



**GLOBAL
EDTECH
TESTBED
NETWORK**

Towards Systemic EdTech Testbeds: A Global Perspective

Effectiveness, Efficacy, and Learning Outcomes

Stefanie Vanbecelaere¹

Taskeen Adam²

Carmen Sieber³

Alison Clark-Wilson⁴

Katie Boody Adorno⁵

Björn Haßler²

¹ Itec, imec research group at KU Leuven, Belgium

² Open Development & Education (UK and Sierra Leone)

³ Swiss EdTech Collider, Switzerland

⁴ UCL IOE Faculty of Education and Society, University College London, UK

⁵ Leanlab Education, USA

January 2023

DOI: <https://doi.org/10.53832/opendeved.0285>

Available under [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/).

Recommended citation	<p>Vanbecelaere, S., Adam, T., Sieber, C., Clark–Wilson, A., Boody Adorno, K., & Haßler, B. (2023). <i>Towards Systemic EdTech Testbeds: A Global Perspective</i>. https://doi.org/10.53832/opendeved.0285. Available from https://globaledtech.org/. Available under Creative Commons Attribution 4.0 International.</p>
Licence	<p>Creative Commons Attribution 4.0 International https://creativecommons.org/licenses/by/4.0/</p> <p>You – dear readers – are free to share (copy and redistribute the material in any medium or format) and adapt (remix, transform, and build upon the material) for any purpose, even commercially. You must give appropriate credit, provide a link to the licence, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.</p>
About GETN	<p>The Global Edtech Testbed Network is a collaborative effort supported by the Jacobs Foundation to advance best practices in the field of EdTech co-development and evaluation through “<i>testbeds</i>” or authentic school environments in which to trial emerging education technologies.</p>

Contents

Chapter 1. Introduction	9
1.1. Terms and definitions	11
1.2. A simple description of an EdTech Testbed	14
1.3. A guide to this report	15
Part 1	
Chapter 2. Unpacking EdTech Testbeds	17
Chapter 3. Pre-conditions for effective EdTech Testbeds	19
3.1. Sustainable funding in place	19
3.2. Supportive education and economic policies	20
3.3. Supportive school environment	20
3.4. Key stakeholders are involved and motivated	21
3.5. Access to the relevant educational and technological expertise	22
3.6. Access to EdTech evidence sources	23
Chapter 4. Key components of EdTech Testbeds: the EdTech, the Test and the Bed	24
4.1. The EdTech component	24
Educational purpose or goal	25
The type of technology	25
Target group	26
Educational content	27
Pedagogy	27
Stage of Innovation (or Development)	28
Selection criteria	30
4.2. The “Test”	31
What to test?	32
Methodologies and methods	33
4.3. Spotlight 1. The EdTech Hub Sandbox Approach	35
Timeline and frequency of use	36
Procedures	36
Leadership: who does the actual testing?	38
Reporting of results	39
4.4. The “Bed”	41
Context	41

Testers	41
Spaces for learning	42
Scale	43
4.5. Spotlight 2: Swiss National EdTech Testbed Program	44
EdTech	44
Bed	45
Test	45
4.6. Spotlight 3: Impulsionar (Brazil)	46
General information	46
EdTech	46
Bed	48
Test	48
4.7. Spotlight 4: Leanlab Education Codesign Product Research	50
General information	50
EdTech	50
Bed	50
Test	51
Chapter 5. Impact goals	52
5.1. Improve learning outcomes and learning experiences	52
5.2. Produce financially viable and scalable EdTech	53
5.3. Develop more effective teaching approaches	53
5.4. Produce contextualised and equitable EdTech	54
5.5. Further academic knowledge on EdTech	54
5.6. Increase stakeholder collaboration	55
5.7. Generate (better) evidence for specific EdTech products	55
Part 2	
Chapter 6. Lessons learned from more system-wide EdTech Testbeds	56
6.1. Lessons learned from the Swiss EdTech Collider	56
Support is important	57
Teachers can be contacted in several ways	57
Real-world feedback is an added value for companies	57
The test phase must be designed individually	57
6.2. Lessons learned from WISE Report (Qatar Testbed)	58
Building relationships with school leaders early	58

Test the EdTech product before actual implementation	58
Identify an “EdTech” champion	58
Involve the EdTech venture in ongoing professional development of teachers	59
6.3. Lessons learned from the Swedish Edtest	59
Blessing from the principal, but passion from the teachers	59
Special support for the school leaders	59
Design Testbeds for a low budget	60
Try to avoid test fatigue	60
6.4. Lessons learned from Leanlab Education	60
Focus on formative research for EdTech tools	61
Ensure the “Bed” and the “EdTech” are prepared to engage in research	61
Don’t neglect data and security protocols	61
Compensate educators and research participants for their work	62
Ensure the bed is diverse, representative, and has buy-in	62

Part 3

Chapter 7. Key challenges for EdTech Testbed design	63
7.1. Low implementation fidelity of EdTech in schools	63
7.2. Difficult to recruit and engage schools, teachers, learners, and parents	64
7.3. Procurement and contract difficulties	64
7.4. Lack of government involvement	65
7.5. Difficult to reach scalable and sustainable EdTech Testbeds	65
7.6. Lack of test procedures	66
7.7. Conflicting goals and roles of different stakeholders	67
7.8. Lack of support after being involved in a Testbed	67
Chapter 8. Conclusion: Towards more effective EdTech Testbeds	68
8.1. Collaboration on evidence creation	68
8.2. Collaboration on EdTech Testbed design	69
8.3. Development of a global EdTech Testbeds network	69
Chapter 9. References	71
Appendix A. Our research approach	82
10.1. Literature search	82
10.2. Stakeholder contributions to the survey	82
Appendix B. Motivating factors for different stakeholders’ involvements in	

EdTech Testbeds	84
Appendix C. EdTech Testbed Methodologies and Methods	86
12.1. Methodology	86
12.1.1. Methodologies from education research	87
Action research	87
Rapid Evaluation Cycles	87
Design-based research (DBR)	88
Design-based Implementation Research	88
Outcome harvesting	89
Sequential Multiple Assignment Randomised Trials (SMARTs)	89
Systemic Mixed-Methods Research	90
Implementation Science	91
12.1.2. Methodologies from tech innovation	91
EdTech Tulna	92
Design thinking for Educators Toolkit	92
Human-centred Design Toolkit	93
Problem Driven Iterative Adaptation (PDIA)	93
National Digital Education Architecture	93
Principles for Digital Development	94
The Sandbox Model	95
Theories of Change and Logic Models.	95
12.2. Methods	96
12.2.1. Methods for brainstorming and information gathering	96
12.2.2. Methods for design, monitoring, and evaluation	97
12.3. Learning more about generating evidence	98
Appendix D. Describing EdTech Testbeds – Some framing questions	99
13.1. What preconditions do the EdTech Testbed satisfy?	99
13.2. Which key components describe the work of the EdTech Testbed?	99
13.2.1. EdTech	100
13.2.2. Test	100
13.2.3. Bed	101

Figures

Figure 1.1. The EdTech Testbed Ecosystem in Europe	10
Figure 1.2. A simple conceptualisation of an EdTech Testbed	14
Figure 2.1. An initial, conceptual framework for EdTech Testbeds	18
Figure 4.1. Innovation spiral as presented in Using Research Evidence: A Practical Guide (↑Nesta, 2016, p. 14)	28
Figure 12.1. Screenshot from https://docs.opendeved.net showing resources for implementation science.	98

Glossary

In this report, we use the following operational definitions of the key terms listed below.

Bed: The physical, online, virtual, or hybrid place where the EdTech Test is conducted.

Caregivers: The main adults who are responsible for children when they are out of school, who are not the children's parents, for example grandparents and foster carers.

Communities: The formal and informal communities that support learners, teachers, parents, and caregivers.

EdTech: the broad range of digital products and services that are developed with the intention of bringing about positive impacts in the education sector and/or enhancing learning.

Methods: *"the techniques or procedures used to gather and analyse data related to some research question or hypothesis"* (↑Crotty, 2020).

Methodologies: *"the strategy, plan of action, process or design lying behind the choice and use of particular methods and linking the choice and use of methods to the desired outcomes"* (↑Crotty, 2020).

Test: The research methodologies developed to evaluate the impacts of EdTech in the real-world setting.

Testbed: The authentic learning environment where the EdTech Test takes place.

Chapter 1. Introduction

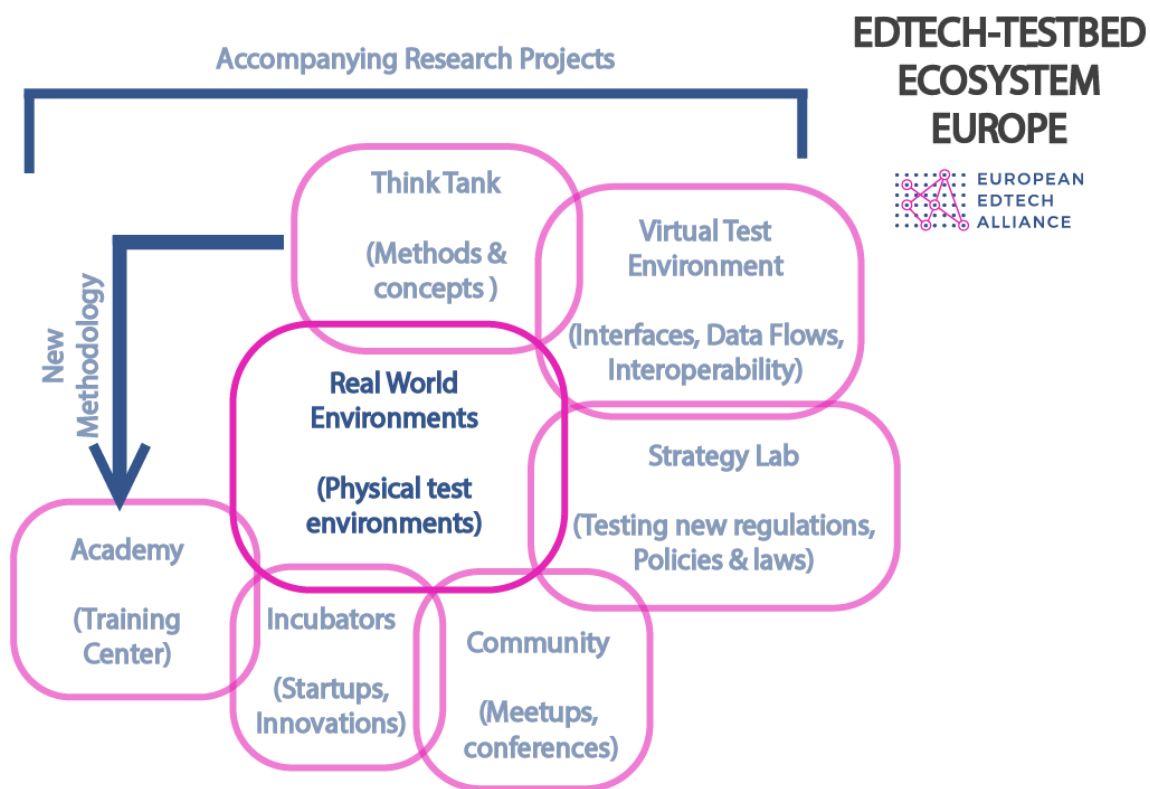
The global growth in the EdTech sector over the past 30 years has resulted in a large array of products and services, from learning management systems to artificial intelligence chatbots (↑[Holon IQ, 2022](#)). The Covid-19 pandemic further reinforced the growth due to a rapid increase in investments in EdTech companies (↑[Holon IQ, 2022](#)). Educational research, however, has shown that the use of various types of EdTech have had mixed results (↑[Escueta and colleagues, 2017](#); ↑[Rodriguez-Segura, 2020](#)). Moreover, ↑[Vegas and colleagues \(2019\)](#) analysed 1640 digital technologies – which is only a snapshot of all education technology innovations – and found that only 11 percent of the EdTech were evaluated externally, while 18 percent were evaluated internally; this highlights that research only captures a small part of what is emerging in the EdTech sector. Therefore, it remains crucial to continue this research and understand which EdTech designs and implementation contexts support more effective use. To realise this, opportunities are required to evidence the educational impact(s) of EdTech and to generate real-world data to improve its design. While the availability of EdTech has expanded rapidly, the ability to evidence its effectiveness has not kept pace.

During the 2010s, the notion of EdTech Testbeds¹ has gained popularity as a way to assess the effectiveness of EdTech (e.g., through co-creation, pilot studies, randomised controlled studies). We appreciate that there are many different roles, purposes, and resources that form the broader concept of EdTech Testbeds. The European EdTech Alliance has written a very helpful blog on this topic, which includes the following diagram ([Figure 1.1. The EdTech Testbed Ecosystem in Europe](#))².

¹In this report, we use the term 'EdTech Testbed'. However, we are aware that this term is a subject of debate in the EdTech community. Others describe EdTech Testbeds as R&D, pilot environments, co-creation communities, sandboxes, etc.

² ↑[Exploring the EdTech Testbed Ecosystem:](#)
<https://www.edtecheurope.org/news/exploring-the-edtech-testbed-ecosystem>

Figure 1.1. *The EdTech Testbed Ecosystem in Europe*



In this report, we focus only on the central component that the European EdTech Alliance refer to as “*real-world environments*” (or “*physical test environments*”), while acknowledging the interrelatedness with other components and the important and critical roles of the other components.

Arguably, EdTech researchers have lacked the resources to consistently test and measure the impact of EdTech, particularly when deployed within an authentic learning environment. Therefore, while new waves of innovations that might have far-reaching implications for increasing global access and educational attainment – such new innovations remain understudied.

The purpose of this report – and, perhaps, the purpose of a thriving future global EdTech Testbed network – is to enable a more sustainable and scaled approach.

The purpose of this report – and, perhaps, the purpose of a thriving future global EdTech Testbed network – is to enable a more sustainable and scaled approach.

1.1. Terms and definitions

For the purposes of this report, we consider EdTech to be the broad range of technology-supported products, services, systems, projects, programmes, processes, or structures that are developed and deployed to bring about positive outcomes in education systems.

EdTech: the broad range of technology-supported products, services, systems, projects, programmes, processes, or structures that are developed and deployed to bring about positive outcomes in education systems.

In this report, we focus on EdTech designed for K-12 learners around the globe, which may also include their teachers, caregivers, and communities. This target group was chosen as governments increasingly require K-12 schools to have access to evidence-based materials. We consider “*K-12 learners*” to be learners of any age acquiring K-12 skills; such learners may be learning a range of subjects, including the usual school subjects, but also subjects falling into the domain of college education, as well as technical and vocational education and training. We do not consider university-based higher education, e.g., leading to Bachelor degrees.

This report aims to capture current thinking and practices about the systematic ways in which we

test “*if*”, “*to what extent*”, “*how*”, “*why*”, “*why not*”, “*in what conditions*”, and “*for whom*” EdTech interventions work

for their intended users within certain settings and places – the **beds**

We go beyond thinking only about the evidence generation process, and the different types of evidence produced (↑[Kucirkova, 2022](#); ↑[Puttick & Ludlow, 2013](#)). Instead, our report focuses on the systems and structures that enable more

sustainable and resource-effective EdTech Testbeds by offering a framework to support stakeholders in their work.

It is becoming more widely appreciated that EdTech Testbeds – “*environment[s] to test and experiment with EdTech in a real-world setting[s]*” (↑[Batty and colleagues, 2019](#)) – offer an approach

1. to ensure the best fit between EdTech and its anticipated users
2. to produce suggestions for improved EdTech design
3. to evaluate products at later stages of development.

At the same time, there is a long history of researcher-led projects in which EdTech is tested in such real-world settings (↑[Higgins and colleagues, 2012](#)). The goals for such studies include

- exploratory studies to test early-stage EdTech interventions, through to
- efficacy and effectiveness studies of more stable EdTech, and finally,
- research to support EdTech implementation in schools, often with a focus on teachers’ professional learning needs.

However, while these studies play an important role in the development of knowledge and practice, they typically do not establish long-term partnerships with schools, industry, researchers, and government. EdTech Testbeds are different from scientific studies, as they aim to establish long-term change through involving all of these actors in the ecosystem. Currently, there is insufficient research funding and human resources to meet the global research needs of the EdTech sector.

Over the last 5 years, several governments and non-government organisations have begun to systemise the process through which multiple types of EdTech can be tested in real-world K–12 school settings. These “*meta*”, or more system-wide, EdTech Testbeds appear to share several common features:

- a sustainable business model,
- a core team,
- national or regional government support (policy and/or funding),

- situated within (or close to) established EdTech incubators/accelerators,
- situated within (or close to) established research organisations,
- situated within (or close to) EdTech investors/funders,
- situated within (or close to) real-world settings (K–12 institutions, teacher education institutions, community settings).

It is clear that, as more system-wide EdTech Testbeds emerge around the globe, there is a need to build human capacity on how (best) to design, implement, and evaluate the different approaches. At the same time, there is an additional global challenge to train and provide ongoing professional learning opportunities for teachers³ (↑[Sailer and colleagues, 2021](#)). As EdTech becomes available to more teachers, governments are increasingly requiring teachers to stay up-to-date with EdTech, and to provide feedback on their experiences to help improve EdTech.

As EdTech becomes available to more teachers, governments are increasingly requiring teachers to stay up-to-date with EdTech, and to provide feedback on their experiences to help improve EdTech.

Hence, EdTech Testbeds offer an obvious opportunity to match EdTech type and schools to serve a dual enabling purpose. Firstly, for the EdTech sector to iterate, evaluate, and improve their products based on real-world data. Secondly, as recognised opportunities to enable and support teachers' lifelong professional learning. If teachers are involved in testing EdTech, it may give them agency in the development process and increase their confidence in implementing future-oriented technologies as the world continues to evolve. Furthermore, learners could get the opportunity to test EdTech, share their experience, and as such, play a role in co-designing future EdTech.

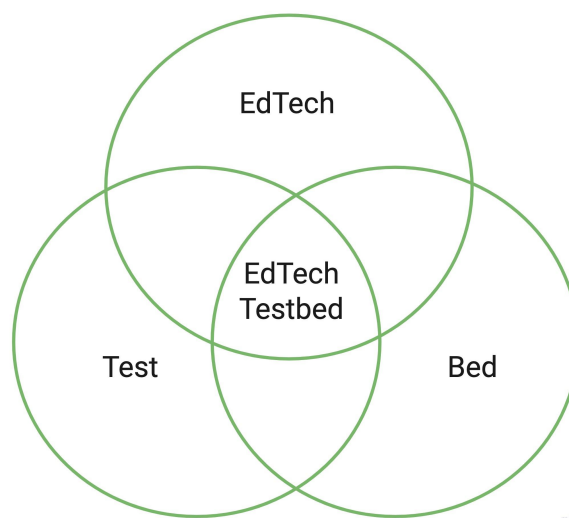
³ Equally, parents and caregivers also require planned opportunities to learn how to maximise the opportunities that EdTech offers for children's learning (see ↑[Clark-Wilson and colleagues, 2022](#)).

1.2. A simple description of an EdTech Testbed

For the purposes of this report, our description of an EdTech Testbed is simple (see [Figure 1.2.](#)). We illustrate the main components of any EdTech Testbed as three intersecting circles.

1. **EdTech:** The broad range of technology-supported products, services, systems, projects, programmes, processes, or structures that are developed and deployed to bring about positive outcomes in education systems.
2. **The Test:** The research methodologies developed to evaluate the impacts of EdTech in a real-world setting.
3. **The Bed:** The physical, online, virtual, or hybrid learning environment where EdTech is being implemented and studied.

Figure 1.2. A simple conceptualisation of an EdTech Testbed



For any EdTech Testbed, there are three key communities:

1. the end-users (learners, teachers, caregivers, communities etc.) who do the testing in the real-world setting;

2. the EdTech developers who provide the product to be tested;
3. the researchers, who advise and support the generation of appropriate evidence ([↑Cukurova and colleagues, 2019](#)).

However, EdTech Testbeds are widely supported by other key stakeholders within the EdTech ecosystem. These include the funders, investors, policymakers, EdTech incubators and accelerators, institutions, teacher unions, and professional communities who all provide human, financial, and tangible resources to enable the Testbed to operate.

Recently, several reports about EdTech Testbeds have been released. [Batty and colleagues \(2019\)](#) produced one of the first reports on EdTech Testbeds describing different models for how to deliver an EdTech Testbed in the UK. [↑Boot and colleagues \(2022\)](#) described EdTech Testbed initiatives in Europe based on the proposed models of [↑Batty and colleagues \(2019\)](#). Furthermore, several EdTech Testbeds produced their own reports in which they describe their procedures and outcomes of pilots (e.g., [↑Global Victoria EdTech Innovation Alliance, 2022](#); [↑Lee & Basma, 2021](#)).

Globally, there have been few opportunities to develop a wider conversation on best practice in the design, implementation, and evaluation of EdTech Testbeds that have been designed to operate at scale.

Globally, there have been few opportunities to develop a wider conversation on best practice in the design, implementation, and evaluation of EdTech Testbeds that have been designed to operate at scale. Recent panel discussions and plenaries on the topic at global events, such as Bett 2022, ASU-GSV, and EdTechX, have been well attended, which demonstrates stakeholders' desires to develop this shared conversation.

1.3. A guide to this report

This report is composed of three main parts.

In the first part ([Chapter 3. Pre-conditions for effective EdTech Testbeds](#), [Chapter 4. Key components of EdTech Testbeds: the EdTech, the Test and the Bed!](#), [Chapter 5. Impact goals](#)), we unpack current EdTech Testbeds and identify their key components and include commonly used methodologies and methods.

In the second part ([Chapter 6. Lessons learned from more system-wide EdTech Testbeds](#)), by focusing on four particular system-wide EdTech Testbeds (Swiss National EdTech Testbed Program, WISE Qatar Testbed, Swedish Edtest and Leanlab Education) we summarise the lessons learned.

In the third part ([Chapter 7. Key challenges for EdTech Testbed design](#)), we conclude with key challenges for effective EdTech Testbed design as we work towards a wider global offer of more system-wide EdTech Testbeds.

The report ends with a number of conclusions ([Chapter 8. Conclusion: Towards more effective EdTech Testbeds](#)), References, and an extensive Appendix.

Chapter 2. Introduction to Part 1: Unpacking EdTech Testbeds

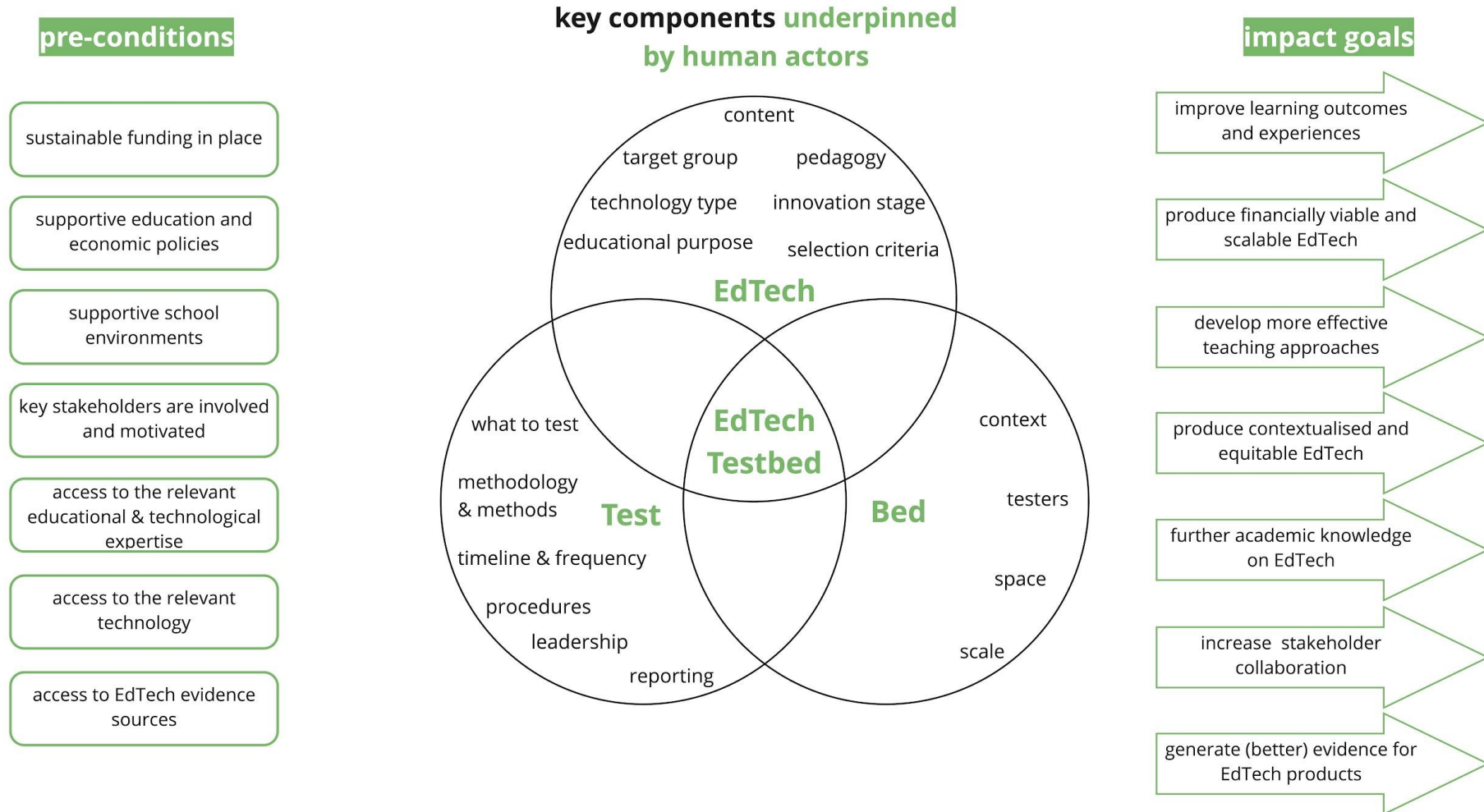
Our initial desk research and associated follow-up interviews⁴ reveal a wide range of similarities and differences in the way that the more system-wide EdTech Testbeds are being resourced, hosted, staffed, and sustained over time. The methods and approaches used to design and implement the EdTech Tests also vary greatly. With a focus on K–12 learners and settings, we begin by offering a tentative framework, described below.

Figure 2.1. below offers an initial conceptual framework for EdTech Testbeds; the figure has three parts:

1. **Pre-conditions** that should be met to increase the likelihood of the Testbed evolving to become an effective EdTech Testbed. Pre-conditions are covered in [Chapter 3. Pre-conditions for effective EdTech Testbeds](#).
2. **Key components** of an effective EdTech Testbed categorised by the EdTech, Test, and Bed, which are all underpinned by the decisions and actions of the human actors involved. Key components are covered in [Chapter 4. Key components of EdTech Testbeds: the EdTech, the Test and the Bed!](#)
3. A possible overarching mission or set of **impact goals** for the EdTech Testbed, i.e., the explicit goals for the EdTech Testbed. Impact goals are covered in [Chapter 5. Impact goals](#).

⁴ See [Appendix A](#) for details of our methods.

Figure 2.1. An initial, conceptual framework for EdTech Testbeds



Chapter 3. Pre-conditions for effective EdTech Testbeds

At a country level, several existing resources can help to determine the ‘readiness’ of an education system to maximise the potential impact of EdTech innovations. For example, the World Bank *EdTech Readiness Index* offers a tool to support country-level strategic planning by focusing on six pillars:

1. School Management
2. Teachers
3. Students
4. Devices
5. Connectivity
6. Digital Resources ([↑World Bank Group, 2021](#)).

We go beyond this to highlight some preconditions for effective, system-wide EdTech Testbeds, that is, Testbed models that can be scaled and sustained over time.

3.1. Sustainable funding in place

In all cases, there are costs involved in EdTech testing that require funding. For system-wide EdTech Testbeds, sustainable core funding is needed for dedicated staff and associated support infrastructures. Our initial research suggests the following main funding sources (separately, or in combination):

- Government funding (from education, business, innovation and enterprise, and technology budgets, separately, or in combination) to
 - School districts and schools

- National projects
- Individual schools and school groups
- EdTech incubators/accelerators
- Universities
- Philanthropic funding from individuals and foundations
- Venture Capital
- EdTech enterprises (from their own R&D budgets, or as a condition for grant or Venture Capital funding)
- Research funders

3.2. Supportive education and economic policies

In many countries and regions, education policies (and resulting budgets) focus on schools, school leadership, and teacher development; curriculum and assessment; and high stakes testing. Alongside, economic policies can lead to grant funding and fiscal support for entrepreneurial and technology initiatives that might favour the EdTech sector. However, it is less common for these policies to align in ways that enable the associated budgets to address a common educational mission.

As an exception, in Finland, the six regional governments aligned their educational and economic policies, to fund *6Aika*, a nationwide EdTech Testbed project that involved 300 companies and approximately 200 pilots.⁵

3.3. Supportive school environment

Globally, the capacity for individual schools (or groups of schools) to participate in EdTech Testing is greatly influenced by their political, social, and cultural contexts. In some cases, the schools may have to satisfy some externally imposed criteria relating to student attainment or leadership capacity.

⁵ †Results: Smart Learning Environments of the Future:
<https://6aika.fi/en/project/results-smart-learning-environments-of-the-future/>

In other cases, it may be that the teachers and/or learners require a certain level of technology access and proficiency before being eligible to test EdTech. Some countries, regions, and districts give total autonomy to teachers to pilot EdTech as they choose, whereas others only encourage EdTech innovation within more tightly controlled settings.

Consequently, before the development of more system-wide EdTech Testbeds, it is necessary to find out about, and seek to understand, the local school environment. Globally, a number of self-audit or assessment tools have been developed to ascertain the digital capacities of teachers, schools, and students. For example, in Europe, the [SELFIE tool](#), which launched in 2018. As of 23 January 2023, SELFIE has been used overall by 4.7 million users from over 37,000 schools, in 91 countries. This includes 3.87 million students, 672,940 teachers and 119,039 school leaders⁶.

In many cases, work will be needed to prepare schools and teachers to be committed and active participants within EdTech Testbeds, which is likely to require dedicated human and financial resources.

3.4. Key stakeholders are involved and motivated

The three key stakeholder groups are the *EdTech developers*, the *researchers*, and the *teachers, learners, parents, caregivers and communities*, within the educational settings.

Each stakeholder group has a role to play, and responsibilities aligned with that role, which captures their involvement within any particular EdTech Testbed instance. Alongside, there will be different motivations for their involvement, some of which are captured in [Appendix B](#). A report by the Global Victoria EdTech Innovation Alliance provides a more comprehensive list of the perceived benefits of one particular EdTech Testbed for different stakeholders ([Global Victoria EdTech Innovation Alliance, 2022](#), p. 11-13).

Beyond these three stakeholder groups, there are several other stakeholders who play key roles to ensure the sustainability and scaling of system-wide EdTech Testbeds. These include:

⁶ [↑Is your school making the most of digital technologies for learning? – https://schools-go-digital.jrc.ec.europa.eu/](https://schools-go-digital.jrc.ec.europa.eu/)

- *EdTech Testbed leaders*, who direct and manage all stakeholders in multiple and diverse EdTech Testbed projects.
- *Wider researchers within academic institutions and organisations*, who innovate research methods to respond to evolving EdTech Testbed needs, and provide developmental research knowledge and support for the whole community.
- *EdTech programme implementers*, who provide guidance and support that aim to ensure the best possible outcomes for the testing.
- *Policymakers*, who have a direct interest or role in the funding and evaluation of system-wide EdTech Testbeds.
- *Investors and funders*, who provide essential funding to support both the sustainability of system-wide EdTech Testbeds and for individual testing of EdTech within their portfolios.
- *Evaluators*, who support the formative and summative evaluation of the EdTech Testbed processes/outputs to enable the evolution of a strong learning culture across all stakeholders.

3.5. Access to the relevant educational and technological expertise

While it is obvious that the technology is provided to all those involved in any particular EdTech Testbed instance, the nature of the relevant educational and technological expertise is often less clear.

While it is obvious that the technology is provided to all those involved in any particular EdTech Testbed instance, the nature of the relevant educational and technological expertise is often less clear.

Research evidence over the past 50 years indicates that it takes a significant amount of time for teachers to develop confident and competent uses of technologies (and associated educational content and pedagogies) that are new to them (↑[Koehler and colleagues, 2013](#); ↑[Vanderlinde & van Braak, 2010](#)).

Educational and technological expertise develops, in and through, classroom practice and this development can be accelerated when teachers are supported within learning communities (↑[Cristol and colleagues, 2019](#)). Teachers bring much educational expertise, which can be amplified when they have opportunities to share this know-how with educational advisors, EdTech developers, researchers, and teacher educators. For teachers to be maximally supported in EdTech testing, they must be able to have opportunities to have a dialogue with others. For example, the EdTech developer who has the responsibility for providing the technological expertise to the Testbed; in this case the developers will need to be sensitive to the educational content and pedagogy that underpins the innovation. The nature of the EdTech “*onboarding*” may need to be carefully designed, in collaboration with educational researchers, and training and support offered throughout the testing period to enable teachers’ practices to evolve with time.

3.6. Access to EdTech evidence sources

Much has been written on the need to ensure that all EdTech is designed, implemented and evaluated in ways that draw from and contribute to educational research (↑[Knight and colleagues, 2020](#)). In addition, a wealth of existing resources support EdTech developers and researchers to develop the growing global evidence base. In the USA, [Digital Promise](#)⁷ has developed the [Research Map](#),⁸ which provides an easy-to-use gateway to research on topics that are relevant to EdTech design, implementation, and evaluation. Additionally, the [EdTech Hub Evidence Library](#)⁹ has an extensive collection of research that has particular relevance for low- and middle-income countries (LMICs).

⁷ ↑[Digital Promise: https://digitalpromise.org/](https://digitalpromise.org/)

⁸ ↑[Digital Promise – Research Map: https://researchmap.digitalpromise.org/](https://researchmap.digitalpromise.org/)

⁹ <https://docs.edtechhub.org/lib/>

Chapter 4. Key components of EdTech Testbeds: the EdTech, the Test and the Bed

In the following chapters, we explain key components of EdTech Testbeds. Furthermore, [Spotlight 4.5](#) illustrates each of the key components of one particular EdTech Testbed, namely the Swiss National EdTech Testbed Program.

4.1. The EdTech component

The first key component of an EdTech Testbed is the Educational Technology (EdTech) itself, which should be distinguishable by its:

- educational purpose or goal
- technology type
- target group (or end-users)
- educational content
- pedagogy
- stage of innovation or development
- quality requirements

Below, we briefly explain these components and highlight why it is important for them to be carefully considered when designing and implementing the Testbed in which EdTech is to be evaluated.

Educational purpose or goal

Fundamental to any EdTech design is a clear understanding of the educational context and how the EdTech can address current educational issues or build skills (e.g., further digital citizenship). For example, EdTech has the potential to realise increased access to education, or improving the quality of learners' or teachers' experiences. Before designing the Testbed approach, a shared understanding of the educational purpose of the EdTech is essential.

Fundamental to any EdTech design is a clear understanding and validation of the educational problem that the EdTech is designed to address.

The type of technology

There is a considerable variety in the available EdTech products and services, which range from small digital interventions to teach defined content - through to large-scale EdTech services, programmes, and systems. Some wide-ranging examples include:

- Educational radio programmes to support out-of-school learners to access learning opportunities (e.g., [Rising Academies](#));
- A digital game to learn to read (e.g., [Graphogame](#));
- An AI WhatsApp chatbot to support teachers with resources and keep them motivated;
- An authoring tool that empowers teachers to build 3D experiences to use in their teaching practice;
- A Geo-information system used to map schools to assist in teacher allocation;
- A national education management system to improve resource allocation and real-time data processing for decision-making (e.g., [NEMIS Kenya](#));

- A platform to provide assistance or advice to teachers prior to or after they have purchased a product.

Given this diversity, some more systemic EdTech Testbeds specialise in particular types of technology (e.g., Exten.D.T.2¹⁰, which focuses on emerging technologies such as artificial intelligence, augmented reality, and educational robotics). This enables the Testbed leaders to build expertise in devising appropriate Testbed approaches, before perhaps broadening the scope. Designing a test that can be applied on a large variety of products, services, programmes, and systems is challenging. Therefore, it is helpful to consider what technology to include in an EdTech Testbed, particularly if the EdTech is at an early stage of development – aiming to scale to serve multiple EdTech types over time is advisable.

Designing a test that can be applied on a large variety of products, services, programmes, and systems is challenging.

Target group

The target group of EdTech refers to the end-user(s) the EdTech is designed to impact educationally. A key tension to navigate in Testbeds is that purchasers are often not the end-users (e.g., school head teachers who decide about purchasing a tool for use in language classes). The target group for the EdTech greatly influences the research methods for the Testbed. This includes what data is collected and how and for what purpose, alongside ethical considerations (↑Lupton & Williamson, 2017).

In this paper, we focus on K-12 education: learners, teachers, parents, and communities. The choice of the target group influences the methods. For example, informing young children about a study and understanding what consent means for this target group is more challenging than informing their

¹⁰ ↑Project: Extending design thinking with emerging digital technologies (Exten.D.T.2): <https://lnu.se/en/research/research-projects/project-extending-design-thinking-with-emerging-digital-technologies/>

teachers about a study and gaining consent from them (or their parents or caregivers). It is also important to consider the particular demographic and socio-economic and cultural characteristics of the target group to ensure that findings are reliable and representative of the Testbed context. Educational researchers bring the expertise to help in this task and, importantly, can advise on the evaluation methods appropriate for the target users, which are most appropriate for the Testbed.

Educational content

Most EdTech that is designed for K–12 learners will be designed to address particular educational content that relates to the curriculum. Consequently, if the Testbed is seeking to evaluate the learning of the target group, it will be necessary to understand the specific educational content and its intended purpose in the curriculum.

Pedagogy

EdTech has been criticised for having more of an economic and techno-centric rationale than an educational / pedagogical one ([↑Ramiel, 2019](#)). The introduction of any EdTech to K–12 classrooms disrupts the classroom norms and places demands on teachers and learners as they grapple with new ways of teaching and learning.

Therefore, EdTech Testbeds can contribute when it comes to testing the impact of the EdTech on the human aspects of the classroom. For example, it can test the

- agency of the teacher and of the students,
- the impact on social interaction and interrelationships,
- ethics, and so on.

Although, the examples in the list above might not be the primary focus of the EdTech Testbeds (as this might not be the goal of a start-up for example), they can play a role in assuring the quality of EdTech products and services by enabling rich discussions about the pedagogical challenges and helping to make more explicit the underlying pedagogical theories of EdTech products or services ([↑Boot and colleagues, 2022](#)). The involvement of both education researchers and

practitioners in the design of the test for a certain product or service can enrich the specific hypotheses concerning the EdTech to be thoroughly tested.

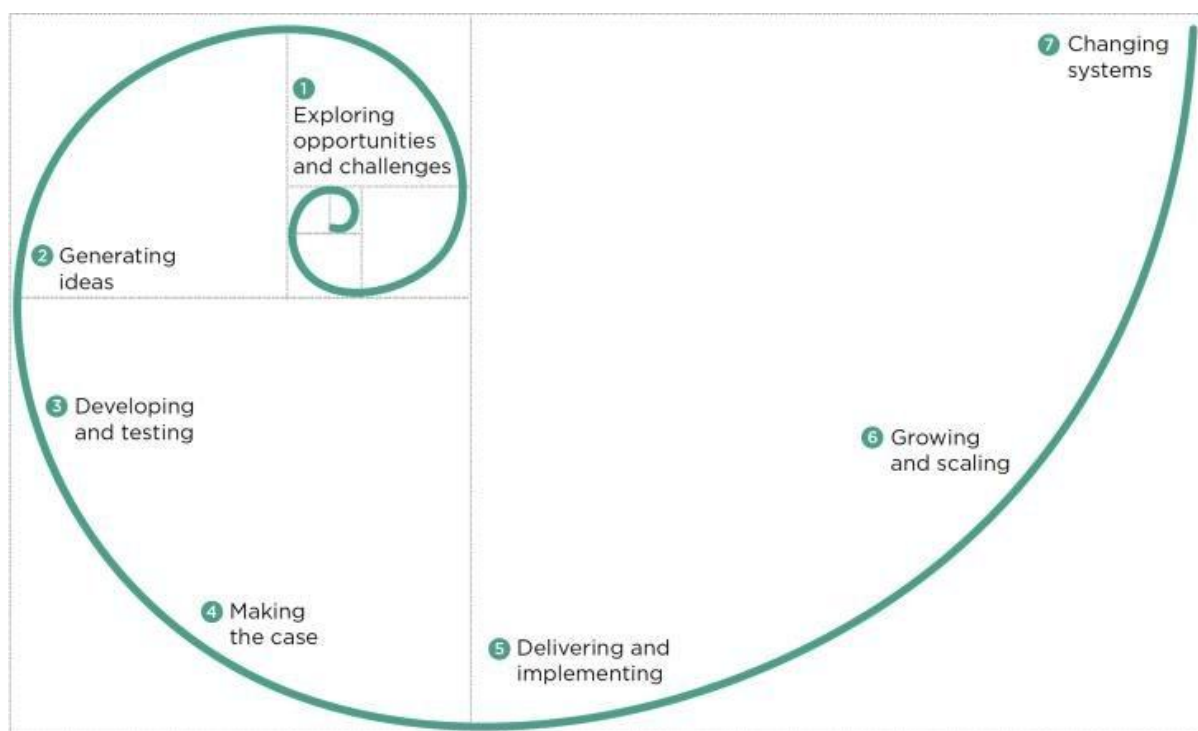
Stage of Innovation (or Development)

EdTech Testbed designers often raise the question:

What is the minimal stage of development for a product so that it can be tested in schools?

The Innovation Spiral developed by Nesta (Figure 4) is useful in responding to this question.

Figure 4.1. Innovation spiral as presented in *Using Research Evidence: A Practical Guide* (Nesta, 2016, p. 14)



The innovation spiral offers seven stages of innovation, which range from *exploring opportunities and challenges* through to *delivering and implementing* and ultimately, to *changing systems*.

What is the minimal stage of development for a product so that it can be tested in schools?

For EdTech innovations, the Testbed methodologies and methods will be very different across the different stages.

An EdTech concept can refer to identified opportunities and challenges in education, and be based on newly generated ideas that address current or future needs. EdTech Testbed designers have questioned whether EdTech products or services in early stages of development are ready to be tested in real-life settings such as classrooms – as the value for the teacher and the learners may be very limited. Others have argued that ideas, concepts and products *should* be tested from an early stage as only when EdTech is co-created with teachers (and other education stakeholders) can it really address their needs.

Once the EdTech concept is transformed to a working product / service / programme, it will need iterative testing to focus on its usability, functionality, and effectiveness to improve its design. Typically, this is a stage when early-stage EdTech companies value a lot of feedback from target users, which can be facilitated by EdTech Testbeds.

The later stages of innovation indicate when EdTech – that might already have been tested in smaller scale tests – needs to be scaled more widely to fully understand the challenges of large-scale implementations.

A recent study indicated that the proven impact of EdTech based on small-scale experiments is not always reproduced when EdTech are scaled up ([Kizilcec and colleagues, 2020](#)). The role of EdTech Testbeds in such contexts will be different as the type of evidence, and scale of the studies might be entirely different.

Selection criteria

All EdTech Testbeds need to select the EdTech to be tested, which might be thought of as “*EdTech readiness*”, and connect to the previously described *stage of innovation*.

Several different frameworks, tools, and criteria exist to support this selection process, and each focus on a different perspective:

For example, the Helsinki Testbed¹¹ uses the selection criteria listed below based on the potential of the EdTech (↑[Boot and colleagues, 2022](#), p. 27):

1. Need for development of an educational solution;
2. Product needs to address a real challenge or needs to be an improvement compared to existing solutions;
3. The product has potential and is scalable;
4. The City of Helsinki itself needs information or has another need related to the solution being tested, and/or resources to enable the tests to be carried out.

To be included in the Swedish Edtest¹², an EdTech enterprise must fulfil two main criteria¹³:

1. The digital learning resources must be newly developed or newly launched and testable (ready to be used in a school);
2. The company should be able to provide a description of the specific part of the resource they are looking to test and what feedback they are looking to receive to improve their resource further.

¹¹ ↑[Testbed Helsinki / EdTech: https://testbed.hel.fi/en/edtech/](https://testbed.hel.fi/en/edtech/)

¹² ↑[Swedish Edtest – Utveckla digitala läresurser med Swedish Edtest: https://edtest.se/sv](https://edtest.se/sv)

¹³ ↑[Swedish Edtest – Frequently asked questions: https://edtest.se/en/frequently-asked-questions](https://edtest.se/en/frequently-asked-questions)

Another example would be Leanlab Education's *Pilot Readiness Audit*¹⁴. This evaluation measures an EdTech company's functionality, usability, interoperability, and their data privacy and security practices. This audit is used to

1. Ensure the product meets a certain standard before it is placed inside a classroom.
2. Suggest an appropriate level of research for the tool.

Another approach to selecting EdTech is to look at the people behind the EdTech enterprise to assess their potential to design impactful EdTech (and to continue to iterate on their design as they scale their enterprise). Based on †Moeini's 3-year doctoral study (†2020), Goldstar Diagnostics¹⁵ reveals otherwise hidden aspects of the company's processes, assets, capacities and mindsets, which provide an indication of trustworthiness. The leadership team of the EdTech enterprise completes an online diagnostic survey that then provides aggregated scores and formative feedback about their Leadership Vision, Learning Culture, Sense of Purpose, Teamwork, Research Know-How, and Action Orientation.

Further, educational researchers investigated what quality criteria can be taken into account when deciding on the quality of particular EdTech. For example, †Kolak and colleagues (2021) developed a tool for assessing the educational potential of apps for preschool children in the UK.

4.2. The "Test"

The second key component of an EdTech Testbed is the "*test*", which is the research design developed to evaluate the impacts of the EdTech. Ideally, the research design will focus on the educational goals for the EdTech. For example, ranging from improving learning outcomes - to contributing to more effective teaching - to other impacts on the classroom. In addition, as there can be substantial hype about the potential impact of EdTech, it is important to build trust in the system by generating better evidence to support (or refute) claims. In what follows, we describe subcomponents that need to be considered when designing an EdTech test.

¹⁴ †Pilot Readiness Audit: <https://www.leanlabeducation.org/pilot-readiness-audit>

¹⁵ †Goldstar EdTech Diagnostics: <https://www.goldstared.com/>

What to test?

Having a clear goal for the test is essential. EdTech Testbeds can play a role as facilitators in identifying the specific (research) questions that will be treated when testing EdTech ([↑Massachusetts Institute of Technology, 2016](#)). EdTech Hub¹⁶ uses the education ecosystem 6Ps framework to ask questions about people, provision, place, product, pedagogy, policy, and product (see [Spotlight 4.3.](#)). Questions can be very diverse; what can be tested can be described with the following keywords: *what, what if, why, who, how, where, and when*. The question(s) will drive the methods and methodologies used to either predict, understand, explore, test, describe, explain, compare, or correlate.

Typically, EdTech companies want to undertake research because they are looking to evidence that they can meet a need, build skills, or have identified a problem or potential obstacle. This will often inform the types of research questions and methodologies used. Leanlab Education uses a tiered research journey (see [Spotlight 4.7.](#)) to guide companies to find an appropriate type of test. These studies range from validating an idea or prototype with a few teachers or parents in the “*sandbox*” stage to correlational study, implementing the product across a school system.

Recently, [Pratihast & Mote \(2022\)](#) developed their Blueprint for Inclusive Research and Development in Education – a framework and toolkit (BIRD-E¹⁷). The blueprint can be used by different partners of EdTech Testbeds to support them in thinking about a research hypothesis, identify and articulate data needs for researchers, solution providers or for practitioners to compile and evaluate the impact of interventions. Through the framework and toolkit, the research can be undertaken in a more structured manner. Research questions can be very diverse and include, e.g.,

- product testing (features, intended outcomes),
- implementation frequency (duration, frequency),
- implementation fidelity (school access, school support, programme fidelity),

¹⁶ ↑EdTech Hub: <https://edtechhub.org/>

¹⁷ ↑BIRD-E: <https://www.bird-e.org/>

- implementation costs (device/student, teacher training time).

Methodologies and methods

Methodologies refer to

“the strategy, plan of action, process or design lying behind the choice and use of particular methods and linking the choice and use of methods to the desired outcomes” (↑Crotty, 2020).

Methodologies used in EdTech Testbeds tend to have two different roots, and in some cases have converged nicely. Given the importance of this subcomponent, we give a complete overview of widely used EdTech Testbed methodologies in [Appendix B](#).

The first root is the field of education research and shares similarities with research methodologies from social sciences. These methodologies may also be found in anthropology, sociology, or psychology. The primary focus of these approaches is to build evidence on EdTech effectiveness and its broader impacts, where contributing to EdTech's future development is secondary.

The second root is the technology innovation sector, where methodologies used share similarities with the language of entrepreneurship and innovation. These methodologies may overlap with those used to develop mHealth, AgriTech or FinTech, and are more often presented as toolkits, principles, or processes rather than (research) methodologies. The primary focus of these approaches is to improve the design and implementation of EdTech, where contributing to global evidence is secondary.

Different methodologies are applicable for different stages of innovation (↑Kucirkova, 2022; ↑Puttick & Ludlow, 2013; ↑Global Victoria EdTech Innovation Alliance, 2022). At conceptualisation stages, methodologies like *“Design thinking for educators”* and *“Human-Centred Design”* can assist with brainstorming and investigating the problem. In development stages, methodologies like Design-Based Implementation Research (DBIR), Problem-Driven Iterative Adaptation (PDIA) or the EdTech Hub’s Sandbox Approach (see [Spotlight 4.3](#).) can assist in weeding out teething issues. This period is arguably the most important and these methodologies emphasise:

- taking a problem-first approach;
- understanding the context and the ecosystem factors;
- involving diverse stakeholders and end-users;
- being agile and iterative, incrementally improving the EdTech in question.

Finally, when considering scaling the EdTech to be used with a larger audience, we recommend first proving that the EdTech can realise its intended goals. Designs such as (quasi-)randomised controlled trials (RCTs) and Systematic Mixed Methods Research (SMMR) have been frequently used and allow for effectiveness to be assessed. Even when RCTs are conducted on a large scale, there can be concerns about implementation fidelity and how generalisable the outcomes might be beyond the original study context.

Across methodologies, a range of methods can be used. Methods refer to

“the techniques or procedures used to gather and analyse data related to some research question or hypothesis” (↑Crotty, 2020).

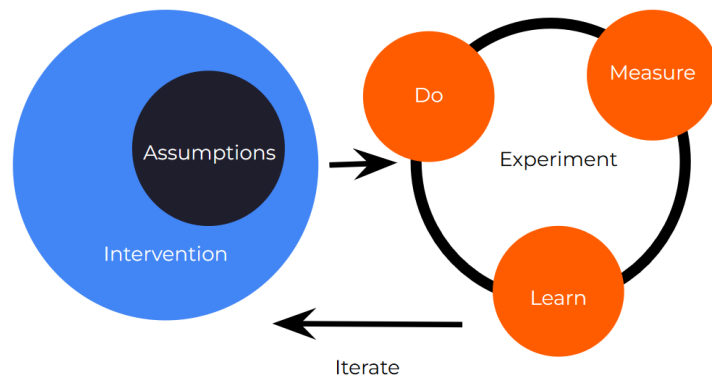
Methods range from those more traditionally used in education research such as interviews, surveys, focus group discussions, and ethnographies to more innovative methods such as mindset mapping, photo journaling, and storyboarding. Methods that are more quantitative include A/B testing or factorial designs. A detailed list of methods with many useful links is given in [Appendix C](#).

4.3. Spotlight 1. The EdTech Hub Sandbox Approach

Building on “*lean impact*” (Chang, 2018) and user-centred design, the sandbox process is designed to help generate real-time evidence. It has two key concepts, outlined below (EdTech Hub, 2022).

1. Experimentation: the learning and iterating cycle

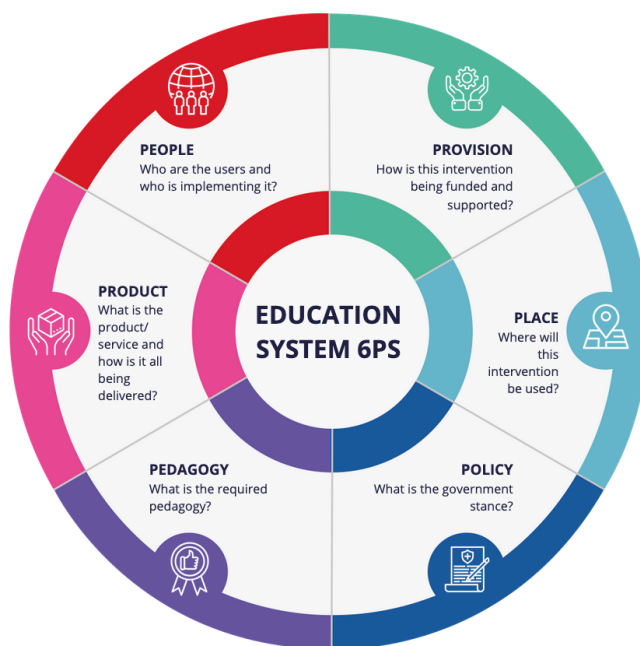
The experimentation framework tests your assumptions that underpin your solution. Through a process of doing, measuring, and learning, you iterate and improve as you go.



2. The Education system 6Ps framework.

EdTech is embedded in a broader ecosystem that needs to work in-sync to reach sustainability and scale.

Past sandboxes include WhatsApp-based learning for refugee children in Lebanon, helping caregivers foster learning at home in Kenya, making radio work better for children in Uganda, scaling personalised learning in Malawi, and using laptops to distribute video sign language lessons to deaf children in Pakistan.



Timeline and frequency of use

The timeline and frequency of use refers to "*when and how often the EdTech is used*". As illustrated by the following examples, EdTech Testbeds greatly vary in the timelines they use to test EdTech:

- Helsinki Testbed: mean duration of test activities ranges from 3–6 months ([↑Boot and colleagues, 2022](#));
- Swedish EdTest: The participants of the Testbed decide how long the test phase will be;
- Swiss National EdTech Testbed Program: the duration of a Testbed depends on the product, but takes a maximum of 6 weeks;
- Leanlab Education Codesign Product Research: duration of a Testbed ranges from 4 weeks to 6 months based on the stage of the EdTech and needs of the school.

Procedures

The procedures for the test include the materials, templates, protocols, contracts, procurement processes, guidelines on methods, scenarios, instruments, among others. Many different procedures are used by Testbeds, ranging from structured support for the implementation of the EdTech, to providing templates for the test, or providing only administrative and technical support.

Many different procedures are used by Testbeds, ranging from structured support for the implementation of the EdTech, to providing templates for the test, or providing only administrative and technical support.

Helsinki Testbed provides support to facilitate the test. i.e., companies or other partners need to have a clear research question when they enrol. Based on their question, the requisite staffing and time are allocated. The Testbed is mainly concerned with bringing the different parties (school / company / researcher) together and taking care of practical arrangements (e.g., venue).

Swedish EdTest provides administrative support, facilitates communication between parties, agreements, and supports General Data Protection Regulation (GDPR) compliance. The following protocol is used:

1. School decides which product/service they want to test;
2. Introductory conversation between parties, agreement;
3. Test procedures are defined, there is a standard test guide;
4. Generate report of testing;
5. Certificate of participation.

Swiss National EdTech Testbed Program provides administrative support and support on site in the classrooms, facilitates the communication and planning of the Testbed with the teachers, evaluates the feedback on the products. The following protocol is used:

1. Teacher/school register for the Testbed; decide which product they want to test;
2. Introductory conversation between teachers and EdTech Collider; planning the Testbed;
3. Conducting the test in schools;
4. Send feedback form;
5. Certificate of participation.

WISE provides monthly professional development and technical support to teachers. Furthermore, WISE observes lessons and stays in communication with participating teachers and school leaders to provide further support.

Leanlab Education includes these elements:

1. Identifies alignment between an EdTech company and a school system;
2. Matches schools and companies and facilitates Memorandums of Understanding, data-sharing agreements, and contracts with all stakeholders;
3. Codesigns parameters of the research study with all stakeholders;
4. Implements the research study and analyses the findings;
5. Produces a final report to be shared with all stakeholders;
6. Issues a certificate to the company; contingent on incorporation of participant feedback into the product design.

Leadership: who does the actual testing?

The test can be led by different stakeholders and can be teacher-led, company-led, or researcher-led ([↑Atherton, 2019](#); [↑Hollands & Escueta, 2017](#)). The test procedures will differ depending on the goals of the test leader. The goal of company-led tests is often focused on how to improve a product, or which features can be added, to lead to more sales. Companies often have limited knowledge about methodologies and methods that can be applied to testing EdTech systematically. Moreover, companies do not always have the resources and time to test at scale.

The goal of researcher-led tests is to investigate the effectiveness of EdTech in general, and examine the impact these technologies (or their characteristics) have on teachers and learners. In the context of a scientific study led by researchers to determine the effectiveness of a particular product, teachers and learners are also often involved. In this case, testing procedures are often strictly prescribed, with little leeway for teachers to fully integrate the EdTech into their current practice.

The goal of teachers is mostly centred around how EdTech can improve their daily classroom practice and, by extension, learning outcomes. This can range from testing learning portals to give more differentiated instruction to learners, or using a tool that helps the teacher with their administration or evaluation. In fact, every teacher implementing EdTech in the classroom is testing it in one way or the other, and they have their own ideas about its potential and the challenges of

using it. Given the different goals of stakeholders, designing one test that is useful for all stakeholders may be overambitious. Nevertheless, EdTech Testbeds can provide the necessary support different stakeholders need to design tests.

The goal of teachers is mostly centred around how EdTech can improve their daily classroom practice and, by extension, learning outcomes.

The Swiss National EdTech Testbed Program is neither teacher-led nor company-led, as responsibility for the Testbed rests with a third non-involved party, the non-profit organisation Swiss EdTech Collider. Leanlab Education employs this model as well. The Swedish Edtest, on the other hand, is company-led, as the management of the Testbed is the responsibility of the company.

Reporting of results

The results of a test should be made available to stakeholders as appropriate/relevant. It is essential to agree on what information, in particular, should be shared and with whom. First and foremost, EdTech developers, researchers, learners, teachers, and caregivers who are directly involved in a test are the primary stakeholders and will benefit directly from the results.

Second, the results may also be of interest to a wider community of stakeholders. For example, the results of testing an app to teach reading in the classroom can also provide valuable insights for other teachers, EdTech developers, and researchers working in the same domain. Again, EdTech Testbeds have a role in discussing what results can be shared with this community (and what information should remain confidential).

There are several papers that provide helpful frameworks describing different types of evidence. [Cukurova and colleagues \(2019\)](#) distinguish scientifically based research, local knowledge, perspectives, and values of users and educators, and professional experience and judgements. The US department of

education's ESSA [Tier framework](#)¹⁸ describes the rigour of the research with Tier 1 being most rigorous (randomised experiments), Tier 2 generating moderate evidence (quasi-experiments), and Tier 3 producing promising evidence (correlational studies). A very recent report by [Kucirkova \(2022\)](#) describes three types of evidence: foundational (literature reviews and desktop research), formative (observational or correlational), and summative (experiments or RCT) evidence. There are also the Nesta standards of evidence ([Puttick & Ludlow, 2013](#), p. 2):

- **Level 1:** describe what you do and why it matters, logically, coherently and convincingly;
- **Level 2:** capture data that shows positive change, but you cannot confirm you caused this;
- **Level 3:** you can demonstrate causality;
- **Level 4:** you have one +independent replication evaluations;
- **Level 5:** you have manuals, systems and procedures to ensure consistent replication and positive impact.

Finally, the [Global Victoria EdTech Innovation Alliance \(2022\)](#) states that evidence consists of three stages, namely effects, outcomes, and impact. Effects are the changes. Outcomes are the specific and measurable effects. Impact refers to the story, focusing on the long-term results or changes.

Test results can be presented in several ways. Written reports are the most traditional way of disseminating test results. Written reports are often for internal use as they contain many details and confidential information. The downside of written reports is that it is very time-consuming to write these reports, and entails the risk that the report will soon be outdated given the often fast evolution of EdTech designs ([Cukurova and colleagues, 2019](#)).

EdTech Testbeds can facilitate and support by giving suggestions about whom to involve in writing the report and by providing resources such as templates to aid in this process. Additionally, more dynamic forms to disseminate results to a wider community are videos, website pages, blogs, infographics, social media, and

¹⁸ [↑Evidence & ESSA: What Is Evidence And How Should We Use It? – http://www.research4schools.org/evidence-essa/](http://www.research4schools.org/evidence-essa/)

digital archives. Results can also be shared through events where different stakeholders meet, such as presentations, poster sessions, workshops, and demonstrations. These formats promote interaction and community building.

4.4. The “Bed”

A third, key component of an EdTech Testbed is “*the Bed*”, which is the setting in which the EdTech is tested. Each setting has its own particular context, testers, learning space, and scale. In some cases, the test design will dictate the beds – the research methods will state the broad context in which the testing will take place. However, it is more usual for the beds to be more opportunistic, in which case the more that is known about them, the more valuable the evidence that they generate. This is particularly pertinent when it comes to creating equitable access to EdTech by ensuring that it is designed and tested for a wide range of learners and teachers.

Context

The context refers to the geographical, institutional, economic, societal, and political conditions that describe the Testbed. For example, testing in the private school sector, compared to public schools, can produce very different results. At the teacher or student-level, it is also important to gather demographic data that might enhance the analysis of the test outcomes. This raises ethical issues concerning data collection, storage, privacy, and ethical consent. However, the quality and potential future impacts of the research findings far outweigh the extra effort needed to establish high ethical standards within any EdTech Testbed model.

Testers

The Testbed is characterised by the people testing the product (or the *research sample* or *participants* in the language of the test designers). As this report is aimed at K-12 education, the testers are mainly students and teachers in preschool, primary, and secondary schools.

The roles and responsibilities of the testers should be widely understood by all involved. For example, the learners may only be required to test the EdTech by

completing tasks and activities as part of their normal day-to-day lessons. Alternatively, they might form focus groups where they are invited to feedback on their experiences.

Similarly, teachers might only be observed while planning, teaching or evaluating lessons involving the EdTech. Another role for them would be as a practitioner-researcher with the responsibility to collect research data during the test.

To conclude, at the level of the EdTech Testbed, the choice of the testers is likely to be influenced by:

- Whether or not the EdTech is perceived to be ready to be tested;
- Whether or not the technology (and associated infrastructure and support) is available to both teachers and learners;
- Whether or not the school leadership is knowledgeable and supportive, ideally offering a professional learning community;
- The number of students available to test the EdTech, their age range, and the perceived benefits to them;
- The number of teachers available to test the EdTech, their backgrounds, capacity to be involved, and the perceived benefits to them;
- Whether or not it is possible to obtain genuinely informed consent from the participants.

Spaces for learning

The huge global diversity in the spaces in which learning takes place needs us all to think beyond what we have come to refer to as “*classrooms*”. The global coronavirus pandemic widely introduced a new vocabulary to the world to include virtual, distance, synchronous, asynchronous, and hybrid learning spaces, among others, which all rely to some extent on digital technologies and digital educational resources. Furthermore, the learning spaces in low- and medium-resource countries look very different to those in highly resourced countries. Different learning spaces are more conducive to different types of

technology and pedagogical approaches, which are factors that need to be carefully considered when selecting Testbeds.

Scale

The number of Testbeds (and testers) for any particular EdTech test will usually be dictated by the research design, particularly for larger, formal research studies. However, as many Testbeds arise in more opportunistic ways, it is important to report the scale of the test in the findings. This would include the numbers of schools, teachers, and learners involved (supplemented by key contextual information).

At the level of the EdTech Testbed, the scale of its work must also be considered. For example, how many different EdTech products can be simultaneously tested, and how can a network of beds be sourced, managed, and involved?

4.5. Spotlight 2: Swiss National EdTech Testbed Program¹⁹

In this spotlight, we use the conceptual framework for EdTech Testbeds (Figure 4.1.) to describe the key components of the Swiss National EdTech Testbed Program.

The Swiss National EdTech Testbed Program was launched in September 2022 by the non-profit-organisation Swiss EdTech Collider. The goal is to create and establish a nationwide “*test and learn*” innovation model that enables rapid EdTech testing, structured feedback loops for EdTech products, connecting technology providers with educators, and strengthening teachers’ digital skills. Through this collaboration between EdTech suppliers and teachers, teachers actively contribute to the development of educational technologies. The tests therefore focus on applications used in the classroom (K–12). The Swiss National EdTech Testbed Program is guided by the “*test and learn*” method (↑Batty and colleagues, 2019).

EdTech

Target group: students from 4–15 years old and teachers

Pedagogy: the pedagogical concept of the EdTech product is evaluated during testing

Technology type: no specific type of technology (platforms, apps, robots)

Innovation stage: the product must be ready to use

Educational purpose: improve quality of learners’ and teachers’ experiences, simplification of various processes in everyday school life (e.g., learning management system, autocorrection)

Selection criteria: readiness, potential of the product, product addresses a real challenge or need, product provides added value for teachers and/or students

¹⁹ ↑Swiss National EdTech Testbed Program | Gebert Rüt Stiftung:
<https://www.grstiftung.ch/de/media/portfolio-grs-074-21~.html>

Educational content: curriculum-related content

Bed

Context: high-resource context, four national languages, only public schools

Testers: students and teachers in preschool, primary, and secondary school

Space: mainly physically in classrooms, or hybrid

Scale: more than 40 tests with five EdTech products so far (January 2023)

Test

What to test: product testing, user testing

Methodology and methods: test and learn approach ([↑Batty and colleagues, 2019](#)), rapid evaluation cycles are used to quickly generate actionable insights to make evidence-based decisions for developing and optimising EdTech products and improving their implementation

Timeline and frequency: depend on the product, short test phase of a few weeks or a few months, repetitive

Leadership: neither teacher-led nor company-led, the responsibility for the Testbed rests with the non-profit organisation Swiss EdTech Collider

Reporting: *“good practice paper”* about the implementation and use of specific EdTech products are made available to interested stakeholders

4.6. Spotlight 3: Impulsionar (Brazil)

General information

The Brazil EdTech Testbed Program, [Impulsionar](#), began in 2021 and consists of a partnership between

- the Lemann Foundation (implementing partner, financial partner),
- IDB Lab, innovation laboratory of the Inter-American Development Bank Group (IDBG) (financial partner),
- Imaginable Futures (financial partner),
- Quintessa (an impact accelerator, is responsible for the articulation, coordination, and acceleration of the partners involved) and
- Instituto Reúna (the pedagogical partner and creator of the Pedagogical Model for prevention and reduction of the gaps that are being implemented in public school networks by the organisations).

The impact goal of the programme is to innovate through inclusion (narrowing educational gaps) and technology (EdTech). EdTech companies receive resources to test and improve their solution in the public sector, plus receive support to develop their Business to Government (B2G) model. Currently, pilots are ongoing, with 5 EdTech products, in 7 municipal secretaries of education across Brazil. The programme will be evaluated by an independent organisation (Germina). In addition to testing EdTech in schools, the EdTech Testbed also supports companies by helping them to take advantage of new laws that facilitate open innovation in the public sector in Brazil (e.g., ‘Marco Legal das Startups’, Lei Complementar nº 182, of 1 June 2021).

EdTech

Target group: K–12 students, public schools, teacher training and implementing organisations, teachers, and Municipal Secretaries of Education

Pedagogy: pedagogy varies according to the type of EdTech and the problem that it tackles (for example evaluation, data management or adaptive learning). In addition, the program’s pedagogical model (created

by Reúna, the pedagogical partner) provides a framework and guidance for municipal secretaries of education, and its public-school network on how to include these technological solutions as a didactic resource – fostering a student-centric approach

Technology type: depends on the EdTech, there are three types:

1. Evaluation -app for helping teachers to create student evaluation,
2. Data Analytics -solution that collects and organise educational data for teachers and for management purposes or
3. AI for adaptive learning – a platform that uses gamification strategies and AI to respond to the student's pace.

The Edtech being used is tablets (in some cases donated by the program), computers, or students' own cellphones – and in some contexts with infrastructure challenges (no internet and / or hardware) a combination of online and offline activities is used.

Innovation stage: startups, with proof-of-concept and previous experience, but with B2G model / operations / product still to be developed or improved. The intention was not to work with established innovations and companies, although an exception is a new EdTech product developed by a big company.

Educational purpose: reduce and prevent learning losses in the public school system in Brazil, specifically in Portuguese and mathematics. Teachers (and the education municipality as a system) develop pedagogical guidance, technical abilities, management processes and uses of technology. The key outcome measures are the % of students attaining grade expectations in mathematics and Portuguese.

Selection criteria: a diagnostic was done with the municipal government to prioritise the problem that was going to be tackled and define criteria for each education municipal system. Criteria were related to price, quality, technical specifications, UX/UI, etc. Then, three public sourcing mechanisms were tested to formalise the public-private (EdTech) pilot:

1. An open call to address identified challenges, launched via the PitchGov platform;
2. Hire of a startup with new public legal framework of innovation startups ('Marco Legal das Startups': Lei Complementar nº 182, of June 1st 2021) or
3. Hire of a startup with regular public procurement framework, selection of 5 to 6 educational technologies (each municipality selected one).

Educational content: EdTech that helps to reduce and prevent learning losses and address the challenge of the education municipal system of the cities participating in the program. Focus on solutions that in some way improves mathematics and Portuguese teaching (student evaluation EdTech, AI learning platforms or data management solutions).

Bed

Context: 7 municipalities in five states - Guaramiranga/Ceará, Domingos Mourão/Piauí, Cabrobó, Igarassu and Bonito/Pernambuco, Volta Redonda/Rio Janeiro, Santa Maria/Rio Grande do Sul: who are now piloting 5 EdTech approaches, challenging context (poor municipalities), mainly in north-east of Brazil, characterised by connectivity/ infrastructure challenges.

Testers: 9000 learners and 240 teachers of 35 middle schools (aged 11–14 years, Grades 6–9).

Space: tests are taking place physically in the school.

Scale: 35 schools, 5 EdTechs, 7 municipalities in 5 states involved.

Test

What to test: how to reduce learning gaps and losses in the education public system and what is the role of technology (EdTech) in this context, while testing open innovation / new public procurement mechanisms in

Brazil. Moreover, the program aims to test how to create and adapt pedagogical models and edtech solutions in vulnerable contexts.

Methodology and methods: pre-post surveys/assessments to students to quantify learning gap and exposure to assess learning impact. In addition, an independent researcher/organisation will evaluate the pilots when the testing is finished. Moreover, there will be qualitative/structured interviews and post diagnosis of how the educational public municipalities finished the program.

Timeline and frequency: EdTech selection: November 2021 – February 2022; Pilots: were initially of 9 months (until December 2022), but they were extended until June 23, due to infrastructure challenges at the beginning of the tests.

Leadership: specialised organisation firm / researchers: Oppen Social conducted a process evaluation during implementation of the pilots and Germina will conduct an outcomes report.

Reporting: monthly progress reports discussed with the municipal secretaries of education looking at the access of the teachers and students to the EdTech tools, satisfaction, and perceived quality. The final outcomes report will start in March–April 2023. Dissemination to be determined, initially present the results to the National Council of Secretaries of Education in Brazil, and other events of the financing/implementing partner (Lemann Foundation).

4.7. Spotlight 4: Leanlab Education Codesign Product Research

General information

Leanlab Education is a non-profit research organisation specialising in connecting educational product developers and school communities to research EdTech tools. Unique to Leanlab's research process is the focus on Codesign – a research framework that emphasises the co-creation of research objectives, methodologies, data collection, and dissemination among researchers, educators, and developers.

EdTech

Target group: K–12 students, teachers, administrators, and parents

Pedagogy: Depends on EdTech tool

Technology type: No specification

Innovation stage: Ranges from prototype to established companies

Educational purpose: Allow educators to teach more effectively and improve educational outcomes for all students.

Selection criteria: Leanlab Education conducts a Pilot Ready Audit on all inbound companies focused on the tools' functionality, usability, interoperability, as well as their data privacy and security practices.

Educational content: Depends on the EdTech tool

Bed

Context: Based in Kansas City, MO, USA, but operates in states throughout the USA. Specific focus in Missouri, Kansas, Texas, Florida, New York, and California.

Testers: K–12 teachers, administrators, students, and caregivers.

Space: Tests are conducted inside virtual, in-person, and hybrid schools.

Scale: 37 tests conducted with 20 companies at 30 schools (23 districts) and 7 states.

Test

What to test: Leanlab uses a range of research offerings to test:

1. The perceived value and use case of a product (Sandbox Study);
2. How easily users can navigate a tool (Usability Study);
3. How easy a tool is to use during the regular classroom experience (Feasibility Study);
4. What functionality is needed for a tool to sustain engagement throughout the academic year (Implementation);
5. The relationship between a product's use and its intended outcomes (Correlational Study).

Methodology and methods: Tests range from a sandbox study, using focus groups to assess perceived value of a prototype or concept, to correlational studies that use non-experimental methods and statistical analysis.

Timeline and frequency: Tests are conducted on a rolling basis throughout the academic year and can range from 4 weeks to a full semester (6 months).

Leadership: Leanlab Education

Reporting: Leanlab creates a final written report distributed to all stakeholders, as well as an action plan for the EdTech company with specific participant recommendations. Findings are also explained to stakeholders via a webinar.

Chapter 5. Impact goals

All EdTech Testbeds are driven by a set of highly interrelated goals that underpin their existence. Our desktop research and interviews have revealed the following goals, although we anticipate that there will be more. We suggest that it is important for the more system-wide EdTech Testbeds to develop a clear set of goals that can be used to support the evaluation of the Testbed itself, over time. We describe each goal below.

5.1. Improve learning outcomes and learning experiences

A fundamental goal for the majority of EdTech is improvement of learning outcomes; it follows that EdTech Testbeds serve to provide the real-world settings to iteratively test whether, to what extent, and how EdTech can improve learning outcomes and enhance learning experiences at increasing levels of scale. This can involve drawing on evidence-based approaches like *“teaching at the right level”* in digitalised forms such as so-called intelligent tutoring systems, gamification, and personalised adaptive learning. Testbeds can measure the effectiveness of the design (e.g., algorithms, frequency and length of use, user interface, etc.) as well as the implementation considerations (e.g., teacher support, infrastructure requirements) in terms of how these impact learning outcomes. In addition, EdTech Testbeds provide the opportunity to test how effective EdTech can be with respect to broader educational goals, for example, the nurturing of learner creativity, curiosity, critical thinking, and self-directed learning.

Broader educational goals include the nurturing of learner creativity, curiosity, critical thinking, and self-directed learning.

5.2. Produce financially viable and scalable EdTech

The perspective of EdTech developers and their funders/investors demands that Testbeds support the generation of evidence to give confidence that:

- the EdTech enterprise is financially viable in one or more markets;
- use of the EdTech can be scaled within, and across, the target markets.

This implies that a more connected network of global EdTech Testbeds, with some widely understood methods and approaches, might accelerate the global scaling of the most promising products.

A more connected network of global EdTech Testbeds, with some widely understood methods and approaches, might accelerate the global scaling of the most promising products.

5.3. Develop more effective teaching approaches

EdTech is not only designed for learners, but also teachers. EdTech can be used to support teachers in their classroom teaching, preparation, and professional development. Even when EdTech is learner-facing, effective EdTech designs ensure that the role of the teacher is clearly integrated. EdTech can improve a teacher's effectiveness through providing tools to develop their knowledge and skills, supporting planning and assessment, and helping teachers to communicate, share and critically reflect ([↑EdTech Hub, 2021](#)). It is important to build on existing evidence and evidence gaps of how EdTech can improve teaching (see [↑Hennessy and colleagues, 2022](#)), and then use the EdTech Testbed to test approaches in specific contexts.

5.4. Produce contextualised and equitable EdTech

A major consideration for the education sector is the development of EdTech that serves the needs of diverse learners, which crosses geographical and cultural borders. If this is a major impact goal for the EdTech Testbed, then considerable resources need to be allocated to support all relevant stakeholders involved in the evaluation of the Testbed's progress towards this goal.

Such Testbeds are also likely to adopt highly collaborative and co-design methods, which draw extensively on research evidence. Such approaches include:

- The Codesign Framework²⁰ (developed by Leanlab Education), which provides resources to centre practitioner voice and consider potential power imbalances that exist, particularly in historically marginalised settings,
- The Learner Variability Project²¹ resources (developed by Digital Promise in the US), and
- The sandbox tools developed by EdTech Hub.²²

5.5. Further academic knowledge on EdTech

Some EdTech Testbeds may have a broader set of impact goals that go beyond efficacy and effectiveness, especially those situated within research organisations. Such Testbeds might also test the impact of the EdTech on the human aspects of the classroom. For example, the agency of the teacher and of the students, impacts on social interactions and interrelationships, issues of social equity, diversity, inclusivity and accessibility, ethics and data sharing, and so on.

²⁰ ↑Codesign Framework: <https://www.codesignframework.com/>

²¹ ↑Learner Variability Project: <https://digitalpromise.org/initiative/learner-variability-project/>

²² ↑Tools | Applications for Decision Makers, Implementers and Researchers: <https://edtechhub.org/edtech-tools/>

5.6. Increase stakeholder collaboration

EdTech Testbeds designed with a major impact goal to increase stakeholder collaboration require dedicated resources to both develop the stakeholder network in the local ecosystem and find ways to actively involve key stakeholders in the Testbed processes. It is particularly important to increase collaboration with the education sector, as it can provide a main route to scaling and sustaining both the EdTech Testbed itself, and the EdTech tested within it.

5.7. Generate (better) evidence for specific EdTech products

Developing the use of (better) evidence for a range of stakeholders that reports the outcomes of individual tests within the Edtech Testbed can be an important goal. Such evidence can provide EdTech enterprises to develop working practices that ensure that they can sustain (and possibly improve) the impact of products and services over time. The outcomes of multiple EdTech tests can then form the evidence to evaluate the impact goal for the Testbed itself.

Chapter 6. Lessons learned from more system-wide EdTech Testbeds

Recently, a number of EdTech Testbeds have been established around the world. They come in different shapes and sizes. Some Testbeds are still in the construction phase, others already have partnerships with companies, researchers and schools, and still others may be doing Testbed-like activities but may not advertise themselves as Testbeds. Other Testbeds have now been ended.

There is a need to learn from EdTech Testbed methodologies and share this knowledge with the global community, as a way to move towards more system-wide EdTech Testbeds. Our desk research revealed few reports of EdTech Testbeds evaluating their own methodologies. Below, lessons learned from four EdTech Testbed methodologies are described:

1. Swiss Edtech Collider,
2. WISE (Qatar Testbed),
3. Swedish Edtest,
4. Leanlab Education.

6.1. Lessons learned from the Swiss EdTech Collider

In this part, we describe the lessons learned from the Swiss National EdTech Testbed Program. The goal of this Testbed is to establish a nationwide *"test and learn"* innovation model that enables rapid EdTech testing, structured feedback loops for EdTech products, connecting technology providers with educators, and strengthening teachers' digital skills.

Support is important

Many teachers are at their limit, so they are more likely to be willing to participate in a Testbed if the extra effort is minimal. In addition, teachers often feel more comfortable using new technologies if someone helps them to implement them in the classroom. So professional support on site and online during the testing is a key component for a successful Testbed.

Teachers can be contacted in several ways

One of the most difficult tasks of a Testbed is reaching teachers who are open to participating in a Testbed. It has been shown that a combination of different ways of contacting teachers is effective. For the first tests, the network of the Testbed facilitator (friends and colleagues) is crucial. These teachers are often very open to participation. To expand the circle of participants beyond one's own network, the following has proven to be effective: participation in teacher events, reaching out to school principals, working together with teacher universities and other important educational organisations. However, it is essential to explain how participating in an EdTech Testbed can be beneficial for teachers, so they understand the value of participating.

Real-world feedback is an added value for companies

Many companies in the EdTech sector lack regular access to schools, and thus, to real-world feedback. Therefore, companies greatly appreciate receiving feedback on their product from the Testbed. However, it is important that the feedback is detailed and specific so that the companies know what they need to adjust about their product in concrete terms. The observations during testing by the Swiss EdTech Collider team are also very helpful for companies (e.g., *"the registration process took too long / was too complicated for this age group"*).

The test phase must be designed individually

Products differ so much from each other that a new test needs to be designed for each product in terms of length of the test phase, type of support, and questions for the evaluation. Moreover, all companies are at different stages and therefore need different types of feedback. While late-stage companies often want

feedback on a specific new feature, early-stage companies want more general feedback on user usability, for example. As a result, it is important to take enough time in advance to find out why a company wants to participate in a Testbed and what they hope to achieve. No company only interested in marketing their product and not in further development should be included .

6.2. Lessons learned from WISE Report (Qatar Testbed)

WISE launched Qatar’s first EdTech Testbed in 2020. The goal was to improve teachers’ understanding of the use of technology in the classroom and to identify how technology can be used to overcome persistent problems in the classroom. The four key lessons learned from implementing an EdTech Testbed during the pilot phase between 2020 and 2021 are outlined in the sections below ([Lee & Basma, 2021](#)).

Building relationships with school leaders early

First, thanks to building relationships with schools early on, there is enough time to demonstrate the benefits of participating in the Testbed and establish clear expectations and norms. The teachers also need sufficient time to plan the upcoming term with the EdTech product in mind.

Test the EdTech product before actual implementation

Second, before an EdTech product is tested in schools, it is essential to check whether the products/platforms are accessible in the schools, if they work, and what the potential broader impacts are. It is helpful to work together with the school IT departments to clear security requirements in advance. This way you don’t lose time and energy during the testing to get the platform running.

Identify an “EdTech” champion

Third, having a “*champion teacher*” can help you encourage other teachers to use EdTech products, brainstorm, and participate in the Testbed. This person can also motivate other teachers to use the EdTech product long enough to evaluate the product’s effectiveness. For a Testbed to be successful, teachers’ willingness

to try something new is essential. They must also be eager to invest their time to learn how to use the new EdTech products.

Involve the EdTech venture in ongoing professional development of teachers

Fourth, by incorporating the EdTech venture into ongoing professional development, teachers will have the opportunity to be trained in the use of a particular product. In addition, they will be encouraged and challenged to consider how to integrate technology with their teaching methods and specific classroom needs. This can result in teachers gaining knowledge in both educational technology and design thinking, leading to more innovative teaching practices.

6.3. Lessons learned from the Swedish Edtest

The Swedish Edtest was launched in January 2020. This initiative aims to improve the digital skills of teachers, develop better digital tools for education, and bridge the gap between customers and suppliers by understanding the real need of users. Over the first two years of the programme, over 200 educators and more than 50 EdTech companies participated. In the following section, some lessons learned shared by Hanna Elving, lead project manager of Swedish Edtest, are summarised.

Blessing from the principal, but passion from the teachers

For a successful implementation of a Testbed, it is essential that both the teachers and principals are fully engaged. The teachers, on the one hand, must understand and believe in the value of the Testbed, and the principals, on the other hand, must provide support and resources for the Testbed to be incorporated into their job or competence development.

Special support for the school leaders

Involving the principals and providing them with support is crucial, while teachers often collaborate and share information among themselves, principals may require a separate platform for exchanging their experiences and learning from

one another. To achieve this, the Swedish Edtest are planning to design a series of workshops where principals can share their lessons learned and discuss any challenges they have encountered.

Design Testbeds for a low budget

Given that the education sector often operates on a limited budget, it is important to design a Testbed that can run with minimal financial resources. This is also important to finance a Testbed sustainably over a longer period of time.

Try to avoid test fatigue

Avoiding test fatigue among teachers can be a real challenge, and it is therefore important to involve teachers meaningfully. By helping teachers to connect with one another and making it clear what benefits they will gain from participating, will mean teachers are more likely and willing to take part in the Testbed multiple times.

6.4. Lessons learned from Leanlab Education

Leanlab Education began 10 years ago, and transitioned from an EdTech accelerator to a Testbed in 2019. It should be noted that Leanlab does not use the terminology “*Testbed*” to describe their work. This is done to avoid the negative connotations, in the USA, of testing EdTech on students and teachers. Their goal is to elevate the expertise of educators, parents, and students in the design and development of EdTech tools through systematic research.

Leanlab does not use the terminology “Testbed” to describe their work to avoid negative connotations.

Focus on formative research for EdTech tools

In 2022 Leanlab Education contracted a third-party researcher (Edtech Recharge) to conduct an analysis²³ on the 14 most recent research studies and identify common themes for usability, feasibility, implementation, and correlational studies. The biggest takeaway from the analysis was the need to weigh and fund formative research more heavily than efficacy research to support stronger foundations in EdTech.

Ensure the “Bed” and the “EdTech” are prepared to engage in research

Leanlab Education has developed the *“Pilot Ready Audit”* for EdTech companies and the *“Research Readiness Index”* for schools to assess the readiness for both parties to engage in research. For schools, these indicators are focused on their *“innovation mindset”*, human capacity, and use of data in decision-making. The audit for companies focuses on the functionality, usability, and interoperability, as well as their data privacy policies and evidence of prior research. This work to vet companies and schools before starting the research process allows Leanlab to conduct research with lower risk of participant attrition and a greater chance that EdTech companies will implement participant feedback in product development.

Don’t neglect data and security protocols

Schools are rightfully hesitant to engage with a third-party researcher and share data when it contains any student information. Researchers have an obligation to handle the access they are given respectfully and allay any hesitations through clear agreements. Leanlab uses multi-year legal agreements with schools and proactively negotiates IT protocols and data-sharing agreements that can help remove roadblocks once studies get closer to implementation. In larger school systems, careful power-mapping and clarification of internal protocols for approvals (including an Institutional Review Board or Ethical Review Board) is necessary.

²³ Edtech Research Trends: <https://www.leanlabeducation.org/edtech-research-trends>

Compensate educators and research participants for their work

Although good Testbeds should attempt to ensure that all participants and stakeholders involved in the research benefit from the test and feel valued and heard throughout the process, it is important to also ensure that the participants are fairly compensated for the work in the form of direct stipends to research participants. Leanlab also provides schools with unrestricted grants for participation in exchange for their willingness to participate in tests aligned to their priorities.

Ensure the bed is diverse, representative, and has buy-in

Leanlab Education has experimented with building an intentionally diverse portfolio of "*Testbed*" school communities that represent both the geographic and contextual diversity of the United States. In an attempt to both increase buy-in from school partners and increase opportunities to run different types of "*tests*". This diversity of school partners also allows for the potential to conduct studies with larger n sizes across different contexts.

Chapter 7. Key challenges for EdTech Testbed design

There are several factors that hold back more system-wide Edtech Testbeds. Based on desk research and survey responses, we have insights into the main challenges key stakeholders face when setting-up and/or running EdTech Testbeds. We summarise these challenges below.

7.1. Low implementation fidelity of EdTech in schools

Stakeholders report low implementation fidelity of EdTech in schools, meaning that schools and teachers implement the EdTech differently than expected. There could be many reasons for this.

First, it could be due to limited access to equipment and the time teachers can spend testing EdTech. Second, teachers usually do not have much experience in generating and acting on evidence. Therefore, a school's ability to use Testbeds for direct professional development for teachers and school administrators needs to be strengthened.

For example, teachers should be better informed about what evidence generation entails and how they can contribute to it (not just at the beginning of the test). As such, capacity among educators and their institution to carry out robust evaluations will be improved.

Third, schools currently test EdTech often for other reasons (e.g. because access to the EdTech is free, or for the kudos gained through their involvement in Testbed initiatives) rather than testing it because it actually meets their defined needs.

Fourth, perhaps the EdTech is badly designed and difficult to implement faithfully – a problem that Testbeds are well-suited to counter through formative research. EdTech developers need to understand how real classrooms work, and should not

automatically expect teachers to change their practice to fit EdTech, rather than the other way around.

7.2. Difficult to recruit and engage schools, teachers, learners, and parents

Stakeholders emphasise the importance of engaging local partners to ensure that the local context is maximally considered when testing EdTech. A recent trends study found teachers' roles to be underdeveloped or limited in EdTech design ([UNICEF, 2022](#)). However, many stakeholders report difficulties finding schools, teachers, and parents to partner with (depending on whether it is an in-school, after-school, or at-home Testbed model). Furthermore, there are many differences between schools. Several stakeholders fear that only a small sample of innovative schools (who are highly motivated) are testing EdTech, which can lead to biased conclusions. Several reasons have been mooted to explain why schools are not willing to engage in EdTech Testbeds. First, evidence generation can feel like it is in the service of developers and not in service of education or learners. Second, company and researchers' discourse is often different from the language teachers use. Teachers use different terminology, the needs/problems/expectations of schools and teachers may be different from EdTech developers. One resource that can be helpful when recruiting schools, teachers, learners, and parents is the [MASTER framework](#) produced by EdTech Hub ([Rahman and colleagues, 2022](#)). They explain several key principles which can be considered when thinking about involving schools.

7.3. Procurement and contract difficulties

EdTech Testbed reports and stakeholders emphasise that the administration during procurement needs to be kept to a minimum. Also, setting up contracts before use of technology by schools is often time-consuming, discouraging schools to engage in EdTech Testbeds ([Eriksson and colleagues, 2018](#)). Furthermore, stakeholders mention that procurement processes often discourage adaptation and testing, leading to a rigid test plan of the EdTech by default. Still further, procurement processes do not focus enough on outcomes and efficacy, which impacts the way purchasing and piloting decisions are made.

7.4. Lack of government involvement

Stakeholders report that governments need more encouragement to see the importance of stakeholders' roles in finding long-term sustainable changes that connect EdTech companies, education providers, and researchers to drive educational innovation through Testbeds and technology. One issue is that governments are under-resourced or have little understanding of how they can promote collaboration within EdTech testbeds.

In the United States, while there are grants available for EdTech companies to conduct research,^{24,25} this funding supports education technology providers working with independent researchers, rather than developing research and development infrastructure directly within schools.

To enable the scaling of EdTech, EdTech Testbeds should be designed, implemented, and evaluated in collaboration with governments.

7.5. Difficult to reach scalable and sustainable EdTech Testbeds

A key issue with EdTech Testbeds is “*pilotitis*” where Testbeds do not scale (↑[Principles for Digital Development, 2022](#)). We define scale of EdTech Testbeds in terms of stage and size. “*Stage*” refers to the maturity of the Testbed, how many people work for the Testbed, and their positions. “*Size*” refers to the number of schools, researchers, and companies involved in the Testbed, and how many EdTech products have been tested. For example:

- Stakeholders report that it is often difficult to have sustained funding to continue EdTech Testbed activities.
- Lack of sustainable infrastructure.

Scaling is a challenge for education interventions in general, and is often catalysed when EdTech is brought into the mix. For example, ↑[Evans & Yuan \(2022\)](#)

²⁴ ↑[Institute of Education Sciences: Small Business Innovation Research Program \(SBIR\): https://ies.ed.gov/sbir/](https://ies.ed.gov/sbir/)

²⁵ ↑[IES Funding Opportunities: Education Research Grants Programs: https://ies.ed.gov/funding/ncer_progs.asp](https://ies.ed.gov/funding/ncer_progs.asp)

reviewed 234 causal education studies in LMICs and found that effect sizes are cut in half when education projects scale from under 500 participants to over 5,000 participants.

7.6. Lack of test procedures

Stakeholders report many difficulties determining the test procedures for EdTech for several reasons, as listed below.

1. Stakeholders report not having sufficient knowledge of the methodologies and methods available. The process of deciding what data to collect and how to automate and process it is challenging.
2. The reproducibility and scalability of tests is limited, EdTech product types can be very different and therefore require different methodologies and/or methods for testing.
3. Test results often do not represent the scale and diversity of target populations to ensure the results can be generalised. This requires tests to be at a scale of at least a few regions, districts (within or across countries).
4. Learner and context characteristics have only been taken into account in a limited way in test procedures thus far. Important characteristics may include language, access to technology (devices, network), learning gaps and learning patterns based on social and economic disparities.
5. Ethical and privacy issues are a major concern due to participation of students and teachers, as is the intellectual property of researchers when they are involved in experimentation.
6. Translating results of tests to practitioners is a difficult process due to a lack of shared language between researchers/EdTech developers and teachers, and due to the difficulty that test results are rapidly outdated given fast changing technology.
7. Rewarding participation in testing by providing certification can be an incentive for companies and schools. However, stakeholders also warn of potential risks of unfair competition between companies ([↑Boot and colleagues, 2022](#)).

7.7. Conflicting goals and roles of different stakeholders

Another common challenge is that stakeholders participate in EdTech Testbeds with their own goals and agendas (see [Appendix B](#)). The differing motivations of various stakeholders' can be difficult to reconcile. For example, while some researchers might prefer multi-year control group trials, EdTech entrepreneurs want short pilots with clear and positive results! EdTech Testbeds can play an important role in aligning different stakeholders by informing them of what to expect and discussing each of their motivations. EdTech Testbeds need to convey a clear Testbed vision and support all stakeholders to engage with this.

7.8. Lack of support after being involved in a Testbed

Often, start-up EdTech entrepreneurs participate in EdTech Testbeds. After being involved in a Testbed, it is important to support early stage companies to build on the experience.

Chapter 8. Conclusion: Towards more effective EdTech Testbeds

In this report, we attempted to determine a pathway towards systemic EdTech Testbeds, bearing in mind the global perspective, with a particular focus on effectiveness, efficacy, and learning outcomes. In the first part of the report, we identified key components of EdTech Testbeds and presented commonly used methodologies and methods ([Chapter 3. Pre-conditions for effective EdTech Testbeds](#), [Chapter 4. Key components of EdTech Testbeds: the EdTech, the Test and the Bed!](#), [Chapter 5. Impact goals](#)).

In the second part ([Chapter 6. Lessons learned from more system-wide EdTech Testbeds](#)), we summarised lessons learned from three particular system-wide EdTech Testbeds (WISE, Swiss National EdTech Testbed Program, Swedish Edtest). The third part ([Chapter 7. Key challenges for EdTech Testbed design](#)), we conclude with key challenges for effective EdTech Testbed design as we work towards a wider global offer of more system-wide EdTech Testbeds.

Given the relatively dispersed efforts in the area of EdTech Testbeds globally, there are several areas where more cooperation could pay off quickly.

What evaluations of EdTech Testbed methodologies have been carried out?
What improvements have been made?

8.1. Collaboration on evidence creation

As noted by many stakeholders, generating rigorous actionable evidence on EdTech is critical to ensure learning outcomes and value for money, as well as ensuring equity across different populations of learners, including those who are

marginalised. A crucial initial question in this is what evaluations of EdTech Testbed methodologies have been carried out. Further, based on the evaluation of the methodologies, what (iterative) improvements have been made?

Evidence creation should be a pluralistic and inclusive process, recognising a range of methods (see [Chapter 12. / Appendix C. Commonly used EdTech Testbed Methodologies and Methods](#)).

8.2. Collaboration on EdTech Testbed design

Rigorous evidence can only be collected if there are opportunities for the collection of that evidence. Therefore, we need to collaborate on EdTech Testbeds, including their design and implementation methods, as well as their evaluation ([8.1. Collaboration on evidence creation](#)). We might say that, on the basis of rigorous evidence, we need to build an ‘implementation science’ for Testbeds: Ranging from simple prompts to audit one specific EdTech Testbed approach and perhaps identify missing aspects, to comprehensive frameworks.

We need to build an ‘implementation science’ for EdTech Testbeds.

To build such an implementation science, several components are needed. It is imperative to develop a shared understanding of EdTech Testbeds. Testbeds do not need to follow identical principles, but we need to make Testbed designs comparable, perhaps even interoperable. In other words, there have to be mutually agreed comparable frameworks. Frameworks should be sufficiently standardised so that evaluations become comparable. With such a systemic understanding, evidence can be compared across approaches, enabling broader benefits from Testbeds around the globe.

8.3. Development of a global EdTech Testbeds network

You, dear reader, might feel that those last two goals are perhaps rather obvious, and that many such calls for collaboration have already been made. The issue at

hand may indeed not be the awareness of those goals expressed in writing, but rather how to meet those goals in practice.

Clearly, it would be beneficial for some kind of community or network to be formed. Specifically, an inclusive community of EdTech Testbed practice and practitioners might be able to achieve those goals. Together, such a community might not only share knowledge, practice, resources, and opportunities but also build relations with policymakers and funders. To form a viable network, participants need to experience a net gain in return for their investment. The focus of the EdTech Testbeds event in March 2023 will be to discuss the creation of such a network, the Global EdTech Testbeds Network (GETN), and to determine the desirable resources and conditions under which such a network can thrive.

Ultimately, such a network could help not only accelerate progress towards sustainable models and approaches, but make a significant contribution to the achievement of Sustainable Development Goal 4, to provide a quality education for everyone.

The focus of the EdTech Testbeds event in March 2023 will be to discuss the creation of such a network, the Global EdTech Testbeds Network (GETN).

Chapter 9. References

This bibliography is available digitally in our evidence library at

<https://docs.opendeved.net/lib/KGWBCB4D>

BIRD-E. (n.d.). BIRD-E. Retrieved February 2, 2023, from <https://www.bird-e.org>.
([details](#))

Codesign Framework. (n.d.). Codesign. Retrieved February 2, 2023, from
<https://www.codesignframework.com>. ([details](#))

Digital Promise – Research Map. (n.d.). Research Map. Retrieved February 2, 2023,
from <https://researchmap.digitalpromise.org/>. ([details](#))

Digital Promise. (n.d.). Digital Promise. Retrieved February 2, 2023, from
<https://digitalpromise.org/>. ([details](#))

EdTech Hub. (n.d.). EdTech Hub. Retrieved February 2, 2023, from
<https://edtechhub.org/>. ([details](#))

Educational Technology Rapid Cycle Evaluations. (n.d.). Office of Educational
Technology. Retrieved February 2, 2023, from <https://tech.ed.gov/rce/>.
([details](#))

Exploring the EdTech Testbed Ecosystem. (n.d.). European Edtech Alliance.
Retrieved February 2, 2023, from
<https://www.edtecheurope.org/news/exploring-the-edtech-testbed-ecosystem>. ([details](#))

Goldstar EdTech Diagnostics. (n.d.). Goldstar Education. Retrieved February 2,
2023, from <https://www.goldstared.com>. ([details](#))

IES Funding Opportunities: Education Research Grants Programs. (n.d.). [Indexes;
Offices]. Retrieved February 2, 2023, from
https://ies.ed.gov/funding/ncer_progs.asp. ([details](#))

Institute of Education Sciences: Small Business Innovation Research Program (SBIR). (n.d.). [Indexes; Offices]. Retrieved February 2, 2023, from <https://ies.ed.gov/sbir/>. (details)

Is your school making the most of digital technologies for learning? (n.d.). Retrieved February 2, 2023, from <https://schools-go-digital.jrc.ec.europa.eu/>. (details)

Learner Variability Project. (n.d.). Digital Promise. Retrieved February 2, 2023, from <https://digitalpromise.org/initiative/learner-variability-project/>. (details)

Pilot Readiness Audit. (n.d.). Leanlab Education. Retrieved February 2, 2023, from <https://www.leanlabeducation.org/pilot-readiness-audit>. (details)

Project: Extending design thinking with emerging digital technologies (Exten.D.T.2). (2022, September 6). Lnu.Se. <https://lnu.se/en/research/research-projects/project-extending-design-thinking-with-emerging-digital-technologies/>. (details)

Swedish Edtest – Frequently asked questions. (n.d.). Retrieved February 2, 2023, from <https://edtest.se/en/frequently-asked-questions>. (details)

Swedish Edtest – Utveckla digitala lärresurser med Swedish Edtest. (n.d.). Retrieved February 2, 2023, from <https://edtest.se/sv>. (details)

Swiss National EdTech Testbed Program | Gebert Rűf Stiftung. (n.d.). Retrieved February 2, 2023, from <https://www.grstiftung.ch/de/media/portfolio~grs-074-21~.html>. (details)

Testbed Helsinki / EdTech. (n.d.). Testbed Helsinki. Retrieved February 2, 2023, from <https://testbed.hel.fi/en/edtech/>. (details)

Albers, B., Shlonsky, A., & Mildon, R. (Eds.). (2020). *Implementation Science 3.0*. Springer. (details)

American Institute for Research. (2018). *Innovative Designs | Sequential, Multiple Assignment, Randomized Trial*. <https://optimizingoutcomes.air.org/innovative-designs.php>. (details)

American Institute for Research. (2018). *Rapid-Cycle Evaluations*. <https://optimizingoutcomes.air.org/rapid-cycle.php>. (details)

- Armstrong, M., Dopp, C., & Welsh, J. (2020). Design-Based Research. *The Students' Guide to Learning Design and Research*.
https://edtechbooks.org/studentguide/design-based_research. (details)
- Atherton, P. (2019). Bridging the chasm – study of the realities of edtech use among trainee teachers. *Teacher Education Advancement Network Journal*, 11(4), 80–95. (details)
- Batty, R., Wong, A., Florescu, A., & Sharples, M. (2019). *EdTech testbeds: Models for improving evidence* (p. 44).
<https://www.science-practice.com/teams/good-problems/projects/edtech-testbeds/edtech-testbeds-report.pdf>. (details)
- Boot, N., Jukema, H., & Stappen, E. van der. (2022). *Onder de motorkap: Onderzoek naar EdTech testbeds in Europa*.
<https://www.versnellingsplan.nl/Kennisbank/onder-de-motorkap-onderzoek-naar-edtech-testbeds-in-europa/>. (details)
- Boujikian, M., Carter, A., & Jordan, K. (2022). The Sandbox Model: A Novel Approach to Iterating while Implementing an Emergency Education Program in Lebanon during the COVID-19 Pandemic. *Journal on Education in Emergencies*, 8(3), 215. <https://doi.org/10.33682/rj45-k7z7>. Available from <https://archive.nyu.edu/handle/2451/63993>. (details)
- Building Evidence in Education. (2023). *Implementation Research in Education – BE2 Guidance Note*. (details)
- Chang, A. M. (2018). *Lean Impact: How to Innovate for Radically Greater Social Good* (1st edition). Wiley. (details)
- Clark-Wilson, A., Bashir, A., Ahmed, S., Mazari, H., Kaye, T., Radford, K., & Otieno, J. (2022). *A Theory of Change for Parents and Caregivers Towards a Technology-Enhanced Education System in Bangladesh*. EdTech Hub.
<https://doi.org/10.53832/edtechhub.0126>. Available from <https://docs.edtechhub.org/lib/GQSKFD62>. (details)
- Cohen, L., Manion, L., & Morrison, K. (2017). *Research Methods in Education* (8th edition). Routledge. (details)
- Cristol, D., Al-Sabbagh, S., Abdulbaki, A., Majareh, M., Tuhin, S. U., & Gimbert, B. (2019). Innovative delivery of education in Bangladesh using mobile

technology. *World Conference on Mobile and Contextual Learning*, 37–45. ([details](#))

Crotty, M. J. (2020). *The Foundations of Social Research: Meaning and Perspective in the Research Process* (First edition). SAGE Publications Ltd. ([details](#))

Crowley, K. (2016). *Design-Based Research and Design-Based Implementation Research* | *InformalScience.org*.
<https://www.informalscience.org/news-views/design-based-research-and-design-based-implementation-research>. ([details](#))

Cukurova, M., Luckin, R., & Clark-Wilson, A. (2019). Creating the golden triangle of evidence-informed education technology with EDUCATE. *British Journal of Educational Technology*, 50(2), 490–504. <https://doi.org/10.1111/bjet.12727>. Available from <https://onlinelibrary.wiley.com/doi/abs/10.1111/bjet.12727>. ([details](#))

Eccles, M. P., Armstrong, D., Baker, R., Cleary, K., Davies, H., Davies, S., Glasziou, P., Illott, I., Kinmonth, A.-L., & Leng, G. (2009). *An implementation research agenda*. BioMed Central. ([details](#))

EdTech Hub. (2021). Curated Tools for Teacher Continuous Professional Development. *EdTech Hub*.
<https://edtechhub.org/edtech-tools/curated-tools-tcpd/>. ([details](#))

EdTech Hub. (2022). *Sandbox Handbook V.2.0*. EdTech Hub. ([details](#))

EdTech Hub. (2022). The Impact of a Tech-Supported, School-Based TPD Model on Learning Outcomes in Tanzania. *EdTech Hub*.
<https://edtechhub.org/evidence/edtech-hub-research-portfolio/study-4/>. ([details](#))

EdTech Tulna. (2021). *Standards – EdTech Tulna*.
<https://edtechtulna.org/standards/>. ([details](#))

Eriksson, E., Heath, C., Ljungstrand, P., & Parnes, P. (2018). Makerspace in school: considerations from a large-scale national testbed. *International Journal of Child-Computer Interaction*, 16, 9–15. <https://doi.org/10.1016/j.ijcci.2017.10.001>. ([details](#))

- Escueta, M., Quan, V., Nickow, A. J., & Oreopoulos, P. (2017). *Education Technology: An Evidence-Based Review* (No. w23744; p. w23744). National Bureau of Economic Research. <https://doi.org/10.3386/w23744>. Available from <http://www.nber.org/papers/w23744.pdf>. (details)
- Evans, D. K., & Yuan, F. (2022). How Big Are Effect Sizes in International Education Studies? *Educational Evaluation and Policy Analysis*, 44(3), 532–540. <https://doi.org/10.3102/01623737221079646>. (details)
- Evidence & ESSA: What Is Evidence And How Should We Use It? (n.d.). R4S. Retrieved February 2, 2023, from <http://www.research4schools.org/evidence-essa/>. (details)
- Fishman, B., Cheng, B., & Penuel, W. (2019). *CIRCL Primer: Design-Based Implementation Research*. CIRCL. <http://circlcenter.org/dbir/>. (details)
- Forman, S. G., Shapiro, E. S., Coddling, R. S., Gonzales, J. E., Reddy, L. A., Rosenfield, S. A., Sanetti, L. M., & Stoiber, K. C. (2013). Implementation science and school psychology. *School Psychology Quarterly*, 28(2), 77. <https://doi.org/10.1037/spq0000019>. (details)
- Global Victoria EdTech Innovation Alliance. (2022). *Activating research-based testbeds in Victorian and international education environments*. (details)
- Hennessy, S., D'Angelo, S., McIntyre, N., Koomar, S., Kreimeida, A., Cao, L., Brugha, M., & Zubairi, A. (2022). Technology Use for Teacher Professional Development in Low- and Middle-Income Countries: A systematic review. *Computers and Education Open*, 3, 100080. <https://doi.org/10.1016/j.caeo.2022.100080>. Available from <https://www.sciencedirect.com/science/article/pii/S2666557322000088>. (details)
- Higgins, S., Xiao, Z., & Katsipataki, M. (2012). The Impact of Digital Technology on Learning: A Summary for the Education Endowment Foundation. Full Report. *Education Endowment Foundation*. (details)
- Hollands, F. M., & Escueta, M. (2017). *EdTech Decision-making in Higher Education*. (details)
- Holon IQ. (2022). (details)

- IDEO. (2013). *Design Thinking for Educators*.
<https://www.ideo.com/post/design-thinking-for-educators>. (details)
- IDEO. (2015). *Design Kit: The Human-Centered Design Toolkit*.
<https://www.ideo.com/post/design-kit>. (details)
- Kaiser, C. (n.d.). *What Is Agile Methodology: A Primer on Moving Fast*. Wellfound (Formerly AngelList Talent). Retrieved December 22, 2022, from
<https://angel.co/blog/agile-methodology-a-primer-on-moving-fast>.
(details)
- Kerwin, J. T., & Thornton, R. L. (2020). Making the Grade: The Sensitivity of Education Program Effectiveness to Input Choices and Outcome Measures. *The Review of Economics and Statistics*, 1–45. https://doi.org/10.1162/rest_a_00911.
(details)
- Khalayleh, A. (2020, April 3). *Arabic version of Design Thinking for Educators*.
<https://opendeved.net/2020/04/03/arabic-version-of-design-thinking-for-educators/>. (details)
- Kim, J. S., Asher, C. A., Burkhauser, M., Mesite, L., & Leyva, D. (2019). Using a Sequential Multiple Assignment Randomized Trial (SMART) to Develop an Adaptive K–2 Literacy Intervention With Personalized Print Texts and App-Based Digital Activities. *AERA Open*, 5(3), 2332858419872701.
<https://doi.org/10.1177/2332858419872701>. (details)
- Kizilcec, R. F., Reich, J., Yeomans, M., Dann, C., Brunskill, E., Lopez, G., Turkay, S., Williams, J. J., & Tingley, D. (2020). Scaling up behavioral science interventions in online education. *Proceedings of the National Academy of Sciences*, 117(26), 14900–14905. <https://doi.org/10.1073/pnas.1921417117>. Available from <https://pnas.org/doi/full/10.1073/pnas.1921417117>. (details)
- Knight, S., Moeini, A., & Clark-Wilson, A. (2020, April 5). Using evidence to help build and evaluate good ideas in education technology. *EduResearch Matters*.
<https://www.aare.edu.au/blog/?p=5259>. Available from
<https://www.aare.edu.au/blog/?p=5259>. (details)
- Koehler, M. J., Mishra, P., & Cain, W. (2013). What is technological pedagogical content knowledge (TPACK)? *Journal of Education*, 193(3), 13–19.
<https://doi.org/10.1177/002205741319300303>. (details)

- Kolak, J., Norgate, S. H., Monaghan, P., & Taylor, G. (2021). Developing evaluation tools for assessing the educational potential of apps for preschool children in the UK. *Journal of Children and Media*, 15(3), 410–430.
<https://doi.org/10.1080/17482798.2020.1844776>. (details)
- Koomar, S., Adam, T., Massam, W. E., Anthony, G., Mrope, W. J., Mtenzi, F., Mwakabungu, F., Komba, A., Hennessy, S., & Barretto, J. (2022). *Design-Based Implementation Research Baseline Data Collection: Technical report*. EdTech Hub. (details)
- Kucirkova, N. (2022). *Understanding evidence: a brief guide for edtech produceers*. (details)
- Leanlab Education. (n.d.). *Edtech Research Trends*. Retrieved February 2, 2023, from <https://www.leanlabeducation.org/edtech-research-trends>. (details)
- Lee, S. S., & Basma, V. (2021). *Piloting Qatar's First Edtech TestBed: The Journey Towards Best Practice, Sustainable Partnerships and Innovation*. WISE, Qatar Foundation.
<https://www.wise-qatar.org/app/uploads/2022/01/2021wise-rr7-report-web-version.pdf>. (details)
- Lupton, D., & Williamson, B. (2017). The datafied child: The dataveillance of children and implications for their rights. *New Media & Society*, 19(5), 780–794.
<https://doi.org/10.1177/1461444816686328>. (details)
- Lurvink, A.-F., Bangura, Z., Fortune, A., George, A. H., Godwin, K., Mason-Sesay, M., Moseray, D., Tholley, I., Turay, E. B., & Haßler, B. (2022). *What are the enablers and barriers for effective School-Based Teacher Professional Development?* OpenDevEd. (details)
- Massachusetts Institute of Technology. (2016). *A Framework for Evaluating Appropriateness of Educational Technology Use in Global Development Programs*.
https://dspace.mit.edu/bitstream/handle/1721.1/115340/Full%20Report_A%20framework%20for%20Evaluating%20Appropriateness%20of%20Educational%20Technology%20Use%20in%20Global%20Development%20Programs.pdf?sequence=1&isAllowed=y. Available from
https://dspace.mit.edu/bitstream/handle/1721.1/115340/Full%20Report_A%20framework%20for%20Evaluating%20Appropriateness%20of%20Educational%2

[0Technology%20Use%20in%20Global%20Development%20Programs.pdf?sequence=1&isAllowed=y](#). (details)

McBurnie, C. (2021). *Navigating the 'Data Revolution': A Case Study on the One Tablet Per School Programme in Sierra Leone*. EdTech Hub. (details)

McNiff, J. (2017). *Action Research: All You Need to Know* (First edition). SAGE Publications Ltd. (details)

Moeini, A. (2020). *Theorising Evidence-Informed Learning Technology Enterprises: A Participatory Design-Based Research Approach* [Doctoral, UCL (University College London)]. <https://discovery.ucl.ac.uk/id/eprint/10117895/>. (details)

Nesta. (2016). *Using research evidence - A practice guide*. https://media.nesta.org.uk/documents/Using_Research_Evidence_for_Success_-_A_Practice_Guide.pdf. (details)

Pratihast, S., & Mote, E. (2022). *Blueprint for Inclusive Research & Development in Education: A Data-Driven Approach to Organize and Articulate Research Data Needs for Equitable Evaluation*. 19–24. <https://doi.org/10.54808/ICSIT2022.01.19>. Available from <https://www.iiis.org/DOI2022/EB878UD>. (details)

Principles for Digital Development. (2022). *Principles for Digital Development*. <https://digitalprinciples.org/principles/>. (details)

Puttick, R., & Ludlow, J. (2013). *Standards of evidence: an approach that balances the need for evidence with innovation*. Nesta. (details)

Rahman, A., Carter, A., Wambûi Kuria, C., Mbugua, C., Simpson, L., & Kably, N. (2022). *Behaviour change key principles for EdTech | MASTER framework*. EdTech Hub. <https://doi.org/10.53832/edtechhub.0129>. Available from <https://docs.edtechhub.org/lib/BT283J6U>. (details)

Rahman, A., Carter, A., Wambûi Kuria, C., Mbugua, C., Simpson, L., & Kably, N. (2022). *Behaviour change key principles for EdTech | MASTER framework* [Working Paper]. EdTech Hub. <https://doi.org/10.53832/edtechhub.0129>. Available from <https://docs.edtechhub.org/lib/BT283J6U>. Available under Creative Commons Attribution 4.0 International. (details)

- Ramiel, H. (2019). User or student: constructing the subject in Edtech incubator. *Discourse: Studies in the Cultural Politics of Education*, 40(4), 487–499. <https://doi.org/10.1080/01596306.2017.1365694>. Available from <https://www.tandfonline.com/doi/full/10.1080/01596306.2017.1365694>. (details)
- Results: Smart Learning Environments of the Future. (n.d.). 6Aika. Retrieved February 2, 2023, from <https://6aika.fi/en/project/results-smart-learning-environments-of-the-future/>. (details)
- Robinson, J. B. (1990). Futures under glass: A recipe for people who hate to predict. *Futures*, 22(8), 820–842. [https://doi.org/10.1016/0016-3287\(90\)90018-D](https://doi.org/10.1016/0016-3287(90)90018-D). Available from <https://www.sciencedirect.com/science/article/pii/001632879090018D>. (details)
- Rodriguez-Segura, D. (2020). EdTech in Developing Countries: A Review of the Evidence. *World Bank Research Observer*, 37(2), 171–203. <https://doi.org/10.1093/wbro/lkab011>. Available from <https://elibrary.worldbank.org/doi/10.1093/wbro/lkab011>. (details)
- Sailer, M., Murböck, J., & Fischer, F. (2021). Digital learning in schools: What does it take beyond digital technology? *Teaching and Teacher Education*, 103, 103346. <https://doi.org/10.1016/j.tate.2021.103346>. Available from <https://www.sciencedirect.com/science/article/pii/S0742051X21000706>. (details)
- Samji, S., Andrews, M., Pritchett, L., & Woolcock, M. (2018). *PDIA toolkit: A DIY Approach to Solving Complex Problems*. Center for International Development at Harvard University. https://bsc.cid.harvard.edu/files/bsc/files/pdiatoolkit_ver_1_oct_2018.pdf. (details)
- Schurr, M. (2020). *Design Thinking for Educators – Designers Workbook (English/Arabic)*. <https://doi.org/10.5281/zenodo.3733139>. (details)
- Schurr, M. (2020). *Design Thinking for Educators – Toolkit (English/Arabic)*. <https://doi.org/10.5281/zenodo.3552082>. (details)

- Siko, J., & Barbour, M. (2016). Building a Better Mousetrap: How Design-Based Research was Used to Improve Homemade PowerPoint Games. *TechTrends*, 60. <https://doi.org/10.1007/s11528-016-0092-x>. (details)
- Simpson, L. (2019, September 18). Behavioural Innovation: embracing behavioural science for lasting impact. *Brink Blog*. <https://medium.com/hellobrink-co/behavioural-innovation-embracing-behavioural-science-for-lasting-impact-85d1d78688f4>. (details)
- Tools | Applications for Decision Makers, Implementers and Researchers. (n.d.). *EdTech Hub*. Retrieved February 2, 2023, from <https://edtechhub.org/edtech-tools/>. (details)
- UNICEF. (2022). *Trends in digital personalized learning in low- and middle-income countries: Executive Summary*. UNICEF. <https://www.unicef.org/globalinsight/media/2756/file/UNICEF-Global-Insight-Digital-PL-LMIC-executive-summary-2022.pdf>. (details)
- USA Department of Education. (2022). *Educational Technology Rapid Cycle Evaluations*. Office of Educational Technology. <https://tech.ed.gov/rce/>. (details)
- Ubaidullah, A. (2016, June 10). *DESIGN BASED IMPLEMENTATION RESEARCH (DBIR) FOR SUSTAINING INNOVATION IN CLASSROOM INSTRUCTION. CHANGING CLASSROOMS -CHALLENGES FOR TEACHERS & TEACHER EDUCATORS*. (details)
- Vanderlinde, R., & van Braak, J. (2010). The e-capacity of primary schools: Development of a conceptual model and scale construction from a school improvement perspective. *Computers & Education*, 55(2), 541–553. <https://doi.org/10.1016/j.compedu.2010.02.016>. (details)
- Vegas, E., Ziegler, L., & Zerbino, N. (2019, November 20). How ed-tech can help leapfrog progress in education. *Brookings*. <https://www.brookings.edu/research/how-ed-tech-can-help-leapfrog-progress-in-education/>. (details)
- Walker, H., Tegha, G., Lurvink, A.-F., McBurnie, C., & McBurnie, H. W., Ghislaine Tegha, Anne-Fleur Lurvink and Chris. (2022, March 31). *Tich Mi Ar Tich Dem - Open Development & Education*.

<https://opendeved.net/2022/03/31/tich-mi-ar-tich-dem/>,
<https://opendeved.net/2022/03/31/tich-mi-ar-tich-dem/>. (details)

Weatherby, K., Clark-Wilson, A., Cukurova, M., & Luckin, R. (2022). The Importance of Boundary Objects in Industry-Academia Collaborations to Support Evidencing the Efficacy of Educational Technology. *TechTrends*, 66(5), 784–797. <https://doi.org/10.1007/s11528-022-00705-0>. Available from <https://link.springer.com/10.1007/s11528-022-00705-0>. (details)

Wilson-Grau, R. (2018). *Outcome Harvesting: Principles, Steps, and Evaluation Applications*. Information Age Publishing. (details)

World Bank Group. (2021). *World Bank EdTech Readiness Index (ETRI)*. <http://documents.worldbank.org/curated/en/811011628250703800/World-Bank-EdTech-Readiness-Index-ETRI>. (details)

Appendix A. Our research approach

We used two main strategies to get insights into current EdTech Testbed initiatives: (1) we conducted a literature search, (2) an online survey was sent to stakeholders in the field. In the next section, both strategies are explained in more detail.

10.1. Literature search

A literature search was conducted in several scientific databases (Web of Science, Google Scholar). Relevant keywords for this search were: *“EdTech”*, *“testbed”*, *“sandbox”*, *“educational technology”*, *“toolkit”*, *“evidence”* etc. We also contacted several key persons in the field via email to ask for relevant paper contributions. If in scope, an online conversation was held to discuss the paper. Specifically, we were interested in papers that described and/or evaluated EdTech Testbed methodologies. Furthermore, we found relevant articles by searching websites of existing EdTech Testbeds or sandboxes. Finally, based on key papers in the field, we searched the literature backwards and forward. Inclusion criteria applied were: (1) *“journal papers”*, organisation reports, blog posts, book chapters, website pages; (2) relates to the process of developing a Testbed AND/OR to an implementation of a Testbed AND/OR to recommendations or learnings about the Testbed process; (3) published in the last 15 years; (4) focused on K12 education, and (5) included a description of the *“test”* and/or description of *“bed”*.

10.2. Stakeholder contributions to the survey

An online survey was sent to all stakeholders in the field. The topic of the survey was about educational Testbeds that involve technology. In total, 20 responses were collected from members/leads of EdTech Testbeds (6), EdTech incubator or

accelerators (5), programme implementer (4), investors (3), researchers (1) and policy makers (1). Participants were asked to describe how they were involved in an EdTech Testbed and whether they could share reports of methods, processes or models they used in their EdTech Testbed. Furthermore, we asked them what they thought were the main challenges for the design, implementation and evaluation of EdTech Testbeds. The contributions of the stakeholders were used as input for this report.

Appendix B. Motivating factors for different stakeholders' involvements in EdTech Testbeds

Stakeholder	Motivating factors for their involvement
Teachers and schools	<ul style="list-style-type: none"> ■ To influence the education innovation agenda and associated policy development ■ To participate in product design (having access to (test) EdTech) ■ To learn through EdTech co-design and collaboration opportunities within professional communities ■ To improve their students' (technology-enhanced) learning experience
Researchers	<ul style="list-style-type: none"> ■ To gain recognition for their work through better dissemination of existing EdTech research evidence ■ To influence the education innovation agenda and associated policy development ■ To ensure a high quality evidence base for available EdTech ■ To lead on innovations for collaborative, participatory and co-design research methods.

Stakeholder	Motivating factors for their involvement
EdTech companies	<ul style="list-style-type: none"> ■ To generate product credibility (able to test products in real academic environments, prove value of solution etc.) ■ To influence the education innovation agenda and associated policy development ■ To collect feedback about product, incorporate feedback in product design, realise proof of concepts ■ To scale their product within and beyond existing markets
National and Regional Governments	<ul style="list-style-type: none"> ■ To grow the size of the EdTech sector to generate national income, job creation and widespread educational impacts. ■ To win elections!

Appendix C. EdTech Testbed Methodologies and Methods

Here, we give a deeper dive into the different methodologies and methods that are used to support the evaluation of EdTech implementations (or tests).

Methodologies refer to

“the strategy, plan of action, process or design lying behind the choice and use of particular methods and linking the choice and use of methods to the desired outcomes” (↑Crotty, 2020).

Methods refer to

“the techniques or procedures used to gather and analyse data related to some research question or hypothesis” (↑Crotty, 2020).

12.1. Methodology

Methodologies used in EdTech Testbeds tend to have two different roots, and in some cases have beautifully converged. The first root is the field of education research, and shares similarities with research methodologies from the social sciences. These methodologies may also be found in anthropology, sociology, or psychology. The primary focus of these approaches is to build evidence on effectiveness, while contributing to the future development of the EdTech intervention is secondary.

The second root is the tech innovation sector, where methodologies used share similarities to entrepreneurial and innovation language. These methodologies may overlap to ones used to develop mHealth, AgriTech or FinTech, and are more often presented as toolkits, principles or processes rather than methodologies. The primary focus on these approaches is to improve the design and implementation, where contributing to global evidence is secondary.

12.1.1. Methodologies from education research

This section details several methodologies used in education research.

Action research

Action Research is a process of systematic inquiry that is led hands-on by practitioners themselves, rather than researchers (↑McNiff, 2017). Practitioners could be teachers, administrators or principals. Consequently, the research questions posed are insider rather than outsider questions, e.g., *“How do I improve my teaching using digital tools in the classroom?”* The process used is generally: observe, reflect, act, evaluate, modify and move into new directions (ibid.).

Example: With technology changing faster than education research can keep up with, ↑Atherton (2019) report on an action research study that investigates the realities of EdTech use among trainee teachers. The study used mixed methods to triangulate between qualitative and quantitative data. In the midst of the big data buzz, the research recommends that more EdTech research needs to focus on teachers sharing ideas, for example, through autoethnographies.

Rapid Evaluation Cycles

Rapid Evaluation Cycles, also known as Rapid Cycle Evaluations (or *“Test and Learn”* in ↑Batty and colleagues, 2019), are rapid evaluations done to generate quick, actionable evidence to make evidence-based decisions for developing and optimising programs, products or strategies as well as to improving their implementation (↑American Institute for Research, 2018). A key feature of REC is the iterative nature, where iterations are likely to happen within weeks or months (rather than years; ibid.). The questions that might be asked relate to *“what works”, “for whom”* and *“under what conditions”* (ibid.).

Example: The ↑USA Department of Education (2022) created “Educational Technology Rapid Cycle Evaluation’s to determine how effectively technology is being used in classrooms.²⁶ A *“Coach”* tool was developed to support district officers, school leaders and educators to conduct rapid cycle evaluation. The aim was to gather evidence to make informed decisions about EdTech acquisitions.

²⁶ ↑Educational Technology Rapid Cycle Evaluations: <https://tech.ed.gov/rce/>

Design-based research (DBR)

Originating in learning science, DBR aims to address the challenges of design innovation in complex environments, while bridging the gap between research and practice (↑Crowley, 2016). The methodological approach involves researchers investigating solutions to education problems through designing and iteratively testing them in real-world settings (predominantly classroom settings; ↑Fishman and colleagues, 2019). Contrasting traditional education research that tends to be disjoint from the real world, DBR forefront the role of context and researchers are agents rather than observers (↑Armstrong and colleagues, 2020). The outcomes of DBR are to develop designs, practices, or processes that work effectively to achieve their goals and to contribute to literature on learning.

Example: ↑Siko & Barbour (2016) (as cited in ↑Armstrong and colleagues, 2020) use a three-cycle DBR process to build on initial literature that indicated that homemade PowerPoint games showed no improvement to test scores, despite it being grounded in proven pedagogical strategies. In each cycle, they adapted based on recommendations. At the end of the study, students who engaged in the creation of PowerPoint games performed statistically higher than those that did not.

Design-based Implementation Research

DBIR builds upon DBR, bringing more complexity through organisations and systems. It works similarly to the iterative cycles of DBR, but moves beyond the classroom setting to look at education reforms across contexts (↑Ubaidullah, 2016). A key feature of DBIR is researchers, implementers (e.g., government) and practitioners (e.g., teachers) working together in equal status to address the challenges of effective implementation (↑Crowley, 2016). Additionally, DBIR goes further than understanding and improving an intervention in particular, it abstracts this to theorise how to better design an implementation process that considers the complexity of the content (↑Fishman and colleagues, 2019).

Example: DBIR was used in the “*Tich me ar Tich dem*” (Krio for “*teach me to teach them*”) research in Sierra Leone. Working with the Ministry of Basic and Senior Secondary Education (MBSSE) and the Teaching Service Commission (TSC) the project team sought to identify the most effective way to deliver low-cost, scalable in-service TPD (↑Walker and colleagues, 2022). The research was done

collaboratively with school-based facilitators (peer-facilitators), school leaders and teachers to identify needs, barriers, and opportunities through a process of iterative, grounded inquiry ([↑Lurvink and colleagues, 2022](#)). The study had 3 iterative cycles (starting with two schools, then four, and finally 10 schools) where data was collected, analysed and used to optimise both the programme implementation and the research instruments ([↑Lurvink and colleagues, 2022](#)).

Outcome harvesting

Outcome harvesting is a highly participatory evaluation process that is used to evaluate identify, formulate, verify, analyse and interpret outcomes of an intervention ([↑Wilson-Grau, 2018](#)). It is particularly useful where cause and effects are not easily measurable or weren't predetermined at the start of the intervention. Thus, it uses a backwards approach through collecting evidence on changes that have taken place and how the intervention contributed to these changes. Outcome harvesting has 6 loose steps or guiding principles: 1) design the harvest, 2) review documentation and draft outcomes, 3) engage with informants, 4) substantiate, 5) analyse and interpret, and 6) support use of findings.

Example: Outcome Harvesting was used in an initial self-assessment of the impact that EdTech Hub's technical assistance had on the design of the One Tablet Per School programme (Sierra Leone's Ministry of Basic and Senior Secondary Education; [↑McBurnie, 2021](#)). The assessment took two sequential steps (document review and stakeholder analysis) to assess if there were changes to approach and whether these changes contributed to the larger goal of expanding data use.

Sequential Multiple Assignment Randomised Trials (SMARTs)

SMART are used adaptive interventions that want to compare the relative effectiveness of different intervention options. Adaptive interventions in this model refers to an intervention that uses sequenced decision-making rules to provide tailored changes to the intervention based on participant responses ([↑American Institute for Research, 2018](#)). An example of this is personalised learning software. The process involves randomly assigning participants to different treatment groups, and based on responses, further allocating them to modified treatments.

The SMART design allows for questions around sequencing and dosage of activities in real-world contexts to be explored, based on different participants needs ([↑Kim and colleagues, 2019](#)).

Example: SMART was used in a study to develop an adaptive literacy intervention composing of personalised print texts and digital activities on a mobile application ([↑Kim and colleagues, 2019](#)). The study consisted of two stages. In the first stage, students were randomly assigned to two groups with different text and app activities. Based on responses, the second stage augmented and intensified the intervention in different ways. The study concluded that adaptive findings needed would not have been possible in a standard RCT with fixed treatment, however RCTs would be more useful down the line to evaluate intervention effectiveness.

Systemic Mixed-Methods Research

Systemic mixed-methods research combines the need for formative, fast feedback loops and larger-scale, longer-term rigour. It combines

“quantitative and qualitative approaches to appreciate both the scale of effects and understand the reasons behind observed phenomena”

([↑Haßler and colleagues, 2021](#)).

This approach was designed specifically with EdTech in mind, as technology is fast changing. The sequential approach begins with an evidence review, followed by more qualitative, agile methods in the first phase (typically, DBIR); this is followed with a traditional mixed methods approach, including quantitative (quasi-)experimental design (RCT / QED) to assess outcomes (e.g., learning outcomes for education system effectiveness). Qualitative aspects of the second phase incorporates further qualitative research (interviews, focus groups, observation, artefact analysis, etc) to ensure the process details are also captured to explain the “*how*” and “*why*” of outcomes.

Example: The EdTech Hub, Aga Khan University and Tanzania institute of Education are using systemic mixed-methods research to research effective and cost-effective sustainable, technology-supported, decentralised, and school-based Teacher Continuous Professional Development (TCPD) models to improve learning outcomes in rural primary schools in Tanzania ([↑EdTech Hub, 2022](#)). The study employs DBIR in the first phase, with two iterative cycles (4

schools and then 8 schools; [↑Koomar and colleagues, 2022](#)). The second phase aims to assess what value technology adds through comparing the same school-based TCPD model with and without mediation by tech provided at school level and against business-as-usual, including measuring cost-effectiveness.

Implementation Science

Recently, the notion of Implementation Science in education has gained ground. Originating in the field of health ([↑Albers and colleagues, 2020](#)), Implementation Science

“... has been defined as the scientific study of methods to promote the systemic uptake of research findings and evidence-based practices into professional practice and public policy” ([↑Eccles and colleagues, 2009](#)).

The notion of Implementation Science is rooted in the experience that research findings do not automatically turn into practice, but also by the experience that certain established research approaches can miss implementation details (see, for example, [↑Kerwin & Thornton, 2020](#)). While recognition of use of Implementation Science in education go back at least a decade (see, for example, [↑Forman and colleagues, 2013](#)), there has been renewed interest, particularly for low- and middle-income countries ([↑Building Evidence in Education, 2023](#)).

12.1.2. Methodologies from tech innovation

This section reviews processes, standards, frameworks, toolkits, and principles that have been developed in entrepreneurial and innovation spaces. Since such processes tend to be used for internal EdTech design and evaluation, there is less emphasis placed on publishing or making reports publicly available. This section is thus lighter on well-documented case studies that explicitly state an approach used and identifies a gap in sharing evidence generated from innovation spaces.

EdTech Tulna

Designed for the Indian context but generalisable far beyond, EdTech Tulna sets out research-based standards to develop and benchmark the design of EdTech products. The aim of the standards are to enable high-quality EdTech design that

meets demand, given the countless number of EdTechs on the market (↑[EdTech Tulna, 2021](#)). The main beneficiaries of the Tulna standards are EdTech product developers, decision makers (e.g., a government official or institution), local community users (e.g., an NGO or group of teachers) and private end users (e.g., the student). The standards fall into three broad categories: Content quality, pedagogical alignment, and technology and design. EdTech Tulna also has an evaluation centre where product evaluations can be viewed and compared.

Website: <https://edtechtulna.org/>

Design thinking for Educators Toolkit

The [Design Thinking for Educators Toolkit](#) is a tool for educators to work together to create equitable change in schools (↑[IDEO, 2013](#)).²⁷ The toolkit supports a community-led, equity-centred, and design-driven process. Design thinking is defined as a mindset that is human-centred, collaborative, optimistic and experimental and can be used to address challenges that

- centre around the design and development of learning experiences (curriculum),
- learning environments (spaces),
- school programs and experiences (processes and tools), and
- system strategies, goals, and policies (systems).

A follow-up [Co-Designing Schools ToolKit](#) is also available to build the capacity of school communities.

Website: <https://www.ideo.com/post/design-thinking-for-educators> and <https://www.codesigningschools.com/>

Human-centred Design Toolkit

HCD is a creative, practical and repeatable approach to problem-solving and arriving at innovative solutions (↑[IDEO, 2015](#)). Empathy is at the core of the design. The guide provides a step-by-step approach to solving complex problems in

²⁷ An Arabic/English version of the Design Thinking for Educators Toolkit is available (↑[Schurr, 2020a](#); ↑[Schurr, 2020b](#); blogpost: ↑[Khalayleh, 2020](#)).

the social sector. It has three phases: Inspiration, Ideation and Implementation.

Website: <https://www.designkit.org/methods>

Problem Driven Iterative Adaptation (PDIA)

Developed in the context of building state capacity, PDIA is a Do-It-Yourself step-by-step approach which helps you break down your problems into its root causes, identify entry points, search for possible solutions, take action, reflect upon what you have learned, adapt and then act again.” (↑[Samji and colleagues, 2018](#)). The approach is based on four principles: local solution for local problems, pushing problem-driven positive deviance (i.e., encouraging experimentation), promoting active experiential and experimental learning (try, learn, iterate, adapt), and scale through diffusion. The PDIA approach follows 6 steps: 1) Initial problem analysis, 2) Identify action steps, 3) take action, 4) check-in, 5) sustain authority and legitimacy, and 6) adapt and iterate.

Website:

<https://buildingstatecapability.com/2018/10/25/introducing-the-pdia-toolkit/>

National Digital Education Architecture

Developed by the Ministry of Education in India, NDEAR is a holistic digital infrastructure for the education ecosystem. The technological framework outlines 36 interoperable building blocks across 12 categories:

1. open Standards & NDEAR Portal,
2. federated identities (i.e., students, teachers, and education institutions),
3. reference data,
4. infrastructure,
5. technology (i.e., software services),
6. governance,
7. administration,
8. content,

9. learning,
10. reference applications/ solutions,
11. open data and analytics, and
12. ecosystem sandbox.

The aim is for ecosystem players from both the public and private sectors to plug-in, use, or build on NDEAR to develop relevant platforms, solutions, and programmes for students, teachers, parents, administrators, and the community. The standards assist in evaluating learning environments. Such an infrastructure addresses many of the EdTech challenges faced by other countries in terms of scaling and aligning to the national education system. An example of a platform built on NDEAR is Diksha, the national infrastructure platform for school education in India, used widely across the country in more than 35 languages.

Website: <https://www.ndear.gov.in/index.html>

Principles for Digital Development

The Principles for Digital Development area set of nine living guidelines targeted at integrating best practices into technology-enabled programs. They provide guidance at different product life cycle phases, as well as resources and online courses. The nine principles ([↑Principles for Digital Development, 2022](#)) are

1. design with the user,
2. understand the existing ecosystem,
3. design for scale,
4. build for sustainability,
5. be data-driven,
6. use open standards, open data, open source, and open innovation,
7. reuse and improve,
8. address privacy and security,
9. be collaborative.

Website: <https://digitalprinciples.org/principles/>

The Sandbox Model

EdTech Hub's sandbox model creates a Testbed to test and iterate interventions in uncertain conditions. It combines research and design methods to support implementers to evaluate their offerings quickly and iteratively. It builds on previous models such as backcasting (↑Robinson, 1990), lean start-up (↑Chang, 2018), user-centred design, agile methodology (↑Kaiser, no date), and behavioural innovation (↑Simpson, 2019; as cited in ↑Boujikian and colleagues, 2022). While flexible, the guiding steps of the sandbox are

1. articulate the desired impact of the intervention and surface critical assumptions,
2. design lean experiments to test the assumptions,
3. reflect, learn, and iterate at regular intervals.

Website: <https://edtechhub.org/sandboxes/>

Theories of Change and Logic Models.

They help generate a simple “*theory*” for how the EdTech is going to achieve its desired impacts, and help to define and make visible the metrics on which any research evaluation would be based.

12.2. Methods

12.2.1. Methods for brainstorming and information gathering

A small selection of methods for brainstorming is selected from design thinking IDEO's toolkits.

- Mindset mapping²⁸ – A process to surface the diverse mindsets, perspectives, and approaches.

²⁸ <https://www.codesigningschools.com/toolkit-phase-one>

- Photo journaling²⁹ – A Photo journal is a straightforward and very visual way to get a glimpse into how a person lives his or her life and also helps empower individuals to tell their own stories.
- The Five Why's approach³⁰ – The Five Whys is a fantastic method to use to get to the core of a person's beliefs and motivations.
- Draw it³¹ – Drawing can also help the person you're designing for organising her thoughts visually and generally spur ideas and conversation in a different way than talking.
- Resource flow³² – By organising and visualising how a person or family spends money, you'll see how it comes in, goes out, and opportunities for more efficiency in the system.
- Journey mapping³³ – A Journey Map allows you to identify and strategize for key moments in the product, experience, or service you're designing.
- Top 5³⁴ – This exercise gives you a break from the deep thinking and simply asks, what are the top five ideas or themes sticking out to you right now. Not only can answering this question as a team help you strategize, but it can also help uncover themes, isolate key ideas, and reveal opportunities for design.
- Storyboarding³⁵ – A quick, low-resolution prototype to help you visualise your concept from start to finish.
- Role play³⁶ – A Role Play is a type of prototype that is not only pretty easy to build, but can also help you get an idea, experience, or product in front of the people you're designing for quickly;

²⁹ <https://www.designkit.org/methods/photojournal>

³⁰ <https://www.designkit.org/methods/the-five-whys>

³¹ <https://www.designkit.org/methods/draw-it>

³² <https://www.designkit.org/methods/resource-flow>

³³ <https://www.designkit.org/methods/journey-map>

³⁴ <https://www.designkit.org/methods/top-five>

³⁵ <https://www.designkit.org/methods/storyboard>

³⁶ <https://www.designkit.org/methods/role-play>

- Theory of change/Logic Modelling³⁷ – Articulate and interrogate your assumptions about how your solution will create positive change. Logic models have been used as tools to support EdTech developers to make their design decisions more explicit to inform future EdTech Testbed research methods (↑[Weatherby and colleagues, 2022](#)).

12.2.2. Methods for design, monitoring, and evaluation

↑[Batty and colleagues, 2019](#) put forward the following methods for co-design and participation: participatory design; design ethnography; joint requirements gathering; storyboarding, iterative prototyping; and contextual design.

Other methods from qualitative feedback, monitoring, and evaluation include focus group discussions, unstructured, semi-structured Interviews, open-end questionnaires, field notes, observations, and document analysis (↑[Cohen and colleagues, 2017](#), ↑[Crotty, 2020](#)).

Some quantitative evaluation methods include A-B testing, pre-post evaluations, end of course surveys, short-duration experiments, factorial designs, rubrics, rating scales, checklists, and surveys (↑[Cohen and colleagues, 2017](#)).

Lastly, methods involving large education data sets and learning analytics include usage stats (e.g. data on numbers of participants, geographical location) (↑[Eriksson and colleagues, 2018](#)), emotional metrics (↑[Rodriguez-Segura, 2020](#)) and progress and achievement metrics (e.g., student learning outcomes, course completion, course fail, course grade, persistence, and retention) (↑[Hollands & Escueta, 2017](#)).

12.3. Learning more about generating evidence

Open Development & Education curates publications in our evidence library, available at <https://docs.opendeved.net>. The library also contains numerous publications and reports regarding research methods, innovation methods, and evaluation.

³⁷ <https://www.designkit.org/methods/explore-your-theory-of-change>

Figure 12.1. Screenshot from <https://docs.opendeved.net> showing resources for implementation science.

THEME

- Implementation science
 - Design-thinking (6)
 - Education (28)
 - General (36)
 - Health (13)
 - Improvement science (2)
 - Overview (2)
- Programme implementation (6)
- Research and evaluation (5)

PUBLICATION YEAR

- Between 2000 and 2023 (52)
- Unknown (6)

RESULTS 58 resources

Sort by: Author A-Z ▾

1 2 3 >

Abstracts

Quality Improvement Approaches: Implementation Science
2017, March 15 - Carnegie Foundation for the Advancement of Teaching
The implementation science approach to improvement in education centers on how to accommodate local school variables and other contextual factors that can impede successful implementation
[View on www.carnegiefoundation.org](#)

Using implementation science to support the implementation of interventions in real world contexts | Self-evaluation | National Improvement Hub
n.d.
[View on education.gov.scot](#)

Design Research in Education: A Practical Guide for Early Career Researchers
Bakker, A. - 2018 - Routledge
Design Research in Education is a practical guide containing all the information required to begin a design research project. Providing an accessible background to the methodological approaches used in design research as well as addressing all the potential issues that early career researchers will encounter, the book uniquely helps the early career researcher to gain a full overview of design research and the practical skills needed to get their project off the ground. Based on extensive...

An introduction to implementation science for the non-specialist
Bauer, M. S., Damschroder, L., Hagedorn, H., ... Kilbourne, A. M. - 2015 - *BMC Psychology*, 3(1), 1–12

Implementation science: What is it and why should I care?
Bauer, M. S., & Kirchner, J. - 2020 - *Psychiatry Research*, 283, 112376

Adaptation in dissemination and implementation science
Baumann, A. A., Cabassa, L. J., & Stirman, S. W. - 2017 - *Dissemination and Implementation Research in Health: Translating Science to Practice*, 2, 286–300

Appendix D. Describing EdTech Testbeds – Some framing questions

This appendix provides framing questions to describe EdTech Testbeds.

13.1. What preconditions do the EdTech Testbed satisfy?

This section provides key questions regarding preconditions.

- Is a sustainable funding model in place? If so, what is the funding model?
- Do the local, regional or national education & economic policies support the EdTech Testbed to thrive? If so, which policies, and in what ways?
- Are the local preschool and school environments supportive to the work of the EdTech Testbed? If so, in what ways?
- Are key stakeholders involved and motivated within the EdTech Testbed? If so, which stakeholders and in what ways?
- Does the EdTech Testbed have access to relevant educational and technological expertise to support its work? If so, what expertise is accessed and in what ways?
- Does the Edtech Testbed access existing evidence to support its work? If so, what sources are accessed, and how are they used?

13.2. Which key components describe the work of the EdTech Testbed?

This section describes the key components that describe the work on 'EdTech Testbed', organised by "*EdTech*", "*test*", and "*bed*".

13.2.1. EdTech

The following key components and key questions relate to the term “*EdTech*”.

- How is the intended educational purpose for the EdTech to be tested made clear to all?
- How is the type of technology described in ways that can be understood by all?
- How is the intended target group for the EdTech (the beneficiaries) made clear to all?
- How is the intended educational content of the EdTech made clear to all?
- How is the intended pedagogy for the EdTech made clear to all?
- How is the stage of innovation of the EdTech established, and how is this communicated?
- What are the selection criteria for the EdTech to be tested, and how are they communicated?

13.2.2. Test

The following key components and key questions relate to the term “*test*”.

- What is the purpose of the Test? (i.e. What is being tested, and why), and how is this purpose communicated to all (or not)?
- What methodology and methods are used, and how are these communicated?
- How are the timeline and frequency for the testing made clear to all?
- What procedures are in place, and how are these communicated?
- Who provides the leadership for the Test, and what does this role involve?
- Is there a clear plan for the reporting of the Test findings to key stakeholder audiences using appropriate media?

13.2.3. Bed

The following key components and key questions relate to the term “*bed*”.

- Are details about the context for the test captured in ways that are both ethical and sensitive to the setting?
- Are details about the Testers captured in ways that enable relevant educational, socio-economic, physical and cultural differences to be considered, and comply with ethical procedures?
- Are details about the learning spaces captured?
- Is the scale of the test being captured, to include the numbers of schools, teachers, and learners involved?