RISE

RESEARCH ON IMPROVING
SYSTEMS OF EDUCATION

The Secret Sauce of Development Professionals: Tools for Assessing TOR Potential to Source Scalable Learning Interventions

by Adam Barton
University of Cambridge

Summary

Terms of reference (TORs) play an outsized role in driving scalable educational programming. These procurement documents shape, constrain, and signal programme priorities and possibilities. Successful funders and implementers across the globe hold rich processual knowledge about this documentation, which they use to draft and assess TORs. This project explores such best-practice knowledge around TOR review, seeking to support the design and implementation of educational programmes that can improve learning at scale in developing contexts.
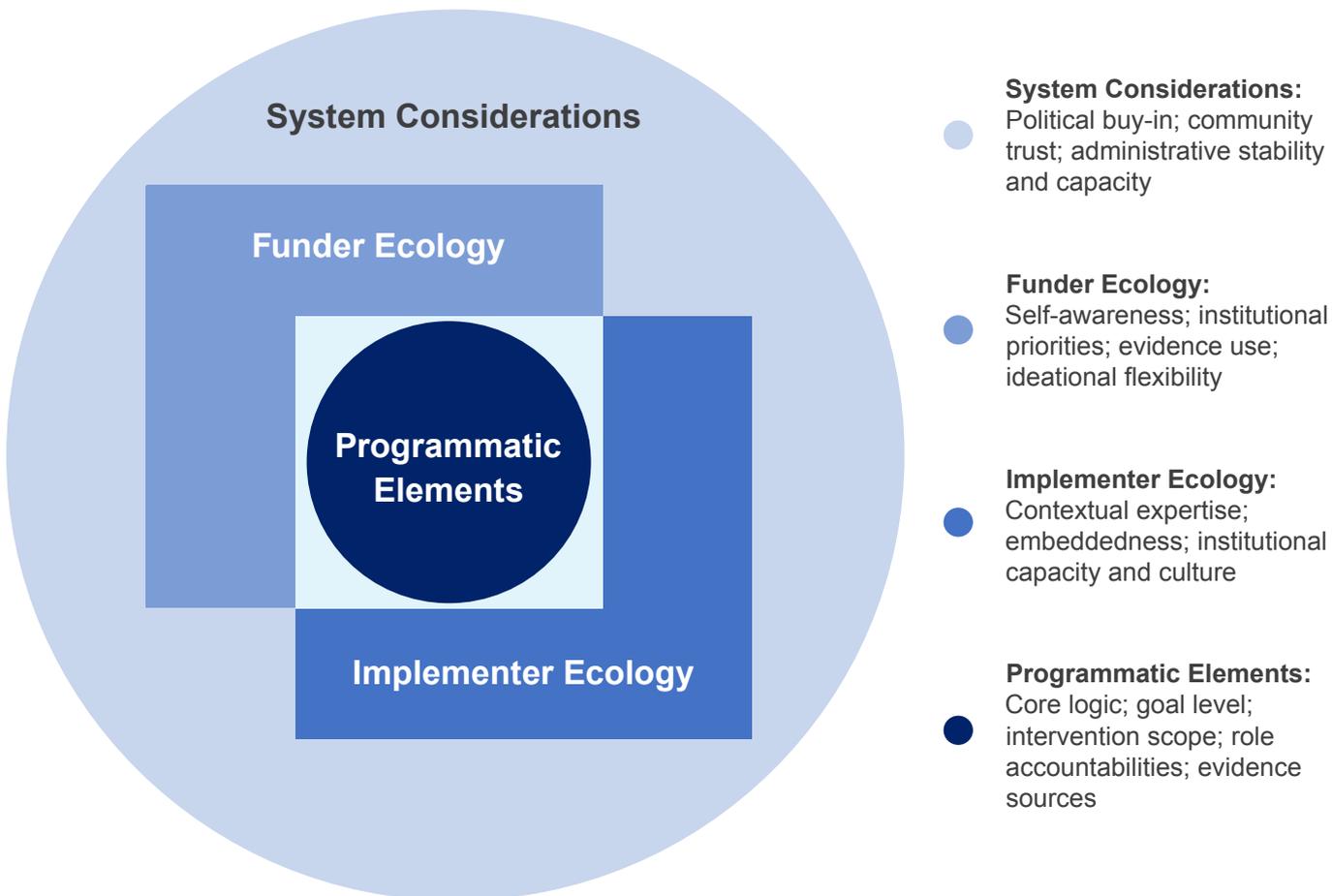
This research builds a practical guide aiming to codify best practice knowledge for effective TOR development. It specifically surfaces the analytical processes and prioritisations that successful practitioners use to evaluate TOR potential. Data emerges from hands-on interviews with 18 expert international funders and implementers who reviewed a real-world TOR. A specialist panel then stress tested preliminary findings. Their insights inspired the creation of two tools: a conceptual framework illustrating a successful procurement ecosystem and a reflective checklist for TOR review.

Development practitioners can use these tools to scaffold their thinking when designing and assessing TORs for scalable learning interventions in developing contexts. The framework and checklist aim to inspire, rather than prescribe, procurement practice. Practitioners should begin by considering the specific aims and needs of their proposed projects. They can then assess TOR potential along the following four framework elements and in line with ten reflection questions.

Key Points

- Terms of Reference (TORs) and other procurement documents directly shape whether and how educational programming functions on the ground.

- Successful development practitioners hold rich knowledge for designing and assessing TORs that can sustainably improve learning at scale in developing contexts.

- This Insight Note codifies experts' evaluative processes for effective TOR creation via two tools: a conceptual framework and a reflective checklist.

- Development professionals structure their thinking in terms of 10 reflective questions across four interlocking levels.

- Practitioners can use this framework and checklist to scaffold TOR development and review in line with experts' best practice processes.

Procurement Ecosystem: A Guiding Framework for TOR Success

System Considerations

Funder Ecology

Programmatic Elements

Implementer Ecology

System Considerations:
Political buy-in; community trust; administrative stability and capacity

Funder Ecology:
Self-awareness; institutional priorities; evidence use; ideational flexibility

Implementer Ecology:
Contextual expertise; embeddedness; institutional capacity and culture

Programmatic Elements:
Core logic; goal level; intervention scope; role accountabilities; evidence sources

Reflective Checklist: Ten Questions for TOR Review

1. To what extent does this intervention **align with existing government** and **civil society activities**?

2. To what extent does this intervention tap into, or propose to cultivate, **political will** and **community appetite**?

3. To what extent does this intervention demonstrate a mentality of **iteration** and **experimentation**?

4. To what extent does the funder adopt the role of a **nurturing collaborator**?

5. To what extent does **implementer selection** prioritise the **organisational characteristics** most likely to **enable systemic change**?

6. To what extent does this intervention logically respond to **a clear problem statement**?

7. To what extent does this intervention **target mindset and behavioural change** to support intended practice improvements?

8. To what extent does this intervention logically lead to **systems-level outcomes**?

9. To what extent is this intervention based on **localised**, **contextual knowledge** and **expertise**?

10. To what extent are **measurement and evaluation** criteria justified by **core programme logic**?

Scope and summary: the power and promise of TORs

The world faces an urgent learning crisis (The State of Global Learning Poverty: 2022 Update, 2022). Despite making great strides in access and attainment, nations are not on track to meet global goals around educational quality (Global Education Monitoring Report Team, 2021). Among many alarming statistics: over one-third of children—some 300 million—will fail to complete primary school in line with the 2030 Sustainable Development Goal targets (UNESCO Institute for Statistics, 2022).

Promising innovations, however, are on the horizon. Governments and civil society organisations across the globe have successfully implemented myriad models for transforming learning at scale.[1] These range from full system reforms, including the New Delhi Happiness Curriculum,[2] to non-profit-led interventions in areas such as peer mentorship for life skills development.[3]

This transformative work requires intensive collaboration between public and private organisations. National governments are—and must remain—the primary drivers of public education. But external funders fundamentally shape educational planning across the globe (Global Education Monitoring Report Team, 2021). These include both government aid agencies, such as FCDO or USAID, and a host of non-state[4] actors financing international programme and policy initiatives. Private foundations are a particularly powerful force for sourcing and supporting novel interventions. They enjoy long-term funding horizons, with ample flexibility to iteratively tinker with and quickly ramp up supportive capacity for new models.

In their search to support global education, funders have grown increasingly preoccupied with questions of scalability and sustainability (Reddy & Narain, 2022). This stems from a host of international insights into the failure of piecemeal and transient educational interventions (Hargreaves et al., 2009). Examples of such challenges abound across the globe, from parental resistance to technical education in San Diego to failed inquiry-based teacher training in Senegal (Winthrop et al., 2018). Consequently, global funders increasingly prioritise packages that partner with civil society and governments for sustainable, scalable transformation.

TORs drive partnerships, which in turn drive global education provision.

International funding collaborations involve a complex array of formal and informal partnership arrangements. Informal activities, such as inter-school learning networks, certainly shape learning ecosystems. However, formal contracting ultimately drives global education collaboration. These agreements entail one party commissioning specific activities or products from another—for example, a British foundation paying a Brazilian non-profit to develop and lead local teacher training initiatives.

Formal contracts involve a process called procurement—methods for sourcing goods or services between actors (Walker, 2015). This document-heavy endeavour requires all parties to codify precise objectives, activities, roles, and outputs. Here, the guiding text is often called the "terms of reference" (TOR)—a formal procurement document that sets out funders' expectations for their implementing partners. In the above example, the British foundation would develop a TOR detailing its guiding rationale, the sorts of teacher training it expects, and all results the non-profit must deliver to receive funding.

---

[1] For inspiration, see horizon scanning efforts by organisations including Catalyst 2030, 2022. Big Change, 2021 and Winthrop, Barton and McGivney, 2018. (*Our 10 Big Hopes for Change: A Review of the Latest Insights and Pioneers Making Impact on the Ground.*, 2010; Reddy & Narain, 2022; Winthrop et al., 2018)

[2] See the Delhi State Council of Educational Research and Training web page for details.

[3] See the CAMFED Learner Guide programme web page for details.

[4] Non-state actors are individuals and organisations with the power to influence education financing, planning, and provision, without belonging to the established institution of an organised government. These include, for example, civil society groups and multilateral funding bodies.

Terms of reference (TORs) constitute the "rules of the game" for funders and implementers as they make sense of, and ultimately intervene in, learning systems. They serve an important signalling function, signposting funders' priorities, rationales, and visions.[5] This idea fits within a broad theoretical camp called "sensemaking," which examines the social and cognitive processes through which individuals interpret and understand a complex object or event—such as development programme implementation (Brown et al., 2015). Here, leaders use various social tools to "give sense" to other actors as they construct and implement shared social projects (Gioia & Chittipeddi, 1991).

Funders, in this case, use TORs to influence the way that procured partners construct and implement their educational interventions. Presenting a long inception period with participatory needs assessment, for example, might indicate that a funder prioritises embeddedness and alignment with existing activities; they would thus attract partners with a similar focus, or lead collaborators to realign their approach. Meanwhile, an emphasis on continuous quantitative monitoring—regular reporting on numbers of textbooks provided, for example, or teachers trained—could lead implementers away from outcome analyses of mindset or behaviour change. Of course, this is to say nothing of the contractual elements fundamentally constraining the partnership, including timelines, budgets, and discrete deliverables. In these many other ways, procurement documents such as TORs directly influence implementers' programmatic conceptions, decisions, and behaviours.

> Terms of reference (TORs) constitute the "rules of the game" for funders and implementers as they make sense of, and ultimately intervene in, learning systems.

TORs represent a key moment to translate evidence on "what works" into high impact practice on the ground.

TORs, then, are key artefacts that codify and communicate conceptions of "what works," "what is missing," and "what to do" to a broad range of collaborating stakeholders. More than bureaucratic hurdles, they represent pivotal opportunities to synthesise best practice, share programmatic rationale, and guide interventions toward sustainable systems outcomes. They are the procurement point at which funders and implementers can align for success—or set off down the wrong track, entirely.

In sum, TORs exert great programmatic influence; they shape how money reaches frontline practitioners, and how those implementers conceptualise educational change on the ground. In practice, however, they are often treated as no more than an administrative box-checking exercise. Organisations rarely afford these key documents the conceptual importance they deserve. This mismatch presents a clear opportunity for institutions and individuals to reimagine TORs that can foster sustainable systems transformation.

At the same time, this opportunity poses an outsized challenge for funders and implementers. TOR design entails representing a dynamic ecosystem[6] on paper. Designers must pin down actors and elements—such as policymakers, donors, and temporary programme structures—that intersect, interact, and diverge. These interconnected components, and their resultant programming, are in constant flux. All the while, designers need to translate and map best practice evidence onto this dynamic ecosystem in a way that is truly implementable.

> In sum, TORs exert great programmatic influence; they shape how money reaches frontline practitioners, and how those implementers conceptualise educational change on the ground.

Further complicating powerful procurement, this small—but important—piece of the evidence ecosystem remains understudied. And practical guidance on TOR design remains exceedingly limited. Though best practice knowledge on procurement processes certainly exists, it remains locked away in the minds and assets of a small cohort of expert funders and implementers.

---

[5] For detailed exploration of the inter-party signaling process through contractual documentation, see Connelly, Certo, Ireland, and Reutzel (2011).

[6] We adopt the term "ecosystem" in this report to reference the extensive interconnection and collaboration between all actors, organisations, and elements involved in supporting learners and learning. These range from parents and principals to community appetite and even fiscal time horizons. Procurement processes are inherently integrated into these complex social and political environments.

Successful development practitioners hold rich procedural knowledge about procurement best practices.

This project stems from the hypothesis that successful[7] professionals[8] hold rich procedural knowledge about "what works" in programme procurement.[9] They tap these sophisticated, experientially honed insights to design and source programming globally. This know-how serves as an intuitive filter, allowing experts to assess the potential of TORs to source programmes that can deliver improved learning outcomes. In this way, development professionals instinctively spot the enablers of and barriers to successful programme implementation.

The question becomes, then: could we tap successful practitioners' intuitive insights about how to design and evaluate TORs for successful programme implementation?

We aimed to capture and codify this TOR design knowledge into practical tools for development professionals designing or reviewing procurement documents at large funding institutions. To this end, we adopted a cognitive sensemaking lens to unpack the reflective processes that successful practitioners use to evaluate TOR promise and potential. This was an attempt to explore how practitioners think about procurement, emphasising the knowledge, assumptions, and intuitions guiding their evaluative practice. We used cognitive interview tactics to highlight the key questions and considerations driving practitioners' decision-making, ultimately aiming to:

- Understand how experienced, successful development practitioners make sense[10] of TORs when designing and procuring scalable learning interventions

- Codify the cognitive processes and key considerations that practitioners use to identify high-leverage[11] or high-risk programme components within TORs

- Highlight discrete examples of procurement best practices and pitfalls

> The question becomes, then: could we tap successful practitioners' intuitive insights about how to design and evaluate TORs for successful programme implementation?

Method: experts making sense of TOR development and evaluation

This research is about tapping the knowledge and processes that successful practitioners bring to bear when designing and assessing TOR potential. Such implicit knowledge, however, is exceedingly difficult to access. Traditional semi-structured interview techniques necessarily divorce thinking processes from their regular context. Asking a practitioner explicitly about how they assess TORs leads them to share only what they consciously prioritise. What results is reflective abstraction, pulling the researcher ever farther from participants' processual realities (Barton, 2015).

To tap implicit knowledge and follow the sensemaking process, then, we stimulated participants' ongoing reflection using cognitive interviews—and, more specifically, stimulated probing (van Braak et al., 2018). Practically, this involved presenting participants with a real-world artifact—a genuine TOR—and asking them to narrate their thought processes as they attempted a real-world exercise: assessing the document's potential, as in a procurement review.

---

[7] We define success as demonstrating a track record of regularly funding and/or implementing sustainable, evidenced programmes that have improved basic education learning at scale in developing contexts. Regularity requires at least one programme annually over the past 5 years. Evidenced entails presenting, at a minimum, internal qualitative evidence of improved academic or socio-cognitive gains for beneficiaries. Learning involves at least an explicit focus on academic- or cognitive-domain growth and improvement. Finally, scale requires programme implementation across at least one entire jurisdictional level—be that a municipality or school chain.

[8] Professionals refers to both funders and implementers of educational programming in developing contexts.

[9] The programme procurement process analysed in this work involves reviewing TORs and other programme documents to assess the feasibility and value of whole educational programmes.

[10] Making sense, in this context, involves uniting experience, intuition, and knowledge to assess the chances that a given intervention will successfully improve learning at scale.

[11] These are components reflected in a TOR that have the greatest bearing on successfully improving learning at scale.

We began by drafting a list of development professionals for expert interview, in collaboration with the RISE team. We prioritised professionals according to their role in the procurement process, track record of success, and diversity of institutional and geographic contexts.[12] We invited 40 participants to collaborate, 22 of whom declined or passed us to colleagues. Ultimately, we interviewed 18 professionals in June and July 2022. Though 12 worked at funding institutions and 6 at implementing organisations, this functional line was quite indistinct; nearly all interviewees had several years' experience successfully implementing educational programming, often transitioning between practitioner and funder roles throughout their careers.

During the one-hour online interviews, we utilised cognitive interview techniques that prompted professionals to narrate their thought processes as they reviewed a real-world TOR (Fisher, 1992). This procurement document, sourced from the public archives of a multilateral development funder, was selected to match research scope around sourcing systems-level basic education programming. We complemented this review narration with several semi-structured interview questions, seeking additional insights into high- and low-leverage TOR components.

Inductive thematic analysis of interview transcripts and notes focused on parsing the underlying insights and processes driving participants as they evaluated TORs (Clarke et al., 2015). We codified these themes into a conceptual framework to assess the various ecosystem elements surrounding TOR implementation. These insights additionally crystallised into a reflective checklist, aiming to mirror the thought processes successful practitioners utilised during their TOR reviews. These tools can scaffold practitioners' TOR assessment in line with experts' evaluative processes. We refined these tools through an online panel discussion, prompting five participants—a subset of the initial interviewees—to use the checklist and framework to assess a TOR and share emergent insights and sticking points.

Research Method

Guiding Questions:

- How do **successful** development **practitioners approach TOR review** during procurement for basic education learning programmes at scale?

- What do **successful** development practitioners know about **programmatic best practices** for basic education learning programmes at scale, particularly as **revealed in TORs** and procurement documents?

- What **cognitive** (evaluative, reflective, knowledge) **processes** do **successful** development **practitioners use** to spot **high-leverage** programme components—or combinations of components—during procurement?

Research Process:

- **Identified** development funders and implementers experienced in high-level TOR drafting, review, and enactment with an evidenced track record of programmatic success.

- **Interviewed** 18 successful development professionals in a semi-structured fashion, prompting participants to narrate their thought processes and reflections while reviewing a real-world TOR.

- **Codified** procurement insights through thematic analysis into a conceptual framework and reflective checklist that mirrors experts' reflective review processes.

- **Piloted** the framework and checklist with a subset of expert participants, using the tool to review a TOR and sharing emergent insights and sticking points.

---

[12] We selected only those professionals with institutional responsibility for reviewing and approving TORs and/or TOR bids. Professionals must have been at institutions serving developing contexts, and we specifically chose one-third of interviewees regularly based in these geographies. We define success as demonstrating a track record of regularly funding and/or implementing sustainable, evidenced programmes that have improved basic education learning at scale in developing context. Regularity requires at least one programme annually over the past five years. Evidenced entails presenting, at a minimum, internal qualitative evidence of improved academic or socio-cognitive gains for beneficiaries. Learning involves at least an explicit focus on academic- or cognitive-domain growth and improvement. Finally, scale requires programme implementation across at least one entire jurisdictional level—be that a municipality or school chain.

Findings: Framing the procurement ecosystem

Development professionals preliminarily structured their thinking in terms of four interlocking levels of analysis: **system considerations**, **funder ecology**, **implementer ecology**, and **programmatic elements**. Experts used this fundamental framing to assess TOR potential, examining possibilities and pitfalls in procurement documents across each of these four dimensions. They used these elements to conceptualise what we term the "procurement ecosystem"—the constellation of actors and factors in TORs that interact to enable or impede educational programme implementation on the ground.

Though these elements may appear discrete and hierarchical, they intersect and influence each other in multiple directions. Implementers' apparent leadership strength, for example, may lead funders to provide greater TOR flexibility; meanwhile, funder capacity may itself affect political appetite at the systems level. We consequently adopt the terms "ecology" and "ecosystem" to reflect interviewees' wide-ranging focus on diverse, interlocking organisational factors— from culture and leadership to embeddedness and flexibility. This draws from organisational studies literature focused, especially, on analysing partnership and collaboration; groups such as funding or implementing institutions organise and evolve in response to unpredictable and ever-changing internal and external factors, from staffing to political climates (Ivery, 2007).

System considerations

Beginning at the highest level, **system considerations** fundamentally constrain implementational possibilities. These include elements such as *political buy-in, community trust,* and *administrative stability and capacity*.

Experts regularly stressed *political buy-in* and *community trust* above all other factors. Quite simply: implementing learning at scale begins and ends with the community's and administration's appetite for change. Political buy-in, here, extends beyond the will of public educational organisations—such as the ministry of education—to also include the broader commitment and alignment of the agenda-setting political regime. Community trust involves a host of affective dimensions, including confidence in implementing institutions, appetite for change, and a sense of shared ownership. Experts therefore assessed the extent to which documentation described concerted efforts to cultivate demand in line with proposed programming, while securing constituent confidence in implementers' intentions and abilities. This involved searching for buy-in-related activities, from stakeholder mapping to dialogic discussion forums.

With scalability and sustainability as principal TOR goals, professionals consistently questioned *administrative stability and capacity* to manage planned change over extended time horizons. They sought evidence that procurement considered ministerial and civil society capacity to sustain scaled programming after the temporary delivery vehicle faded away. This included, for example, explicit reflections on regime change or trainings for ministry officials. As one funding expert put it:

> " Transforming education systems is largely a political challenge; no amount of technical expertise can overcome a lack of community trust or political will [across administrations]. "

Funder ecology

Experts further reflected on **funder ecology** when assessing TOR potential. Key components involve funder *self-awareness, institutional priorities, evidence use,* and *ideational flexibility*.

Self-awareness, here, involves demonstrating clear reflection on what has been tried, what has not worked, and what remains unknown. This is a matter of "intellectual honesty" or "institutional humility," as experts framed it. They assessed whether funders were transparent in their procurement prose about extant efforts, past struggles, and institutional learnings. This acknowledgement of historical challenges and incomplete evidence drives positive procurement by signalling a desire for genuinely open partnership.

Interviewees additionally stressed that funders must reflect on and share the *institutional priorities* driving procurement. Successful procurement, according to experts, involves mobilising actors by aligning their diverse priorities and preferences. This is a reciprocal act. Professionals must, first, identify opportunities in line with their

institutional goals; they then seek common ground with community priorities, identifying shared interests and weaving partnership possibilities. This alignment and signaling process drives sustainable programme design; opening the black box of intervention decision-making breeds the buy-in necessary for shared. As one funder noted, "While recognising that they have clear institutional cultures and priorities, [funders] must principally demonstrate that they're open to being wrong—and undoubtedly have been in the past."

This consequently connects with notions of *evidence use* and *ideational flexibility*. The first analytical question, here, involves the type of evidence funders seek and cite within the TOR. Experts assessed whether funders presented their rationale as grounded in "best-bet evidence"—that is, data from local or global sources that can give partners some faith in the programme's promise. Interviewees observed that the lowest bar for assessing TOR potential is simply observing whether procurement language is grounded in citable "best practice" learnings from similar contexts. But, beyond this, TORs must justify why these data are, in fact, relevant for their proposed context; powerful procurement language involves reflecting on the strengths and limitations of such evidence, such as whether localised testimonials might be more compelling than randomised control trial data from a different geography.

The second question assesses a spirit of prescriptivism: whether funders adopt a rigid stance toward partnership and programming—or present TORs in a co-constructive fashion. Experts analysed, for example, whether the funders developed the TOR through an open review period with outside feedback. They further analysed whether the document expressed an openness to shift priorities in accord with implementer analyses. According to one implementer-turned-funder:

> "Funders overthink what needs to be done; they think it their responsibility to devise solutions and simply find an executor to deliver results—when they really should be demonstrating openness to outside ideas and evidence."

Implementer ecology

Practitioners extensively analysed **implementer ecology**, as well. Central elements include *contextual expertise, embeddedness*, and *institutional capacity and culture*.

Searching for this *contextual expertise* and *embeddedness*, experts assessed the extent to which TORs discussed implementers' ongoing relationships with and positionality toward local communities and administrations. Procurement documents are most successful when explicitly seeking implementers with long-term visions of and commitments to the target context. Indeed, multiple experts stressed the importance of one simple line in a TOR: "We prioritise organisations that are locally-based with a proven track record of relationships and implementation in this specific context."

Institutional capacity and culture of the implementing organisation build on these dimensions. The question experts ask: who is it that ultimately leads this work? What are their connections to, experiences with, and capacity to support the programme community? In assessing these questions, they searched for the ways that a TOR conceptualised what we term "relational evidence"—factors relevant to an organisation's community embeddedness, diversity of connection, and experience successfully cultivating and maintaining relationships. This further included relational analyses within the implementing organisation, itself: matters of leadership culture, succession planning, and articulated passion in line with proposed programming.

Additional considerations of *co-design* specifically emerged at the union of funder and implementer ecologies. Interviewees regularly probed what TORs revealed about the ways implementer and funders interact. They framed these as interlocking domains, each heavily relying on the other as they interact within and influence their surrounding system. The implemented programme, here, emerges from constructive efforts between both parties. Experts assessed how the TOR might allow one party to dominate the other, through one-sided responsibility for conception or limited opportunities to externalise and reflect on organisational thinking. Reflective questions

ultimately involved co-design for genuinely supportive collaboration; as one funder framed it:

> " How does the [TOR] demonstrate a support structure to help organisations on both sides rise to meet the problem? "

Programmatic elements

Programmatic elements related, specifically, to the proposed intervention emerged as the final analytical unit. These comprise *core logic, goal level, intervention scope, role accountabilities,* and *evidence sources*.

Chief among these are reflections on the intervention's *core logic*—how the TOR theorises the guiding programmatic rationale. Procurement documents, here, should "show their work," as in a mathematics proof, by logically linking local challenges and demands with proposed approaches using external evidence. Echoing other experts, one practitioner noted that the biggest driver of programmatic success is "getting right that core rationale behind the programming—why we think this will actually work in practice this time, when so much else has been or could be attempted."

This links directly to conceptions of programmatic *goal level*. Experts principally examined whether the TOR articulated ultimate goals in systems-level, interpersonal terms—focusing on behaviours, mindsets, and systemic sustainability. Questions included analysis of the contextual background—if these goals and activities "logically come next in the country context"—as well as genuine aims toward sustaining scale, such as how programmatic aims strengthen, subvert, or incorporate into long-term systems processes. Perhaps most importantly, they analysed whether the TOR articulated these goals in terms of outcomes at the heart of education: teaching and learning. That is, "how will these activities actually be taken up to solve pedagogical challenges and improve learning?"

Experts consistently pointed out, however, that none of these components matter if the TOR misconceives *intervention scope*. Fundamentally, this involves the extent to which the timeframe and deliverables are realistic for the proposed processes and partnerships. Reflecting on the pitfalls of over-ambitious TORs, one multilateral funder proposed a simple calculation:

> " A key indicator of TOR potential involves determining the time and resources available per deliverable. "

This calculus directly ties to *role accountability* structures: the clear distribution of responsibility for executing articulated activities. In this way, TOR review involved assessing the clarity of each actor's role in delivery—and their capacity to deliver. Interviewees deemed a TOR promising only when it outlined discrete tasks for each actor, clear appraisals of actor capacity, and plans for training or supporting actors, as needed.

Finally, experts questioned how the document conceived of legitimate *evidence sources*. This is, first, a matter of defining what constitutes success—and why these measures should be compelling. Interviewees therefore examined how the proposed monitoring and evaluation metrics reflected, contradicted, or even impeded intervention logic. The best TORs, they observed, contemplate flexible outcome reporting—instead of focusing on rigid outputs. These documents further justify how these metrics are contextually-relevant, aligning with local needs and realities. Lastly, they consider how monitoring activities complement ongoing analysis and documentation relevant to the programme, itself, rather than inventing new practices.

Findings: A reflective checklist for TOR development and review

A reflective checklist emerges from the various intersections of the above framework. These guiding questions can scaffold TOR development and review in line with experts' best-fit practice evaluative processes. This tool mirrors the questions successful professionals ask themselves when evaluating TOR potential to source programming that can successfully improve learning at scale in developing contexts. Development professionals can use this list to systematically assess the ways in which their procurement documents demonstrate—or could be altered to demonstrate—focused attention on each question.

It is worth reiterating, however, that these questions aim only to prompt reflection. They do not prescribe documentary or programmatic practice. No TOR has the ability or aspiration to perform perfectly across each section. Before using this list, practitioners must first reflect on the purpose(s) of their planned intervention; a TOR procuring teacher training materials, for example, demands different evidence considerations than one partnering for parent-involved life skills development. With these nuanced intentions in mind, professionals can prioritise the questions most aligned with, and most feasibly addressed by, their planned practices.

> These guiding questions can scaffold TOR development and review in line with experts' best-fit practice evaluative processes.

This section presents the ten reflection questions guiding successful procurement. Various sub-questions follow from each major question.

Aligning with existing government and civil society activities

1 To what extent does this intervention **align with existing government** and **civil society activities**?

→ Does this intervention contemplate **long-term funding** arrangements?

→ Does it propose ways **to complement or augment**, rather than reinvent, **existing activities**?

→ Does it seek to **strengthen**, rather than circumvent, **public systems, organisations, and processes**?

Experts regularly began their review by assessing how the TOR addressed or assessed ongoing initiatives in the target community. This was, in part, an effort to avoid reduplicating efforts. But, more centrally, it was a matter of ensuring programmatic sustainability at scale. Practitioners examined the ways in which the TOR demonstrated systems strengthening, activity alignment, and longitudinal planning. They continuously returned to one red flag: proceed with caution if a document makes no mention of existing initiatives, or if it is framed without long-term organisational processes spanning multiple administrations.

> " Education projects don't begin with a blank slate. What existing structures do governments or communities have in place for us to leverage? "

Tapping political will and community appetite

2 To what extent does this intervention tap into, or propose to cultivate, **political will** and **community appetite**?

→ Does this intervention explicitly **acknowledge the ongoing need** to cultivate political will and community appetite, across various levels of action?

→ Does it **map and differentiate** the **specific change agents** necessary to sustain will and appetite?

→ Does it demonstrate knowledge of, or a drive to assess, specific areas of **educational demand**?

➡ Does it propose **specific processes for cultivating will** towards the intended activities?

Building on questions of alignment, interviewees stressed the importance of documents considering demand for the intended intervention. TORs should identify or seek key change agents who are already excited about proposed programming; these will become key advocates who can speak in political or community circles when funders and implementers are not in the room. This will, however, does not necessarily need to be secured before TOR elaboration; experts stressed that programmatic activities can and should include cultivating demand by showing partners what has worked, where, and why.

> " There needs to be some deep willingness there. Is this document articulating this as a problem that the government and communities in the ecosystem want to solve? "

Demonstrating an iterative and experimental mentality

3   To what extent does this intervention demonstrate a mentality of **iteration** and **experimentation**?

➡ Does this intervention identify specific **moments to pause, reflect, and iterate** on proposed activities and emergent best practices?

➡ Does it explicitly acknowledge the **lack of silver bullets** and **quick fixes**?

➡ Does it explicitly reflect on what remains **unknown** or **unclear**?

➡ Is it framed as a **best-bet attempt**, given current knowledge, while remaining **open to new learnings and inputs**?

"Open," "adaptive," "nimble," and "iterative" proved among the most common words practitioners used to describe TOR strength. They ultimately classified TORs along a spectrum of prescriptivism and adaptability. The core question, here, is whether procurement documents provide flexibility to scope, challenge, trial, and rethink programming, in line with emergent experience and evidence. This is about demonstrating self-awareness alongside a tinkering mindset; documentation should reflect upon past failures, challenges, and learnings, making space for emergent best-bet ideas. As such, procurement language should explicitly prioritise and plan for action-reflection cycles between all parties—from funders and implementers to politicians and community members.

> " You should be able to see a clear feedback loop between designers, administrators, and participants—some opportunity to pause and reflect in both design and implementation. "

Cultivating implementer capacity

4   To what extent does the funder adopt the role of a **nurturing collaborator**?

➡ Does the funder commit to intentionally **nurturing implementer capacity and networks**?

➡ Does it establish a clear and transparent process for **external experts to question and shape** the proposed intervention?

➡ Does it delineate **clear role accountabilities and responsibilities** for each partner, including government actors and community members?

➡ Does it recognise and value the **topical and technical expertise** of selected partners?

Experts assessed the extent to which TORs described collaborative and nurturing funder-implementer relationships. The TOR should ideally focus on cultivating genuine partnerships—identifying the right actors for the challenge, and scaffolding their potential through financing, connection, and training. This is all about the relationships and roles between funders and implementers. Quite simply, the TOR should position funders as open to—and indeed inviting—shifts in their theories and processes, with ample room for co-developing ideas and activities before, during, and after procurement—such as through open calls to weigh in on the TOR, itself. The document, then, might set out a community-defined problem, and invite potential partners to pitch their unique approaches, positions, processes, and needs. But it must also define clear role accountabilities and expectations, mapping precisely who is responsible for what activities, their capacity to execute, and any plans to support the actor's development.

> " I'm looking to see if this TOR describes nurturing partnership. Where is the support structure in the TOR system to help an organisation rise to meet the challenge? "

Prioritising implementer organisational characteristics relevant to systemic change

5 To what extent does **implementer selection** prioritise the **organisational characteristics** most likely to **enable systemic change**?

→ Does the selection process provide room for organisations to demonstrate their **unique core approach and change philosophy**?

→ Does it recognise, and seek to **minimise, the linguistic, financial, and technical burdens of application**?

→ Does it require organisations to demonstrate **compelling "relational evidence"** within the system, including strong local networks and a proven track record with partners across multiple leadership levels?

→ Does it enable organisations to highlight **characteristics of their leadership teams**, such as continuity plans and clarity of vision?

→ Does it push organisations to **demonstrate alignment** between **institutional mission**, organisational **track record**, and **proposed programming**?

Experts further assessed how the TOR understood the implementer characteristics most relevant to implementing change. A central measure, here, is "*relational evidence*" demonstrating organisational embeddedness, diversity of connection, and experience successfully cultivating relationships. This goes beyond foundational data on partners' track records—what they have achieved and learned in similar systems. It assesses partners' political and cultural capital, as well as their organisational culture. These might be assessed by reviewing references from systems leaders or past partners, while sourcing a diverse team including external experts for TOR development and application review. Centrally, this requires adding language about partners' leadership teams, whose passions and drive prove nearly as important as their sustainability structures and programmatic visions.

> " Good [procurement documents] sound like they're setting up a sophisticated job interview. I'm asking, 'how is this [TOR] focusing on important partner considerations, like organisational leadership, capacity, and skills?' "

Targeting mindset and behavioural change

6   To what extent does this intervention logically respond to **a clear problem statement**?

→ Do these changes relate **directly to teaching and learning practices**?

→ Do they emerge from an analysis of **classroom-level challenges**?

→ Are they intentionally sequenced to gradually **build momentum and shift mindsets**?

→ Do they actively **involve constituents** to build long-term community **ownership**?

Improving learning at scale requires shifting pedagogical behaviours. As such, experts assessed whether TORs framed programming in terms of identified pedagogical challenges and targeted teaching and learning changes. This involves seeking classroom- and learning-based analyses, defining deliverables in pedagogical terms, and recognising the human dimension of change management. As one funder put it: "All educational projects are behaviour change projects." Consequently, experts continuously sought evidence that programmatic success was framed in terms of teaching and learning behaviours—with programme logic addressing their underlying mindsets and knowledges.

"
Shifting practice is really difficult. How does this plan to get at those deeply held beliefs and experiences?
"

Responding to a community-based problem statement

7   To what extent does this intervention **target mindset and behavioural change** to support intended practice improvements?

→ Is this problem statement based on a recent assessment of **participant barriers and needs**—or does it propose a participatory scoping period?

→ Does it reference **previous solution attempts**, and hypothesis about why they failed?

→ Does it highlight and contrast **challenges and solutions from comparable contexts**?

→ Does it include **key barriers** for those at the core of proposed changes, namely **teachers and students**?

TOR review regularly began with one key reflection: what, exactly, is the perceived "problem" this programme seeks to solve? Successful practitioners questioned the clarity, contextualisation, evidence, and logic of the fundamental problem statement. Considerations included whether the TOR mapped—or sought to map—similar challenges and solutions in other locales. The document's background section often houses this comparative knowledge around successes and failures. Tying in with the question of mindset and behaviour, experts further questioned whether the TOR successfully articulated, using community-derived evidence, discrete barriers for those at the heart of proposed changes—teachers and students.

"
Am I confident this TOR is grounded in the local context, having clearly identified local problems?
"

Framing in terms of systems-level outcomes

8 To what extent does this intervention logically lead to **systems-level outcomes**?

→ Are systems-level outcomes designed with reference to **long-term durability**?

→ Are they described using a **sequential, processual lens**?

→ Are they compellingly justified with **narration of the evidence** driving the proposed change process?

→ Do they involve a strategy to **build system capacity** and **strengthen local talent pipelines**?

Experts consistently interrogated how the TOR's core idea could lead to enduring systems change. In their framing, only such a systems lens empowers procurement for long-term learning outcomes. This boils down to a focus on what we might term "systems-level outcomes": programmatic consequences linked to student learning, tied to existing structures within the learning ecosystem, and with the capacity to endure beyond intervention delivery. These contrast with an output focus on discrete deliverables, or an outcome orientation towards impacts with a programme-bound time horizon and implementer-bound role scope. Systems-level outcomes necessitate adoption of a process lens; that is, procurement documents should consider the sequenced steps of a relevant change process—such as systems strengthening and teacher training take-up—to determine how partners can apply their expertise so that the system might continue driving learning outcomes into the future.

> " I'm asking myself: how is this document making or seeking a solid argument for why their core idea should lead to enduring systems outcomes that will then engender long-term learning outcomes? I'm looking for [systems and process] evidence—the approach's success in similar contexts, the organisation's track record, clearly identified systems levers, et cetera. "

Grounding in localised knowledge and expertise

9 To what extent is this intervention based on **localised**, **contextual knowledge** and **expertise**?

→ Does this intervention build in **time for (re)scoping with participants**, before implementing activities?

→ Are there multiple opportunities for **government and civil society constituents to question and challenge** programme logic?

→ Does programme knowledge emerge from **dialogic design practices**, such as community steering committees or forums?

→ Does this intervention **emerge from contextual realities**, rather than beginning from desired programming components?

Successful TORs explicitly ground programming in the local context by prioritising regional expertise. Professionals assessed the extent to which documented analysis emerged from, or sought to tap, local sources—from community leaders to long-term implementing partners. Red flags, here, involve proposals framed in terms of institutional priorities or extant practices, without logical links to community context. Centrally, then, interviewees sought explicit language around dialogic and participatory design processes, such as community steering committees or interactive systems mapping. In line with an iterative mentality, these activities also require protracted inception periods and ongoing opportunities for local reflection.

> " I'm asking myself, quite simply: how is this demonstrating and searching for experience in the local educational ecosystem? Is there space for the implementer to bring their contextual expertise? "

Tying measurement and evaluation to internal programme logic

10 To what extent are **measurement and evaluation** criteria justified by **core programme logic**?

→ Do the **resources** offered, including funding and facilitation, **reasonably enable the proposed timeline, scope, and scale** of activities?

→ Are **deliverables** considered in terms of the **time and resources required to execute each**?

→ Is each **deliverable** logically linked to the **long-term change strategy**, building a sense of momentum toward long term change?

→ Do reporting requirements **complement or build on existing deliverable activities**?

→ Are the **types of evidence** sought **flexible and contextually relevant**?

> TOR language should logically tie problem statements, activities, and outcomes to relevant programme evaluation criteria. Pragmatically, this refers to a read of ambition and scope, where pie-in-the-sky documents commit partners to unattainable goals. Experts consequently conducted a simple analysis of the resources—including time and financing—available per deliverable. They further assessed whether monitoring and evaluation procedures aligned with both existing programme activities and contextual demands. First, they assessed whether deliverables and metrics meshed with programming—or whether they required reduplicated or unnecessary efforts. They additionally assessed whether the adopted metrics were fit for purpose; whether a randomised control trial was too stringent, for example, or if ongoing testimonials might overburden participants. Easy red flags in this category include a focus on outputs rather than outcomes—or a focus on many tiny deliverables, each of which require unique reporting mechanisms.
>
> " I'm wondering if this document is thoughtful about its deliverables. Are they as meaningful as possible for both funders and implementers? Which make the most difference, and which are just busywork? "

Conclusion

This research presents tools that can help development professionals to assess and refine TOR potential to support scalable learning interventions in developing contexts. The framework and checklist aim to facilitate practitioner reflection, mirroring experts' cognitive processes during TOR review.

Of course, these tools cannot capture the full breadth of procurement best practice. This is, first, a matter of institutional capacity and priorities. Simply put, no TOR can or should aspire to excel across all dimensions. Some interventions, for example, may aim—and receive funding—to map learning needs and pedagogical barriers across an entire nation; others may simply complement existing activities through targeted training in a specific locality. Though the reflections captured by this analysis focus specifically on scalable systems interventions for basic learning in developing contexts, procurement focuses will necessarily vary according to diverse demands.

Secondly, these tools are limited by our research methods. Despite the many benefits of cognitive interviewing, this technique does not eliminate fundamental qualitative data collection challenges, such as social desirability bias.[13] Further, although we interviewed only experts with a demonstrable track record of programmatic success, participating practitioners represent a small subset of the wider ecosystem. Education development practitioners adopt a great variety of approaches, based on institutional priorities and possibilities. For example, some foundations enjoy adaptive, unrestricted funding, allowing them to spend months nimbly cultivating partnerships on the ground; meanwhile, government agencies may be constrained by bureaucratic processes for transparency and fiscal responsibility.

---

[13] Social desirability bias refers to the tendency of participants to respond in ways they believe the interviewer or future audiences might favor. For further discussion, see Grim (2011).

Nonetheless, the framework and checklist prove useful tools for inspiring powerful procurement practices globally. They can serve as a reflective guide—rather than a prescriptive package—for practitioners aiming to source systems programming that can improve learning at scale. Ultimately, they push professionals to ask additional questions, drive deeper discussion, and more intentionally reflect on the ways in which their interventions can systemically tackle the learning crisis for youth across the globe.

References

Barton, K. C. 2015. Elicitation Techniques: Getting People to Talk About Ideas They Don't Usually Talk About. *Theory & Research in Social Education*, 43(2), 179–205. https://doi.org/10.1080/00933104.2015.1034392

Brown, A. D., Colville, I., and Pye, A. 2015. Making Sense of Sensemaking in Organization Studies. *Organization Studies*, 36(2), 265–277. https://doi.org/10.1177/0170840614559259

Clarke, V., Braun, V., & Hayfield, N. 2015. Thematic analysis. In Qualitative Psychology: A Practical Guide to Research Methods (p. 248).

Fisher, R. P. 1992. Memory-Enhancing Techniques for Investigative Interviewing: The Cognitive Interview. Springfield, Illinois: Charles C Thomas Publisher, 1992.

Gioia, D. A., and Chittipeddi, K. 1991. Sensemaking and Sensegiving in Strategic Change Initiation. *Strategic Management Journal*, 12(6), 433–448.

Global Education Monitoring Report Team. 2021. Global Education Monitoring Report, 2021/2: Non-State Actors in Education: Who Chooses? Who Loses? (Second edition; p. 573). UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000379875

Grimm, P. 2011. Social Desirability Bias. In J. N. Sheth and N. K. Malhotra, Wiley International Encyclopedia of Marketing. Chichester : Wiley, 2011.

Hargreaves, A., Lieberman, A., Fullan, M., and Hopkins, D. (Eds.). 2009. Second International Handbook of Educational Change. Springer Netherlands. https://doi.org/10.1007/978-90-481-2660-6

Ivery, J. M. 2007. Organizational Ecology: A Theoretical Framework for Examining Collaborative Partnerships. Administration in Social Work, 31(4), 7–19. https://doi.org/10.1300/J147v31n04_03

Our 10 Big Hopes for Change: A Review of the Latest Insights and Pioneers Making Impact on the Ground. 2010. Big Change. https://www.big-change.org/wp-content/uploads/2021/07/Horizon-Scanning-FINAL-JUL2021-1.pdf

Reddy, K., and Narain, S. 2022. Pathways to Transforming Education: Proven Solutions from Social Entrepreneurs. Catalyst 2030 Secretariat. https://catalyst2030.net/wp-content/uploads/Pathways-to-Transforming-Education-Catalyst-2030.pdf

The State of Global Learning Poverty: 2022 Update. 2022. World Bank, UNESCO, UNICEF, USAID, FCDO, Bill & Melinda Gates Foundation. https://www.unicef.org/media/122921/file/State%20of%20Learning%20Poverty%202022.pdf

UNESCO Institute for Statistics. 2022. Setting Commitments: National SDG 4 Benchmarks to Transform Education 2022. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000382076

van Braak, M., de Groot, E., Veen, M., Welink, L., and Giroldi, E. 2018. Eliciting Tacit Knowledge: The Potential of a Reflective Approach to Video-Stimulated Interviewing. *Perspectives on Medical Education*, 7(6), 386–393. https://doi.org/10.1007/s40037-018-0487-9

Winthrop, R., Barton, A., and McGivney, E. 2018. Leapfrogging Inequality: Remaking Education to Help Young People Thrive. Brookings Institution Press.

World Bank. 2018. Development Report 2018: Learning to Realize Education's Promise. World Development Report. Washington, D.C., Boston, Massachusetts: World Bank, Credo Reference.

World Bank. 2019. Ending Learning Poverty: What Will It Take? World Bank, Washington, DC. https://openknowledge.worldbank.org/handle/10986/32553

Acknowledgements

Adam Barton is an implementation and change scientist studying how to make education innovations work in practice. As a consultant, Adam regularly advises global organisations and policy leaders on designing, aligning, and implementing sustainable policies and programmes.

He is currently a Cambridge International Scholar and Stamps Fellow at the University of Cambridge. His work focuses on the psychological and social dimensions of educational change: the shifting mindsets, attitudes, and pedagogical behaviours of students, teachers, and parents.

Adam most recently served as a lead researcher at the Center for Universal Education at the Brookings Institution. There, he directed multi-million-dollar research projects on community beliefs and family engagement, including a mixed-methods study of nearly 30,000 family surveys, 100 engagement case studies, and artificial intelligence topic analytics. His publications include Leapfrogging Inequality, a co-authored book on education innovations.

Other recent work includes advising the Delhi ministry on reform implementation, directing systems change research for OECD network partners, and leading research on co-creative innovation as a visiting fellow and Luce Scholar at the Asia Pacific Initiative—a Tokyo think tank.

Citation:

Barton, A. 2023. The Secret Sauce of Development Professionals: Tools for Assessing TOR Potential to Source Scalable Learning Interventions. 2023/054. https://doi.org/10.35489/BSG-RISE-RI_2023/054

Please contact info@riseprogramme.org for additional information, or visit www.riseprogramme.org.

RISE is funded by: