Co-Designing for a Healthy Edtech Ecosystem: Lessons from the Tulna Research-Practice Partnership in India

Aastha PATEL[a*], Chandan DASGUPTA[a], Sahana MURTHY[a] & Rashi DHANANI[b]
[a]*Indian Institute of Technology, Bombay, India*
[b]*Central Square Foundation, India*
*aasthapatel@iitb.ac.in

Abstract: The demand and supply of EdTech products have surged in the past decade and especially post-Covid19. Yet key challenges exist, such as inadequate quality standards and lack of reliable product evaluations. Consequently, stakeholders in the EdTech ecosystem such as schools, teachers, parents, governments, philanthropists, and investors face difficulty in making informed adoption decisions and feel the need for a systematic quality evaluation framework of educational technology (EdTech) products. In such a scenario, we analyze a Research-Practice Partnership between educational researchers, government decision-makers, and a non-governmental organization working on policy and strategy, who collaborated on designing and implementing EdTech quality standards. We examine the co-design process of the 'Tulna EdTech evaluation index' at various stages in the partnership. We adopted the Design-based Implementation Research (DBIR) approach for guiding our partnership. We examine what can be learned from the process of co-design in our RPP that might be useful for our ongoing partnership going forward, as well as for other RPPs. We found that the stakeholders navigated through tensions and iteratively negotiated the design of the evaluation index based on their individual expectations, perspectives, and expertise. Our retrospective, qualitative analysis supports understanding of how researchers and practitioners might engage in co-design and the role co-design might play in establishing a healthy EdTech ecosystem.

Keywords: EdTech in India, Design-based Implementation Research (DBIR), EdTech quality standards, Co-Design, Research-Practice Partnership

1. Introduction

Today's educational settings all over the world have seen the influx of several hundreds of educational technology (EdTech) products. The EdTech market is growing rapidly, even in developing countries. The demand for EdTech is also high and is driven by governments, private organizations, schools as well as individual decision-makers such as parents. The variety of EdTech based solutions is enormous in terms of the intended goals, target audience, technologies, pedagogical design, cost, use case, and so on. At the same time stakeholders in the EdTech ecosystem face key challenges in understanding what constitutes good quality EdTech. Research studies have given rise to recommendations for the design and implementation of EdTech solutions, for instance on addressing goals such as personalization through appropriate pedagogical strategies and technology affordances (Peng, 2019), or on scaffolding students for various learning needs (Quintana et al., 2004). There exist a few evaluation frameworks and instruments (LORI; Nesbit & Leacock, 2007; LOESS; Kay & Knaack, 2009) but are generic across different types of EdTech products, may not have been tested for validity and reliability, and often have not been implemented with products on the market.

In developing countries, this problem of practice is further exacerbated due to inadequate quality standards that take into account the needs arising from the local EdTech ecosystem, and a lack of unbiased reliable product evaluations. In economic terms, this information asymmetry leads to ad hoc and inefficient decision making for adoption, without regard to the impact on learning (Ferreyra, M., & Liang, P., 2012) In such scenario, stakeholders such as schools, teachers, parents, governments, philanthropists and investors feel the need for a systematic quality evaluation framework of EdTech products to make informed decisions.

The problem of lack of EdTech quality standards is systemic in nature involving multiple stakeholders in their individual roles and collective roles. To address this issue, a solution needs to be designed that actively engages all the stakeholders and entities involved who need to not only align but also collaborate. Such a collaborative, systemic effort to design and implement quality standards contribute towards creating a healthy EdTech ecosystem (Omidyar Network, 2019). For a healthy ecosystem to emerge and sustain, the solution needs to integrate the requirements and practices of the various stakeholders. For example, the solution should include research-based quality standards that are robust as well as usable. It should encourage demand for evidence-based recommendations and also garner acceptance from the product supply perspective.

One model of such integration is a co-design process within a Research-Practice Partnership (RPP). RPPs enable the use of research evidence in decision-making for policymakers and funders (Tseng, 2012). In this paper, we describe an RPP formed as part of the *EdTech Tulna* initiative in India (Tulna, 2021). The partnership is between an EdTech research group from a research university, a non-governmental organization working on policy and strategy, and a government agency involved in EdTech procurement decisions. A design-based implementation research (DBIR) approach (Fishman et al., 2013) guided us in the co-design of the 'Tulna EdTech evaluation index' at various stages in the partnership. We unpack the co-design process from the lens of RPP and DBIR and present a case study involving an analysis of three co-design episodes. The lessons learned from this RPP illustrate how researchers and practitioners might engage in co-design and work towards establishing a healthy EdTech ecosystem in India.

2. Background and Context

2.1 EdTech Ecosystem in India

The past two decades have seen a strong emergence of technology-enhanced learning in India in all sectors. Large-scale initiatives by the Government of India such as the National Mission on Education through ICT (NMEICT, 2021) focused on access, quality and equity through providing connectivity and devices as well as on content generation. In addition, traditional academic publishing houses reoriented towards a technology division, adding repositories of ICT-based content. Non-governmental organizations (such as Central Square Foundation (CSF, 2021)) and international foundations have devoted tremendous efforts towards improving the school education system, especially in low-income communities with an emphasis on technology-supported solutions for scaling and access. Investors (for example, Omidyar Network (2019)) look to EdTech based solutions to address the problem of quality education. In addition, in recent years, the market has witnessed the influx of several hundred EdTech products from for-profit product companies as well as not-for-profit organizations funded by philanthropists and foundations. The National Educational Policy (NEP, 2020) emphasizes technology use and integration by establishing a 'National Educational Technology Forum' that will advise the leadership of educational institutions and facilitate decision-making based on induction, deployment, and use of EdTech.

Within such a dynamic scenario, adoption decisions are a key challenge. Government decision-makers are not equipped to differentiate between various types of EdTech and to examine their quality from different perspectives. It is also time-consuming for government panels to conduct proof of concept evaluations. Thus what typically gets emphasized is infrastructure, cost, or convenience, and what gets left out are parameters related to learning and teaching. Reputed product companies value unbiased and credible evaluations both as a pathway to adoption as well as a formative feedback mechanism to improve the product. A prior study of 12 EdTech companies in India has shown the need for public evaluations and a systematic approach to evaluate a large number of products on a regular basis (CSF EdTech Lab 1 Report, 2019). Individual decision-makers such as parents are faced with an array of products. They have to navigate these choices and make financial decisions mostly on the basis of word-of-mouth recommendations, while at the same time facing strong advertising from vendors.

In mature EdTech markets, various models of evaluation and quality frameworks exist (for example EdReports (EdReports, 2021)). Such frameworks, standards, or evaluations that are relevant to the local context are not yet part of the Indian EdTech ecosystem.

2.2 EdTech Tulna Initiative and the Partners in The RPP

The *EdTech Tulna* initiative (Tulna, 2021) in India has been created as a public good to address the challenges of information asymmetry around EdTech quality. *EdTech Tulna* consists of a research-based framework to set quality standards and a corresponding product evaluation index. The index is applied to evaluate existing products in the ecosystem and the evaluations are intended for public use. The aim is to promote demand for evidence-based decision-making and encourage the supply of interventions to meet the demanded quality standards. EdTech Tulna index focuses on evaluating the design of EdTech products along three constructs: Content Quality, Pedagogical Alignment, and Technology & Design to capture a holistic view of the quality of the product design. Each construct further comprises multiple criteria.

EdTech Tulna is created out of a partnership in India between the Educational Technology department at IIT Bombay, a premier research institute in India, and the Central Square Foundation, a non-governmental organization focusing on educational policy and strategy. The RPP discussed in this paper has another partner, a state government that is involved in making decisions for EdTech adoption at scale. The state government, along with a team of consultants that facilitate product procurement and adoption are considered as the practitioners in the RPP. This triadic partnership model brings a wide range of knowledge, skills, and practices that are essential for addressing the EdTech quality problem. Fig. 1 depicts the areas of expertise and the roles of individual partners in this RPP as well as those emerging at the collaboration interface.

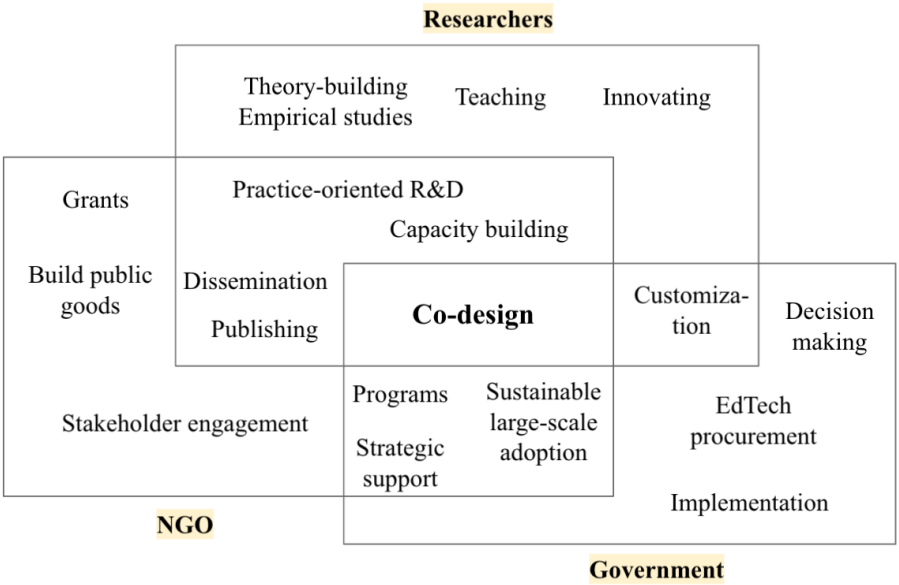

Figure 1. Co-design emerging from the Triadic Partnership between Researchers, NGO and Government.

The researchers have an academic background and engage in research involving theory-building and empirical studies. A key area of practice for academic researchers is teaching and training students in disciplinary knowledge and skills. Through these activities, academic researchers are engaged in discovering new knowledge and innovating solutions. For the NGO, the impact areas of focus are foundational learning, bringing innovation in education through technology-based solutions, and improving learning outcomes at scale for low income children. Towards these, they engage with other stakeholders in the ecosystem to build public goods, run programs and create research-based solutions. They also support organizations through grants to drive sustainable impact in these areas. The government's role in this context includes making decisions regarding EdTech adoption for large numbers of schools within its jurisdiction and implementing the processes for procurement. The collaboration between the researchers and the NGO has given rise to practice-oriented research and development projects with an emphasis on data collected from field implementations. It has also led to capacity-building efforts and the publication of research reports and guidelines for various stakeholders. For this project, the NGO helped the researchers better understand the context for the design. The NGO

collaborates with the government and the consultants in providing strategic and technical support along with the management and implementation of programs, with a focus on scaling and sustainability. The collaboration interface between the researchers and the government consultants in this RPP was a novel exercise in the customization of a research tool to address the requirements of the government's policies and procedures. What emerged at the heart of this three-way partnership was an effort in co-designing a solution for EdTech quality evaluation that drew from the diverse expertise and addressed diverse needs. This partnership includes educators who have experience working in educational institutions but we have not formally engaged with schools as partners yet due to logistical challenges.

3. Theoretical framework

Research practice partnerships (RPPs) are long-term collaborations between practitioners and researchers that are organized to investigate problems of practice and solutions for improving system outcomes. There are multiple types of RPPs such as Research Alliances, Network Improvement Communities, and Design Research Partnerships (Coburn, Penuel, & Geil, 2013). Our partnership is categorized as a 'Design research partnership' as it focuses on informing practice and research by designing, developing, and testing an educational innovation - EdTech Tulna - an EdTech Evaluation Index. Design Research partnerships typically have the following characteristics. These are place-based, have dual goals of informing practice and research, emphasize co-design, and rely on close collaboration at every stage in the process (Coburn et al., 2013). The principle of co-design is central to this type of partnership. By co-design we refer to "a highly facilitated process that engages people with diverse expertise (e.g., research, curriculum, professional development, teaching) in designing, developing, and testing innovations." (Coburn et al., 2013, p.10).

Design research partnerships are often formed using the Design-Based Implementation Research approach (DBIR) (Fishman et al., 2013) since this approach caters to both design and implementation of innovative interventions involving multiple stakeholders. We use DBIR as a framework to analyze the co-design process. The core principles that characterize DBIR are:

- *Principle 1*: Teams form around a focus on persistent problems of practice from multiple stakeholders' perspectives.
- *Principle 2*: To improve practice, teams commit to iterative, collaborative design.
- *Principle 3*: To promote quality in the research and development process, theory guides both learning and implementation through systematic inquiry.
- *Principle 4*: DBIR is concerned with developing capacity for sustaining change in systems.

These principles guide our partnership and help make our co-design process more reliable in terms of the outcomes (Penuel, 2019).

4. Method

We draw anecdotes from the ongoing Tulna RPP focused on establishing EdTech Quality standards in India. The data were collected over a period of 8 months and documented systematically in the form of meeting notes, documents, email communication, versions of designed artifacts, and participant observation. Meetings were conducted twice a week at various levels such as internal research meetings, partner meetings, and meetings with external stakeholders. Due to covid, all meetings happened virtually and collaboration was completely remote. We analyzed notes and transcripts from the meetings. In the first round of analysis, we identified the *episodes of co-design* where multiple stakeholders were involved in creating something emergent. Three main episodes of co-design emerged from our analysis and they were selected based on their chronological occurrence and stakeholders involved in co-design. One episode each was selected from the three phases: (i) Project Scoping (ii) Development of the Evaluation index and (iii) Customisation of the index. In the second round of our analysis, we analyzed each episode using the core principles of DBIR. The analysis was conducted by the research team, the authors of the study were the active members in the co-design process, and presented the first-person account of co-design (Coburn & Penuel, 2016). In the following section, we

first describe the three episodes, highlighting what each stakeholder brought for co-design and how it led to key emergent developments in the partnership.

5. Case Study: Co-Design between Educational Researchers, Non-Governmental Organization, and Government Consultants

5.1 Co-Design Episode 1: Systematic Layered Approach to EdTech Evaluation

Aligned to the first guiding principle of DBIR, the Tulna partnership between researchers and the policy NGO brought in multiple stakeholders' perspectives to solve the pressing problem of the lack of quality standards in EdTech. The partners aligned on the broad problem of practice and further carefully identified root causes, key change drivers to develop a long-term vision, and practical theory of action. At this stage, partnership often gets challenging as researchers and practitioners work at different scales. Researchers start by investigating small discrete elements in controlled settings with few users whereas practitioners are likely to use those measures to investigate the problem at scale in-situ with many users (Tseng, 2012). In this case, the research team preferred to focus on the fundamental design of EdTech products and defining quality standards at a product level, since research indicates that a product can succeed at scale only if it is designed well. On the other hand, the NGO's emphasis was on evaluating the product at scale, taking into consideration infrastructural issues that may affect the use of the product on the ground. Both these perspectives are valid in the case of evaluating EdTech products. However, negotiating these multiple perspectives held by the two stakeholders was important in this partnership in order to make progress. Both the research team and NGO worked together to determine the scope of the project that might allow both perspectives to co-exist. Numerous meetings were held between these two stakeholders to jointly design what this evaluation approach might look like. Both the stakeholders were also trying to negotiate the timeline by which the evaluations would be done. Research suggests that there is a difference in researcher's and practitioner's timelines of work, Practitioners often work under stringent timelines for immediate implementations of innovations, whereas researchers proceed slowly with relatively flexible timelines through cycles of inquiry and analysis before they are ready for recommending action (Penuel & Allen, 2015). The dual negotiation points of what to evaluate and when to evaluate informed the joint decisions taken by the stakeholders.

A layered approach to quality evaluation (Fig. 2) emerged from these joint discussions. This approach enabled multiple perspectives to co-exist and presented a new roadmap for the partnership that had not existed before. As per this approach, the research team started the evaluation process by focusing on the design of products, i.e. the foundational layer in Fig. 2. The long-term evaluation plan now also consisted of two intermediate layers - evaluating the interaction of learners and teachers with the product (curated user studies) and evaluating learning outcomes in the field setting (in-situ studies). The final stage of the evaluation program comprised evaluation of the outcomes at scale.

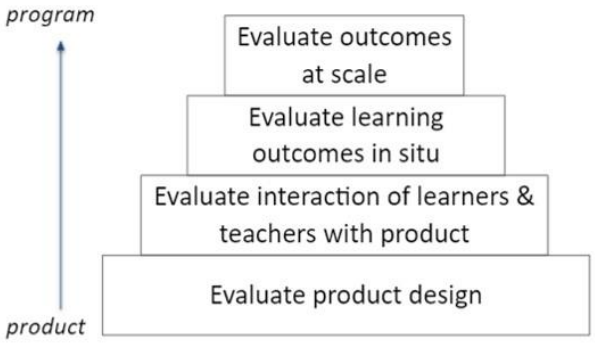

Figure 2. A Systematic Layered Approach to Evaluation.

This approach was accepted by both partners and helped develop a long-term vision for sustainable outcomes. The rationale provided by the researchers of understanding both the outcomes and the underlying mechanisms convinced the practitioners to come on board, and helped build

confidence that all aspects will be addressed but not just immediately. This also helped the partners navigate the different scales and time frames between research and practice (Tseng, 2012).

5.2 Co-Design Episode 2: Re-Defining the 'Contextualization' Construct

This episode is aligned with the third principle of DBIR and describes the co-design process for developing the evaluation index. In the first version of the index, the researchers had included 'contextualization' as one of the constructs for evaluating the EdTech products. Contextualization was defined as a measure of support provided to the teachers and learners to enhance its potential effectiveness in the users' local context. The need for the idea of contextualization in EdTech products is agreed upon by all stakeholders in the ecosystem. This becomes especially important for India because of its diversity in context across different states, 100+ languages spoken in the country, and the disparity in socio-economic classes. However, discussions with the NGO soon made it clear that different stakeholders interpreted contextualization differently. The NGO conveyed that 'contextualization' is a commonly used terminology in large-scale implementation programs where practitioners conduct pilot studies for contextualization before scaling programs. The research team argued for including it as an evaluation construct from the pedagogical, content, and interaction point of view where each product's design needed to be contextually relevant for the target teachers and students. While the NGO agreed with this point of view, it considered infrastructural issues and training and support available for teachers and students as more important factors determining the effectiveness of a product in a context. The research team acknowledged that these factors were relevant for evaluating a product for contextualization. Along with the multiple interpretations of contextualization that emerged from the discussions with the NGO, the researchers also realized that it is not possible to evaluate contextualization entirely from the product's design. One needs to conduct user studies (top three layers in Fig 2) to evaluate a product for contextualization - from the pedagogical, content, and interaction perspective as well as infrastructural and training perspective.

After further negotiations, it was agreed that there were some aspects of contextualization that could be evaluated within the product design. The researchers reviewed the literature to identify these which were then absorbed under the other three constructs of the Tulna index - Content quality, Pedagogical alignment, and Technology & design. Thus, these three constructs were co-designed to now cater to a broader scope that covered contextualization. Aligning to the third principle of DBIR, both partners argued for the extent to which contextualization should be included in the index and the systematic layered approach to evaluation (Figure 2) gave them the flexibility to push it to the upper layers of evaluation i.e., later on in the program. This high-level co-design between the partners evolved the index to be more robust and univocal.

5.3 Co-Design Episode 3: Quantification of the Index

This co-design episode aligns with the second and fourth DBIR principles. Once the Tulna index was in the advanced stages of development, the NGO facilitated interactions with the state government through a group of government consultants. This was the first instance of the adoption of the index by a state government. The Tulna index was considered as a comprehensive tool to evaluate the quality of EdTech products for the state's large-scale EdTech adoption for 2000+ schools. In the initial phases, the government consultants conveyed the criteria that they considered important for the evaluation of EdTech products in their setting. The research team took these criteria into consideration while designing and validating the index. However, when this adoption process began, all the three partners realized that the government needed an index that i) enabled rapid decision-making, and ii) was capable of providing numerical scores. Several rounds of discussions between the partners highlighted a tension between research rigor captured in the index and the need for ease of use by government officials. The original version of the index was designed to be used by trained evaluators. All partners realized that the index was not suitable as is to be used by government officials and in order to make it usable they had to customize it.

Aligning with the second guiding principle of DBIR the researchers and the government consultants collaboratively adapted the index to a more usable and contextualized format for this specific state government. This process of partnership went beyond the usual division of labor where

the researcher would design the index, while practitioners would test and implement it (Coburn & Penuel, 2016). The government consultants, as well as the NGO, played a vital role in bringing out the need for quantification and suggested the researchers assign numerical values to the existing 3-scale descriptive rubric along with a comprehensible numerical target descriptor. For example, they requested to rephrase *"Higher Order Thinking Skills (HOTS) are being sufficiently addressed in the content, examples, activities"* to *"80% or more HOTS are being addressed in the content, examples, activities".* The research team conveyed to the other stakeholders that quantification at this level was not feasible as there was no basis for claiming that a certain percentage of content was sufficient to make a product effective. Further, the government and the NGO preferred an index that represented the evaluation of an EdTech product in the form of one numerical value for easy decision-making, rather than a descriptive qualitative rubric. The researchers understood the need for quantification but emphasized that a qualitative rubric enables bringing out pros and cons in the product's design that would be lost in a quantified number.

Negotiations between the government consultants and the research team resulted in the research team agreeing to assign numerical values in a way that helps the consultants synthesize and make decisions. A quantitative scoring scheme was devised having three nonlinear scales i.e. Exemplary (score 30), Valuable (score 15), and Potential to Improve (score 5). This nonlinear scaling came from a research perspective. Thus, in order to bring structure and objectivity to this complex process, a quantified version of Tulna index was created for this state government. The quantified version of the Evaluation index was designed keeping in mind the practitioners as the primary users of the index. Further, clearly articulated reviewer guidelines for step-by-step guidance for each evaluation criterion were added in this quantified version of the index. Supplementary material was provided to aid the understanding of disciplinary concepts and also help navigate through the index. In order to reduce the cognitive overload of the government officials, important steps were emphasized using black bold text and some tacit guidelines were explicitly mentioned in grey text for anyone who needed guidance at a granular level. Fig. 3 illustrates the characteristics of the quantified version for one criterion.

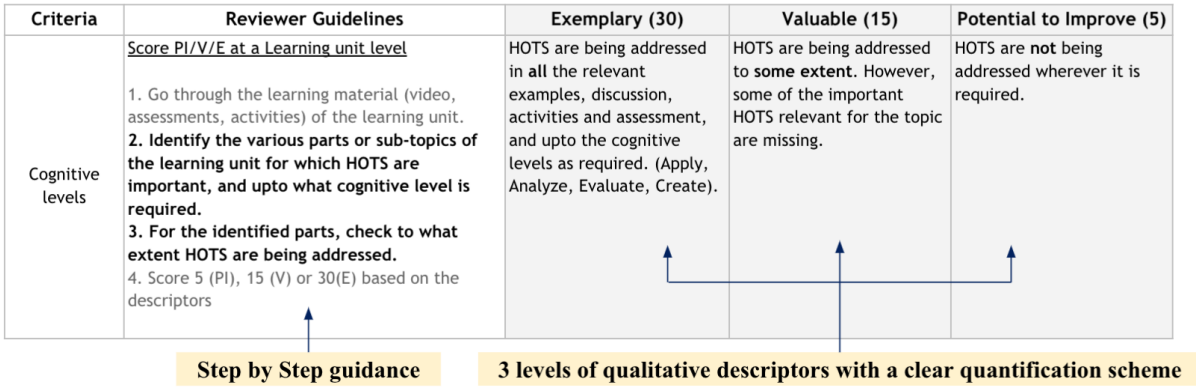

Figure 3. Screenshot from the Quantified Evaluation Index.

At the end of the co-design process, the researchers realized the importance of quantification and ease of use in order to enable the adoption of the Tulna index by multiple state governments. The state government realized the importance of the multi-dimensionality of EdTech evaluations that could not be captured by a single numerical value for the entire product. The outcome of this co-design resulted in the design of a more usable, accessible, and practitioner-centered evaluation index. This new quantified version of the index can be easily carried forward into future practice and decision-making in other contexts. This aligns with the fourth guiding principle of DBIR. The researchers also conducted a short 2-day training workshop with the actual users of the index. They built on this experience to refine and develop a set of training resources that could be used in future collaboration with other state governments.

6. Discussion

In this paper, we have described and analyzed an ongoing design research partnership between a research team, NGO, and government consultants in India working towards defining EdTech quality standards, setting up the process and index for evaluating EdTech products, and enabling research-based decision-making by the stakeholders. The long-term vision shared by all stakeholders is to establish a healthy EdTech ecosystem in India that is characterized by demand for and supply of good quality EdTech products. We used the DBIR approach to describe what this partnership looked like using the three representative episodes. Co-design forms the core of this partnership (Coburn et al. 2013). As highlighted in the literature, this co-design process was place-based i.e. influenced by the state government (represented by the government consultants) that wanted to first adopt the Tulna index and informed the creation of the quantitative version of the index. The process informed research and practice. The research team developed a better understanding of the contextualization construct that was embedded within other evaluation constructs. The NGO and government understood the importance of these evaluation constructs containing contextualization. The state government's evaluation process for procuring EdTech products also changed as a result of this partnership. They used Tulna index's multi-dimensional evaluation criteria for evaluating products' design during the procurement process. Instead of using a single number for describing the quality of the product, they now considered quantitative scoring across the three dimensions of content quality, pedagogical alignment, and technology and design. This partnership has been marked by numerous episodes of close collaboration between the stakeholders that resulted in defining a long-term roadmap for the evaluation program consisting of a tiered approach starting with evaluating a product's design.

Specifically, our findings highlight the following lessons for sustaining co-design between multiple stakeholders aiming to establish a healthy EdTech ecosystem in India.

Lesson 1: Leverage multiple interpretations

Co-design involves numerous instances of multiple interpretations due to differences in the stakeholder's goals and cultural influences defining the way the organization works. In our partnership, we observed multiple perspectives about the desired approach for evaluating EdTech products, contextualization as an evaluation dimension, and the need for quantification of evaluation results. Such episodes became the anchor points for our co-design process. Each stakeholder brings with them a unique perspective and it is essential to enable these perspectives to co-exist from the very beginning of the partnership. In RPPs, it is essential to bring diverse perspectives together in productive ways while maintaining mutualism (Coburn et al., 2013). Both researchers and practitioners must acknowledge that differences are bound to exist but leveraging these differences will likely lead to a collaborative endeavor where the sum is bigger than the combination of the parts contributed by each stakeholder. For instance, in the first co-design episode, multiple interpretations of the approach to evaluation resulted in the creation of a layered approach that presented a new roadmap for the partnership that had not existed before. In the second episode, the stakeholders had different interpretations of contextualization as an evaluation construct. These gave rise to a refined evaluation index where key aspects of this construct were embedded within content quality, pedagogical alignment, and technology and design constructs for evaluating EdTech products. In the third episode, multiple perspectives for the need for quantification resulted in a multi-dimensional quantified version on the evaluation index that served the purpose of the state governments as well as researchers.

Lesson 2: Negotiate tension between research rigor and usability

Research-based educational interventions are guided by multiple disciplinary theories. Tulna index has deep roots in the theories of learning sciences, educational disciplinary knowledge, Human-Computer Interaction, and educational psychology. This may result in the creation of research-based tools that are usable only by researchers or highly trained individuals. This high entry barrier prevents the adoption of these tools by all the stakeholders. In RPPs, the dual goals of improving research rigor and pushing out research to practice (Tseng, 2017) lead to tension between research rigor and usability. Tension develops as researchers are trained to work towards high rigor and the practitioners are concerned about usability, effective adoption, and scaling of various tools. Prior research highlights one such tension between research rigor and timelines (Coburn et al. 2013). Rather than disrupting design, these design tensions can serve as opportunities for productive designs and they represent the need to be addressed in co-design (Severance et al., 2014; Tartan, 2007). Stakeholders need to negotiate this tension between the research rigor and practical usability. While all stakeholders may agree that both rigor and usability

are important to have, the negotiation process should establish the extent to which each can be preserved. For instance in the third episode, while all stakeholders agreed that quantification was needed, quantification to the extent of specifying the percentage of presence of higher-order thinking skills (suggested by government stakeholder) was not possible because it was not supported by research. Instead, a quantitative scoring scheme was devised having three nonlinear scales i.e. Exemplary (score 30), Valuable (score 15), and Potential to Improve (score 5). In addition, detailed reviewer guidelines and evaluation descriptors that can help practitioners assign these numerical scores for decision-making were also designed. Capacity-building training programs were also conducted for the practitioners to help them use this quantified index. Similarly, in the first episode, the layered approach emerged as an outcome of negotiation between the need for a rigorous evaluation of a product's design and the need to evaluate the effectiveness of products at scale. In the second episode, the scope of the contextualization construct was reduced and merged with other constructs to make the evaluation index more usable by practitioners.

Lesson 3: Develop shared expectations

During the co-design process, stakeholders are likely to have a different understanding of each other's as well as their own roles in the partnership. We have outlined some of the unique and shared roles of the three stakeholders engaged in this co-design process in Fig. 1. The shared roles provide opportunities for developing shared expectations in the partnership. Prior research suggests that partnerships need to periodically revisit their shared understanding about their role and synthesize shared expectations and vision around the work that they are engaged in doing (Farrell, Harrison, & Coburn, 2019). The importance of this was reflected in the co-design process observed in our context as well. For instance, in episode 3, significant traction in terms of developing shared expectations was achieved between the researchers and the government agency while engaged in the shared work of customization and fine-tuning of the quantitative version of the Tulna evaluation instrument. Researchers drew on their role of theory-builders and domain experts and the government agency drew on their role involving EdTech product procurement to define the shared work around customization and quantification during the co-design process. The research team had begun with the expectation that only trained experts will be the evaluators using the Tulna index. However, with the new shared expectation of state agencies doing the actual evaluation of EdTech products quickly, the customization process gave rise to new tools and resources. Similarly, in Episode 2, shared expectations developed around the feasibility of data collection and evaluation using the contextualization construct. This process enabled the NGO to draw on its primary mandate and experience with stakeholder engagement and researchers to draw on their knowledge of theories of learning to give rise to the co-design process.

7. Conclusion

In this paper, we began with the objective of unpacking the co-design process and sharing lessons learned from an ongoing RPP to illustrate how researchers and practitioners might engage in co-design for establishing a healthy EdTech ecosystem in India. In the absence of examples of such partnerships in India, we had to create one from scratch. We have highlighted three lessons that emerged from this partnership, namely - leverage multiple interpretations, negotiate the tension between research rigor and usability, and develop shared expectations. These lessons are likely to help future stakeholders to align with each other as they come onboard. Furthermore, we hope to build on and refine these emerging lessons with the help of interactions between all the existing and new stakeholders. While these lessons emerged in the EdTech research-practice partnership, we believe that these will hold true for other RPPs in India as well which share similar stakeholder characteristics. Finally, given the goal of establishing a healthy EdTech ecosystem in India, we believe that these lessons will initiate discussions amongst the policymakers and national education strategists on how to sustain and institutionalize this ongoing work. Towards this long-term goal, the above co-design episodes and lessons demonstrate the importance of having multiple stakeholders aligned around one shared aim who are generating and sharing knowledge with each other. We hope that this work will inspire innovative ideas driven by policymakers and strategists for sustaining such RPPs where the voices of all stakeholders are represented.

Acknowledgments

We acknowledge the collective effort of the EdTech Tulna team from IIT Bombay, the Central Square Foundation, and the International Innovation Corps team working with the state government officials in India.

References

Coburn, C.E. & Penuel, W.R. (2016). Research-practice partnerships: Outcomes, dynamics, and open questions. *Educational Researcher, 45* (1), 48-54.

Coburn, C.E., Penuel, W.R., & Geil, K. (2013). Research-practice partnerships at the district level: A new strategy for leveraging research for educational improvement. *William T. Grant Foundation.*

Central Square Foundation - CSF (2021). www.centralsquarefoundation.org, retrieved 25 May 2021.

CSF EdTech Lab 1 Report (2019). Retrieved on 25 May, 2021 from https://centralsquarefoundation.org/wp-content/uploads/EdTech%20Lab%20Report_November%202019.pdf

EdReports. (2021). Retrieved 25 May 2021, from https://www.edreports.org/

Farrell, C. C., Harrison, C., & Coburn, C. E. (2019). "What the hell is this, and who the hell are you?" Role and identity negotiation in research-practice partnerships. AERA Open, 5(2), 2332858419849595.

Ferreyra, M., & Liang, P. (2012). Information asymmetry and equilibrium monitoring in education. *Journal of Public Economics*, *96*, 237–254.

Fishman, B.J., Penuel, W.R., Allen, A.-R., Cheng, B.H., Sabelli, N. (2013). Design-based imple- mentation research: An emerging model for transforming the relationship of research and practice. In B. J. Fishman, W. R. Penuel, A.-R. Allen, & B. H. Cheng (Eds.), *Design-based implementation research. National Society for the Study of Education Yearbook, 112*(2), 136–156.

Kay R. & Knaack L. (2009): Assessing Learning, Quality, and Engagement in Learning Objects: The Learning Object Evaluation Scale for Students (LOES-S). Edu. Tech. Research and Development, 57(2), 147-168

Leacock, T. L. & Nesbit, J. C. (2007): A framework for evaluating the quality of multimedia learning resources. Journal of Educational Technology and Society, 10(2), 44.

National Education Policy (2020). https://www.mhrd.gov.in/sites/upload_files/mhrd/files/nep/NEP_Final_English.pdf

NMEICT(2021). Retrieved May 31, 2021 from https://nmeict.ac.in/technology-enabled-learning/national-mission-in-education-through-ict/

Omidyar Network. (2019). Scaling access and impact: Realizing the power of EdTech. Omidyar Network. https://omidyar.com/scaling-access-impact-realizing-the-power-of-edtech/

Peng, H., Ma, S., & Spector, J. (2019). Personalized Adaptive Learning: An Emerging Pedagogical Approach Enabled by a Smart Learning Environment. In *Lecture Notes in Educational Technology* (pp. 171–176).

Penuel, W., Allen, A.-R., Coburn, C., & Farrell, C. (2015). Conceptualizing Research–Practice Partnerships as Joint Work at Boundaries. *Journal of Education for Students Placed at Risk (JESPAR)*, *20*, 182–197.

Penuel, W. (2019). Co-design as Infrastructuring with Attention to Power: Building Collective Capacity for Equitable Teaching and Learning Through Design-Based Implementation Research (pp. 387–401).

Quintana, C., Reiser, B., Davis, E., Krajcik, J., Fretz, E., Duncan, R., … Soloway, E. (2004). A Scaffolding Design Framework for Software to Support Science Inquiry. *J. of the Learning Sciences*, *13*, 337–386.

Severance, S., Leary, H., & Johnson, R. (2014). Tensions in a Multi-Tiered Research-Practice Partnership. In *Proceedings of International Conference of the Learning Sciences, ICLS* (Vol. 2).

Tatar, D. (2007). The Design tensions framework. Human-Computer Interaction, 22, 413-451

Tseng, V. (2012). Partnerships: Shifting the dynamics between research and practice. New York, NY: William T. Grant Foundation.

Tseng, V., Easton, J., & Supplee, L. (2017). Research-Practice Partnerships: Building Two-Way Streets of Engagement. *Social Policy Report, 30*, 1–17.

Tulna (2021). https://edtechtulna.org/, retrieved 25 May 2021.