.12.2.p663527490513071>
- Gibson, C. B. (2017). Elaboration, generalization, triangulation, and interpretation: On enhancing the value of mixed method research. *Organizational Research Methods, 20*(2), 193–223. <https://doi.org/10.1177/1094428116639133>
- Gift, E. H. (1984, August). *Impediments to curriculum engineering for change: The case of Trinidad and Tobago*. Paper presented at the Multicultural and Interdisciplinary Development Education Conference, Mona, Jamaica.
- Gift, E. H. (1987). A systems model for curriculum engineering in a third world county-Trinidad and Tobago. *Perspectives in Education, 3*(4), 193–214.
- Glaser, B. G. (1965). The constant comparative method of qualitative analysis. *Social Problems, 12*(4), 436–445.

- Glatthorn, A. A., Boschee, F., Whitehead, B. M., & Boschee, B. F. (2018). *Curriculum leadership: Strategies for development and implementation*. SAGE Publications.
- Gouëdard, P., Pont, B., Hyttinen, S., & Huang, P. (2020). Curriculum reform: A literature review to support effective implementation. *OECD Education Working Papers No. 239*. Paris: OECD.
- Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11*(3), 255–274. <https://doi.org/10.3102/01623737011003255>
- Grundy, S. (2002). Big change questions—Is large-scale educational reform possible? *Journal of Educational Change, 3*(1), 55. <https://doi.org/10.1023/A:1016539814319>
- Halkier, B. (2011). Methodological practicalities in analytical generalization. *Qualitative Inquiry, 17*(9), 787–797. <https://doi.org/10.1177/1077800411423194>
- Hall, G. E., Dirksen, D. J., & George, A. A. (2006). *Measuring implementation in schools: Levels of use*. SEDL.
- Harachi, T. W., Abbott, R. D., Catalano, R. F., Haggerty, K. P., & Fleming, C. B. (1999). Opening the black box: using process evaluation measures to assess implementation and theory building. *American Journal of Community Psychology, 27*(5), 711–731. <https://doi.org/10.1023/A:1022194005511>
- Harn, B. A., Parisi Damico, D., & Stoolmiller, M. (2017). Examining the variation of fidelity across an intervention: Implications for measuring and evaluating student learning. *Preventing School Failure: Alternative Education for Children and Youth, 61*(4), 289–302. <https://doi.org/10.1080/1045988X.2016.1275504>
- Harré, R., Moghaddam, F. M., Cairnie, T. P., Rothbart, D., & Sabat, S. R. (2009). Recent advances in positioning theory. *Theory & Psychology, 19*(1), 5–31. <https://doi.org/10.1177/0959354308101417>
- Harry, S. (2007). Teachers' concerns about the CAPE Communication Studies innovation. In L. Quamina-Aiyejina (Ed.). *Reconceptualising the agenda for education in the Caribbean: Proceedings of the 2007 Biennial Cross-Campus Conference in Education, pp.203–215* UWI, St. Augustine.
- Harry, S. N. (2019). *The implementation of the Caribbean Advanced Proficiency Examination (CAPE) Communication Studies innovation in secondary schools in Trinidad and Tobago: Teachers' perspectives*. Doctoral dissertation, University of Leicester.
- Herriott, R. E., & Firestone, W. A. (1983). Multisite qualitative policy research: Optimizing description and generalizability. *Educational Researcher, 12*(2), 14–19. <https://doi.org/10.3102/0013189X012002014>
- Hesse-Biber, S. (2022). Mixed methods research designs in qualitatively driven research. In U. Flick (Ed.), *The Sage Handbook of Qualitative Research Design* (pp. 618–635). SAGE.
- Hickling-Hudson, A. (2010). Curriculum in postcolonial contexts. In P. Peterson, E. Baker, & B. McGraw (Eds.), *International Encyclopaedia of Education* (3rd ed). (pp. 299–306). Elsevier.
- Hizli Alkan, S. (2022). "Are we making a quilt, with lots of ill-fitting cloths in here?" Teachers' internal conversations on curriculum making. *Journal of Educational Change, 24*(2), 457–478. <https://doi.org/10.1007/s10833-022-09452-8>

- Hurrell, S. A. (2014). Critical realism and mixed methods research: combining the extensive and intensive at multiple levels. In P. K. Edwards, J. O'Mahoney, & S. Vincent (Eds.), *Studying organizations using critical realism: A practical guide* (pp. 241–263). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199665525.003.0013>
- Huybrechts, I., Declercq, A., Verté, E., Raeymaeckers, P., & Anthierens, S. (2024). How does the external context affect an implementation processes? A qualitative study investigating the impact of macro-level variables on the implementation of goal-oriented primary care. *Implementation Science, 19*(1), 32.
- Jansen, J. D. (1998). Curriculum reform in South Africa: A critical analysis of outcomes-based education. *Cambridge Journal of Education, 28*(3), 321–331. <https://doi.org/10.1080/0305764980280305>
- Jennings, Z. (1993). Curriculum change in school systems in the Commonwealth Caribbean: Some implications for the management of curriculum development. *International Journal of Educational Development, 13*(2), 131–143.
- Jennings, Z. (1994). Innovations in Caribbean school systems: Why some have become institutionalised and others have not? *Curriculum Studies, 2*(3) 309–331. <https://doi.org/10.1080/0965975940020303>
- Jennings-Wray, Z. D. (1984). Implementing the 'integrated approach to learning': Implications for integration in the curricula of primary schools in the Caribbean. *International Journal of Educational Development, 4*(4), 265–278. [https://doi.org/10.1016/0738-0593\(84\)90012-9](https://doi.org/10.1016/0738-0593(84)90012-9)
- Jennings, Z. (2001). Teacher education in selected countries in the Commonwealth Caribbean: The ideal of policy versus the reality of practice. *Comparative Education, 37*(1), 107–134. <http://www.jstor.org/stable/3099735>
- Jennings, Z. (2017). Interventions in schools' curricula to achieve quality in learning: Experiences from the Commonwealth Caribbean. *Compare: A Journal of Comparative and International Education, 47*(2), 1–17. <https://doi.org/10.1080/03057925.2017.1331120>
- Kenny, J. D., & Cirkony, C. (2022). Using systems perspectives to develop underlying principles for educational reform. *Australian Journal of Teacher Education (Online), 47*(1), 80–99. <https://doi.org/10.14221/ajte.2022v47n1.6>
- Kessler, I., & Bach, S. (2014). Comparing cases. In P. K. Edwards, J. O'Mahoney, & S. Vincent (Eds.), *Studying organizations using critical realism: A practical guide* (pp. 168–184). Oxford University Press.
- Lam, C. C., Alviar-Martin, T., Adler, S. A., & Sim, J. B. Y. (2013). Curriculum integration in Singapore: Teachers' perspectives and practice. *Teaching and Teacher Education, 31*, 23–34. <https://doi.org/10.1016/j.tate.2012.11.004>
- Lemke, J. L., & Sabelli, N. H. (2008). Complex systems and educational change: Towards a new research agenda. *Educational Philosophy & Theory, 40*(1), 118–129. <https://doi.org/10.1111/j.1469-5812.2007.00401.x>
- Lewy, A. (1990). Curriculum try-out. In H. J. Walberg, & G. H. Haertel (Eds.), *The International Encyclopedia of Educational Evaluation* (pp. 203–205). Pergamon Press.
- Lincoln, Y. S., & Guba, E. G. (1980). The distinction between merit and worth in evaluation. *Educational Evaluation and Policy Analysis, 2*(4), 61–71.

- Local graduates ignored. (2013, March 2). *Trinidad and Tobago Guardian*. <https://www.guardian.co.tt/article-6.2.396498.b0aa6a4736>
- Maxwell, B., Sharples, J., & Coldwell, M. (2022). Developing a systems-based approach to research use in education. *The Review of Education, 10*(3), e3368. <https://doi.org/10.1002/rev3.3368>
- Maxwell, J. A. (2021). The importance of qualitative research for investigating causation. *Qualitative Psychology, 8*(3), 378–388. <https://doi.org/10.1037/qup0000219>
- Maxwell, J. A. (2022). Critical realism as a stance for designing qualitative research. In U. Flick (Ed.), *The Sage Handbook of Qualitative Research Design* (pp. 142–154). SAGE.
- Maxwell, J. A. (2024). A Realist Approach for Mixed Methods Research. In Y. Shah (Ed.), *Philosophical Foundations of Mixed Methods Research* (pp. 152–168). Routledge.
- Mertens, D. M., & Wilson, A. T. (2012). *Program evaluation theory and practice: A comprehensive guide*. Guilford Press.
- Meyer, S. B., & Lunnay, B. (2013). The application of abductive and retroductive inference for the design and analysis of theory-driven sociological research. *Sociological Research Online, 18*(1), 86–96. <https://doi.org/10.5153/sro.2819>
- Modell, S. (2009). In defence of triangulation: a critical realist approach to mixed methods research in management accounting. *Management Accounting Research, 20*(3), 208–221. <https://doi.org/10.1016/j.mar.2009.04.001>
- Montero-Sieburth, M. (1992). Models and practice of curriculum change in developing countries. *Comparative Education Review, 36*(2), 175–193. <https://doi.org/10.1086/447098>
- Morrison, K. (2003). Complexity theory and curriculum reforms in Hong Kong. *Pedagogy, Culture & Society, 11*(2) 279–302. <https://doi.org/10.1080/14681360300200174>
- Morgan, D. L., & Hoffman, K. (2021). Searching for qualitatively driven mixed methods research: a citation analysis. *Quality & Quantity, 55*, 731–740.
- Morse, J., & Niehaus, L. (2009). *Mixed method design: Principles and procedures*. Left Coast Press.
- Morse, J., & Cheek, J. (2014). Making room for qualitatively-driven mixed-method research. *Qualitative Health Research, 24*(1), 3–5.
- Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement and validation. *American Journal of Evaluation, 24*, 315–340. <https://doi.org/10.1177/109821400302400303>
- Mukumbang, F. C. (2023). Retroductive theorizing: A contribution of critical realism to mixed methods research. *Journal of Mixed Methods Research, 17*(1), 93–114. <https://doi.org/10.1177/15586898211049847>
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 international results in Reading*. TIMSS & PIRLS International Study Center. Boston College. <http://timssandpirls.bc.edu/pirls2016/international-results/>
- Nakabugo, M. G., & Siebörger, R. (2001). Curriculum reform and teaching in South Africa: Making a “paradigm shift”? *International Journal of Educational Development, 21*(1), 53–60. [https://doi.org/10.1016/S0738-0593\(00\)00013-4](https://doi.org/10.1016/S0738-0593(00)00013-4)

- Newton, E. H. (1993). The secondary headship: perceptions, conceptions, performance and reactions of headteachers in Barbados. *Journal of Educational Administration, 31*(2), 22–40. <https://doi.org/10.1108/09578239310041864>
- Nilsen, P. (2015). Making sense of implementation theories, models, and frameworks. *Implementation Science, 10*(1) 53–65. <https://doi.org/10.1186/s13012-015-0242-0>
- Nilsen, P. (2020). Overview of theories, models and frameworks in implementation science. In P. Nilsen, & S. Birken (Eds.), *Handbook on Implementation Science* (pp. 8–31). Edward Elgar Publishing.
- O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research, 78*(1), 33–84. <https://doi.org/10.3102/0034654307313793>
- Onwuegbuzie, A. J., Slate, J. R., Leech, N. L., & Collins, K. M. (2007). Conducting mixed analyses: A general typology. *International Journal of Multiple Research Approaches, 1*(1), 4–17. <https://doi.org/10.5172/mra.455.1.1.4>
- Onwuegbuzie, A. J., Slate, J. R., Leech, N. L., & Collins, K. M. (2009). Mixed data analysis: Advanced integration techniques. *International Journal of Multiple Research Approaches, 3*(1), 13–33. <https://doi.org/10.5172/mra.455.3.1.13>
- Palinkas, L. A., Mendon, S. J., & Hamilton, A. B. (2019). Innovations in mixed methods evaluations. *Annual Review of Public Health, 40*, 423–442. <https://doi.org/10.1146/annurev-publhealth-040218-044215>
- Parra, J. D., Said-Hung, E., & Montoya-Vargas, J. (2021). (Re) introducing critical realism as a paradigm to inform qualitative content analysis in causal educational research. *International Journal of Qualitative Studies in Education, 34*(2), 168–182.
- Peurach, D. J. (2011). *Seeing complexity in public education: Problems, possibilities, and success for all*. Oxford University Press.
- Polit, D. F., & Beck, C. T. (2010). Generalization in quantitative and qualitative research: Myths and strategies. *International Journal of Nursing Studies, 47*(11), 1451–1458. <https://doi.org/10.1016/j.ijnurstu.2010.06.004>
- Polly, D., & Hannafin, M. J. (2011). Examining how learner-centered professional development influences teachers' espoused and enacted practices. *The Journal of Educational Research, 104*(2), 120–130. <https://doi.org/10.1080/00220671003636737>
- Porter, A., & Smithson, J. L. (2001). *Defining, developing, and using curriculum indicators: CPRE Research Report Series* [Report No. RR-048]. Consortium for Policy Research in Education. https://www.cpre.org/sites/default/files/researchreport/788_rr48.pdf
- Rogan, J. M. (2007). An uncertain harvest: A case study of implementation of innovation. *Journal of Curriculum Studies, 39*(1), 97–121. <https://doi.org/10.1080/00220270500532448>
- Rosenmund, M. (2006). The current discourse on curriculum change: A comparative analysis of national reports on education. In A. Benavot, & C. Braslavsky (Eds.), *School knowledge in comparative and historical perspective: Changing curricula in primary and secondary education* (pp. 173–194). CERC Studies in Comparative Education, No 19. University of Hong Kong & Comparative Education Research Centre. https://doi.org/10.1007/978-1-4020-5736-6_11
- Saunders, R. P. (2016). *Implementation monitoring and process evaluation*. Sage.
- Sayer, A. (2000). *Realism and social science*. Sage.

- Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, 72(3), 387–431. <https://doi.org/10.3102/00346543072003387>
- Stake, R. (1976). To evaluate an arts program. *Journal of Aesthetic Education*, 10(3/4), 115–133. <https://doi.org/10.2307/3332066>
- Stake, R. (2004). *Standards based and responsive evaluation*. Sage.
- Stake, R. (2006). *Multiple case study analysis*. Guilford Press.
- Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.
- Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research. Integrating quantitative and qualitative approaches in the social and behavioural sciences*. Sage.
- Teddlie, C., & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research*, 1(1), 77–100. <https://doi.org/10.1177/1558689806292430>
- Terry, G., Hayfield, N., Braun, V., Clarke, V. (2017). Thematic analysis. In C. Willig & W. S. Rogers (Eds.), *The SAGE Handbook of Qualitative Research in Psychology* (pp. 17–37). Sage Publications.
- Thirsk, L. M., & Clark, A. M. (2017). Using qualitative research for complex interventions: The contributions of hermeneutics. *International Journal of Qualitative Methods*, 16(1), 1–10. <https://doi.org/10.1177/1609406917721068>
- Trinidad and Tobago Ministry of Education [TTMoE]. (2013a). *Primary school curriculum instructional toolkit-Infants 1 to Standard 4–5*. CPDD. <http://moe.gov.tt/Education/Primary/Instructional-Toolkit/Toolkit-Print-Version>
- Trinidad and Tobago Ministry of Education [TTMoE]. 2013b. Primary school guides-Infants 1 to Standard 5. Curriculum Planning and Development Division. <http://moe.gov.tt/Education/Primary/Curriculum-Guides>
- Trinidad and Tobago Parliament. (2014). *Eighth report of the Joint Select Committee on ministries, statutory authorities, and state enterprises: Primary education*. Trinidad and Tobago Government.. <http://www.ttparliament.org/reports/p10-s4-JSC1-20140516-PSE-r8.pdf>
- Wallace, F., Blase, K., Fixsen, D., & Naoom, S. (2008). *Implementing the findings of research: Bridging the gap between knowledge and practice*. Education Research Service.
- Webb, P. T., & Gulson, K. N. (2013). Policy intensions and the folds of the self. *Educational Theory*, 63(1), 51–68. <https://doi.org/10.1111/edth.12009>
- Yan, C. (2012). 'We can only change in a small way': A study of secondary English teachers' implementation of curriculum reform in China. *Journal of Educational Change*, 13(4), 431–447. <https://doi.org/10.1007/s10833-012-9186-1>
- Younas, A., Pedersen, M., & Inayat, S. (2023). Practical strategies to identify and address discordant findings in mixed methods research. *Research Methods in Medicine & Health Sciences*, 4(1), 24–33.

Funding

This work was part of an independent evaluation of the PCR funded by the Government of the Republic of Trinidad and Tobago [IDB Loan 2138/0C-TT/ RFP # IRPC. 2/].

Supplemental Table 1

Instrumentation for Projects 1 and 2

Variable and measure (Project 1)	No. of items (α)	Instrument source/ construction
Teacher individual characteristics	7	Author constructed
Use of curriculum core components (activities)	6 (.74)	Author constructed
Use of curriculum materials	8 (.74)	Author constructed
Use of curriculum subject areas	10 (.74)	Author constructed
Levels of use	8	Adapted (Hall et al., 2006)
Integration strategies (multi, inter, & trans-disciplinary)	10 (.87)	Based upon Drake (2012)
Curriculum orientations	30 (.93)	Used directly (Cheung, 2000)
Perceptions of characteristics of the curriculum	26 (.91)	Adapted (Flight et al., 2011)
Attitude towards curriculum	16 (.92)	Author constructed
Teacher professional learning	6 (.79)	Adapted (Geijsel et al., 2001)
Uncertainty measure	10 (.84)	Adapted (Geijsel et al., 2001)
Organizational readiness (leaders' support)	10 (.94)	Author constructed
Perception of system training	19 (.95)	Adapted (Yates, 2007)
Leadership measure	8 (.95)	Adapted (Bamburg & Andrews, 1991)
Variable and process (Project 2)	Data collection method	Analysis method
Stakeholder's perceptions	Focus groups/interviews	Thematic analysis
Antecedents/transactions/ outcomes	Observation/video/field notes/ interviews	Thematic analysis
Event logs/classroom lessons	Document analysis	Content analysis
Classroom processes and climate	Surveys/observation	Descriptive data/ themes

References for Supplemental Table 1

- Bamburg, J. D., & Andrews, R. L. (1991). School goals, principals, and achievement. *School Effectiveness and School Improvement, 2*(3), 175–191. <https://doi.org/10.1080/0924345910020302>
- Cheung, D. (2000). Measuring teachers' meta-orientations to curriculum: Application of hierarchical confirmatory factor analysis. *The Journal of Experimental Education, 68*(2) 149–165. <https://doi.org/10.1080/00220970009598500>
- Drake, S. M. (2012). *Creating standards-based integrated curriculum: The common core state standards edition (3rd ed.)*. Sage.
- Flight, R. L., Allaway, A. W., Kim, W. M., & D'Souza, G. (2011). A study of perceived innovation characteristics across cultures and stages of diffusion. *Journal of Marketing Theory and Practice, 19*(1) 109–126. <https://doi.org/10.2753/MTP1069-6679190107>

Geijsel, F., Slegers, P., van den Berg, R., & Kelchtermans, G. (2001). Conditions fostering the implementation of large-scale innovation programs in schools: Teachers' perspectives. *Educational Administration Quarterly*, 37(1), 130–166. <https://doi.org/10.1177/00131610121969262>

Hall, G. E., Dirksen, D. J., & George, A. A. (2006). *Measuring implementation in schools: Levels of use*. SEDL.

Yates, S. M. (2007). Teachers' perceptions of their professional learning activities. *International Education Journal*, 8(2), 213–221.

Supplemental Table 2

Cross-Case Analysis of Antecedents of Curriculum Implementation (Project 2)

Antecedent	Site A	Site B	Site C	Site D	Site E	Site F	Site G	Site H	Site I
Positive	Just do it teacher attitude <i>Training & support (coaching)</i>	Teacher as problem-solver <i>Constant collaboration</i> <i>Curriculum committee structure</i> <i>Effective physical layout</i>	<i>Teacher attitude</i> <i>Informal collaboration</i>						
Negative	Lack of resources	<i>Time constraints</i> Inadequate training Lack of alignment with LSA	Lack of resources Inadequate training and support	Lack of time, space and planning Lack of training	Time, resources, & leadership deficiencies Low quality training/unwanted autonomy <i>Imperfections of curriculum</i> <i>Teacher beliefs and priorities</i>	Lack of resources <i>Teacher attitudes</i> Ineffective teacher training <i>Increased workload</i>	<i>Lack of understanding</i> <i>Alternate teacher priorities</i> <i>Physical structure</i> Training and support Lack of alignment with LSA	<i>Expectations unrealized</i> Inadequate curriculum materials <i>Increased workload</i> Deficit thinking	<i>Teacher attitudes</i> Training & support <i>Physical environment</i> Lack of alignment with LSA

Note. LSA = Large Scale Assessments. Bold type indicates repeated concerns.

Supplemental Table 3

Category Zooming: Distribution of Collaboration, Resources, and Training Antecedents across 9 Sites (Project 2)

Case (quantitative classification)	Collaboration	Resources	Training
<i>High implementation sites</i>			
Site A	Constant collaboration	Lack of physical resources	
Site B	Formal structures	Time constraints	Inadequate training/ support
Site C	Informal collaboration	Lack of physical resources	Inadequate training
<i>Medium implementation sites</i>			
Site D		Lack of physical resources	Inadequate training
Site E		Lack of physical resources	Inadequate training
Site F		Lack of physical resources	Inadequate training
<i>Low implementation sites</i>			
Site G		Lack of physical space	Inadequate training/ support
Site H			Inadequate training/ support
Site I			Inadequate training/ support