REPORT

EDTECH IMPACT EVALUATION FRAMEWORKS: SUMMARY 2023

2023

Natalia I. Kucirkova
Janine Campbell
Anna Lindroos Cermakova

WiKIT

CONTENTS

CURRENT EDTECH CONTEXT

The natural experiment and acceleration of EdTech's global implementation during COVID-19, has tested the underlying ideals of the field. National governments, schools, and families have witnessed first-hand that the idea of **technology serving as an equity catalyst**, and enabling **scalable personalised education** for all, may be a false hope, inadvertently excluding those who require this support the most.

The period of post-pandemic reflection has resulted in a series of reports that delve into some adverse effects of inappropriate data use on marginalised communities, offering recommendations for potential strategies to address these issues. A common theme echoed throughout these reports is the paramount importance of **placing technology usage in the hands of learners and teachers**, rather than letting commercial entities dictate the terms. Decisions regarding when and by whom technology is utilised should, therefore, be grounded in **robust and rigorous evidence of its impact, backed by thorough research [1].**

Issues around data safety, interoperability of technologies [see 2] and recently, the use of generative AI, have garnered considerable national funding and dedicated international attention. This has not been the case for 'EdTech Evidence' — the research and scientific proof behind an EdTech solution. The landscape is continuously shifting, driven by **tightened school budgets, increasingly discerning teachers and families, and an oversaturated EdTech market [3].** The most effective, cost-efficient, and impactful tools need to be selected for use, and these choices need to be guided by robust evidence of their effectiveness.

Countries vary in their proactivity in this regard. While some are still debating the question of "whether" to use technology in schools [4], others have taken a more proactive approach by establishing criteria and systems to **systematically evaluate what works best** within their educational systems. In the USA, for example, there has been a notable upswing in interest surrounding EdTech certifications and the 'Evidence movement', with the pivotal year of 2023 witnessing the introduction of **new certifications** to the market, such as those offered by Digital Promise. This development signifies that these organisations can serve as models for desired product design, granting their seals of approval to products that meet their criteria. Additionally, the adoption of the ESSA Standards of Evidence has become more widespread, thanks in part to the US Department's Office of Educational Technology releasing guides and supporting documents for both EdTech providers and procurement teams. These standards have become intricately linked to district funding, prompting several EdTech providers to reevaluate their marketing and scaling strategies. They are now **shifting from an emphasis on volume and rapid scaling, to prioritising rigorous research and external verification** to substantiate their claims.

In every country, the significance of evidence demonstrating a positive impact on learning is complemented by the national and international mandates for secure data utilisation and interoperability of individual solutions deployed in classrooms. For example, in the UK, EdTech's compliance with the Children's Code is essential before evidence can be assessed [5]. However, there is no hierarchical order in how distinct quality requirements for EdTech, pertaining to both data integrity and evidence of effectiveness, should be approached. They are equally essential.

> It would be counterproductive to utilise an EdTech solution that does not contribute to learning, or one that, despite its effectiveness, improperly exploits children's data for commercial purposes.

While the responsibilities for data safety and privacy have predominantly been assigned to legal experts, data researchers, and policymakers, the primary stakeholders in substantiating the effectiveness of technology are researchers, scientists, teachers, and educators, who **assess these solutions in real classroom settings [6].** They provide firsthand accounts and scientific evidence of what works effectively in the classroom. It is perhaps the **complexity of the academia-practice-industry collaboration** that explains why EdTech Evidence has been neglected for so long [7]. Rather than being prioritised, it has been delegated to individual EdTech-researcher partnerships, or individual teachers testing solutions in their classrooms, with no systematic oversight.

THE OPPORTUNITY AND CHALLENGE OF EVIDENCE-BASED EDTECH

We see the drive for EdTech evidence as **beneficial for all stakeholders**: educators, researchers, and EdTech founders.

Educators can foster product improvement and enhance student learning outcomes, and therefore have a unique opportunity **to shape the future of EdTech.**

Researchers can harness their analytical expertise, generate real-world applications that amplify the impact of their work within specific subject domains, and apply their research and validation skills to practical use, **reaching millions of learners**.

EdTech founders can develop routines for systematically embedding research into their design and scaling, with **verified impact metrics informing their business** strategies.

Despite these benefits, the **EdTech evidence movement** is challenged by ill-informed, broad debates that lump all EdTech solutions together, failing to explore their contextual effectiveness [4]. Instead of nuanced discussions about **what works for specific learners and situations**, these debates tend to oversimplify technology as inherently good or bad. This inclination is notably evident in universal technology bans, where students lose access to particular apps or platforms once their smartphones or Chromebooks are removed from schools. Unfortunately, this reverts the focus from the **three digital divides that EdTech addresses**—access, use, and content/design of learning [8]—back to the issue of access for those with and without technology. This situation raises **equity concerns,** as access to technology serves not only important educational objectives but also safety and well-being goals [9].

Several internationally recognized groups (such as UNESCO, UNICEF, and the World Bank) have begun to take a stand on **how evidence in EdTech should be coordinated on a national level, incentivised, and/or validated.** Nevertheless, effective adoption of these recommendations is hindered by three factors:

01 There are no statutory national policies or international legal guidelines on EdTech Evidence.

02 There is limited professional training in skills for EdTech evaluations, or rigorous EdTech research and development for leveraging such training.

03 There is no easily comprehensible overview of various EdTech Evidence evaluation tools, frameworks, and recommendations.

SUMMARY

The current context offers a notable opportunity for the EdTech sector to take a leadership role in education, and to **establish robust evidence-based practices and effective collaboration between industry, academia, and educators.** There are also some challenges inherent in this endeavour [10].

Our report seeks to initiate the first steps required to compile a **comprehensive overview of global EdTech evidence standards**. With EdTech founders as our target audience, our goal is to create a readily accessible resource that consolidates various evidence frameworks and types of evaluations commonly applied to EdTech. We aim to concisely and coherently present the shared characteristics of these diverse frameworks and guide EdTech providers in **strategically engaging with those that are most relevant** for their solutions.

THIS REPORT

From the perspective of an EdTech company, provider, or designer (EdTech founder hereafter), the **EdTech Evidence landscape can appear overwhelming**. There is a marketplace of competing claims and approaches for measuring and demonstrating the impact of solutions. In this report, we will present a short description of the various certifications and standards on the market, and provide an **overview of the options available.** We position this overview primarily for companies operating in the English-speaking market.

Our primary objective is to support EdTech founders to be more evidence-based and, as such, we hope that this report serves as a resource for **understanding the evidence criteria** and for preparing accordingly to meet these standards.

While evidence that a solution works as intended is relevant for all types of EdTech, it is mostly those that target **teaching and learning** (rather than education management or employability) that require documented and rigorous evidence of positive impact on learners. It is this type of EdTech that we focus on in this report.

EdTech vary in their goals and mission, but they all try to improve learning, either through **community building in classrooms, increased learners' engagement, or improved teacher efficiency and effectiveness for impacting student learning**. In this report, we list frameworks that are used on a national and international basis to address these **four aspects of learning**.

We discuss three broad framework categories — **pedagogical certifications, national frameworks, and international frameworks and standards** — and we provide several examples for each of these categories. We recognise that the resulting list is not exhaustive. When choosing examples for this report, we have selected well-known and established examples within the field. We conducted a systematic search using conventional literature research methods. In addition, we relied on our own knowledge of the industry and reached out to scholars within the EdTech field for recommendations.

The academic search led to several academic frameworks developed by researchers to evaluate EdTech in specific subject areas (e.g. literacy apps) or EdTech for specific age groups (e.g. use of touchscreens for two-year-olds). These frameworks typically contain detailed criteria that researchers use when evaluating existing resources in a specific domain, but not as official or national criteria for evaluating the quality of EdTech. We therefore excluded them from this report, but **recommend that EdTech founders familiarise themselves with such frameworks, especially if targeting specific pedagogical areas and aspiring to align their design with latest scientific evidence.**

We also excluded frameworks that emphasise **data safety and interoperability**, even though we acknowledge their significance in the comprehensive evaluation of EdTech quality. Our specific focus is on certifications related to **educational impact**, where both pedagogical and research-based certifications take precedence.

Recognizing the dynamic nature of the EdTech landscape, we acknowledge that our list of frameworks cannot be all-encompassing and serves as a snapshot of the current landscape, subject to periodic updates. **We eagerly encourage contributions from the EdTech community** to enrich and broaden our list over time.

The frameworks we discuss in this report broadly map on three business models followed by EdTech:

- Business-to-Consumer **pedagogical certifications** relevant for teachers,
- Business-to-Business **national frameworks** for schools,
- Business-to-Government **international standards of evidence.**

PEDAGOGICAL CERTIFICATIONS

A **pedagogical certification** is a **systematic framework** of evaluations by practising teachers and/or experts, with scoring usually aligned with established learning science theories.

The integration of EdTech in classrooms is most often mediated by teachers. In some countries, teachers serve as gatekeepers and have the authority to make direct decisions regarding which resources are used in the classroom and how they are employed in the teaching process. Although the degree of public trust in the teaching profession may differ from one country to another, teacher influence holds global significance. The recent rise of teacher influencers and online marketplaces, where educators share or monetize their educational resources, highlights the necessity for EdTech markets to closely monitor trends within teacher networks to maintain their relevance.

It is also well-known that teachers trust other teachers and want products that are **based on solid theories and practice, ideally with a package of recommended activities**. Studies also show that teachers make discerned choices when they download apps from the App store and **these choices are aligned with the learning sciences**. As such, teachers' voices in choosing and evaluating an EdTech product, should not be perceived as replacing a pedagogical certification but rather as an additional insight into what they deem most appropriate for their individual contexts. Many providers of pedagogical certifications rely on teachers to score and evaluate the individual products. Pedagogical certifications can be thus likened to an evaluation by selected teachers; however, they follow a systematic framework and align their scoring with established learning science theories.

Not all pedagogical certifications deliver both evaluation and recommendation for use, but most provide information that facilitates teachers' decision-making around the trustworthiness and quality of a given EdTech resource. To obtain a pedagogical certification, EdTech founders need to apply to a specific provider, who typically conducts a product review and evaluates the product's features according to a predefined rubric. Some providers award badges and certifications, while others provide the founders with internal reports. Some providers share the evaluation process through open-access explanations of each criterion, while others use a proprietary framework with criteria known only to the provider. Some providers actively recruit practising and local teachers in completing each evaluation, while others complete the evaluation with help of their own experts trained in the specific framework.

Generally speaking, engaging teachers in the evaluation of EdTech tools is an effective way of **integrating EdTech pedagogy into teachers' professional development**. The evaluation process should not be a hasty checkbox activity; instead, it should involve thoughtful discussions among colleagues about their individual interpretations of quality and the underlying reasons for their perspectives. This is the reason why, in some countries, notably those with greater teacher autonomy, pedagogical certifications are less well-received, and EdTech providers are better to solicit teachers' independent reviews.

The EdTech Impact platform aggregates **independent teachers' reviews** on various educational technologies from across the world with all ratings shared in a transparent manner.

Counter-arguments to pedagogical certifications include that not all teachers are technologically savvy and most teachers are too busy to evaluate each app or platform available for public download.

With **50,000 apps** marked as "educational" and an average US district using up to 1,500 different EdTech applications, a professionally conducted evaluation can ease teachers' burden.

EdTech providers can also opt for a British official mark of quality and reliability ('Kitemark'), or a seal of approval, distinguishing their solution from others according to objective criteria. In a crowded EdTech market brimming with numerous alternatives, this is an **opportunity that many EdTech want to invest in**. Given that pedagogical evaluations focus on existing design according to a set of theory and practice criteria, they are **easier to obtain than badges of evidence.** As such, they represent a "lower-hanging fruit" opportunity to leverage an EdTech's evidence journey.

When selecting the **key examples of pedagogical evaluations** for this report, we chose those that provide a clear answer to t**he vital pedagogical question** — how does the solution help learners to learn? Pedagogy is thereby understood as **both the theory and practice** of how to teach best. It follows that the pedagogical evaluations need to focus on the **EdTech's content and design**, and the design's alignment with principles of effective pedagogy and learning sciences.

We will present three example of pedagogical certifications:
Education Alliance Finland, Digtial Promise, and **ISTE Standards.**

EDUCATION ALLIANCE FINLAND

Education Alliance Finland (EAF) certifies the pedagogical quality of an EdTech with **EAF's Quality Certificate for Learning Solutions**, using a framework developed by Finnish teachers and researchers. The framework is not aligned with a specific evidence standard or curriculum, but with the **actual pedagogy** of a given solution. EAF has created **the national EdTech quality criteria** for Finland's National Agency for Education and recently merged with EdTech Impact (EI) to provide pedagogical certifications under the EI's Evidence Framework. The EAF evaluation parameters are based on a theoretical paper developed by **researchers at the University of Helsinki** and are used by **EAF affiliated teachers** for each evaluation.

EAF's pedagogical approach looks at **four parameters of effective learning with EdTech:**

- Passive / Active
- Linear / Non-linear
- Individual / Collaborative
- Rehearse / Construct

Each of the four parameters receives a score on the scale of 1-5 (with the second option of each pair of criterion being the higher scoring). The EdTech company receives an evaluation report that describes the EdTech product in pedagogical terms, and the match between the product's design and EAF evaluation criteria. Certified products are listed in EAF's own catalogue and on the EdTech Impact's website with a dedicated badge.

For more information see:
https://educationalliancefinland.com/

DIGITAL PROMISE

Digital Promise is a non-profit organisation that provides certifications for '**Research-Based Design**' and '**Learner Variability**'. In 2023, they added the new certifications: '**Prioritising Racial Equity in AI Design**', '**Research-Based Learning Analytics**' and '**CAST UDL Product Certification**'.

Research-Based Design

Assesses the extent to which **the product's design is based on rigorous research** (defined by Digital Promise as based on a peer-reviewed study with a minimum of 30 learners) and has a **theoretical framework** showing the research used by the EdTech developers in designing their product. In addition, the EdTech company is expected to publicly share their research basis, in either a publication, blog, or video. The research that underpins the product does not need to be an effectiveness or efficacy study directly conducted with the solution, but can be a research study that informs the design and development of the solution.

The assessment rubric classifies products at three levels:

- **Promised Commitment** — when the product team shows commitment to incorporate research findings to inform product design;
- **Honourable Mention** — when the product is informed using at least two research studies that are directly linked to the product design and that are peer reviewed in reputable journals; and
- **Certified** — when the product demonstrates grounding in rigorous research. Companies are requested to support this with an annotated bibliography with specific links to product points and design elements.

At the time of writing it is not clear whether Digital Promise accepts new applications for this certification. For updates, follow:

https://productcertifications.microcredentials.digitalpromise.org/explore/research-based-design-product-certification-2

Learner Variability

This certification evaluates whether an EdTech solution supports **learners with different ways of thinking, feeling, and personal backgrounds**. It gives clear examples of how an EdTech solution can benefit different kinds of learners, and what features are included that learners can adapt to fit their needs. It evaluates whether accessible information about how the product supports all kinds of learners is available. A certified product needs to showcase how designers and developers incorporated teachers' feedback to ensure the product supports diverse learners.

Research-based learning analytics

The certification is awarded to EdTech products that can demonstrate evidence of **comprehensive research on learning analytics**, and how this has influenced and steered their design choices. Similar to the Research-based certification, EdTech solutions become certified when they have a theoretical framework that showcases their research base, and make that research readily available and publicly accessible, such as in a white paper, blog, or video.

Prioritizing racial equity in AI design

This certification is intended to showcase EdTech products that prioritise and emphasise **the concept of racial equity** at every stage of their creation and development process. It is a recognition given to EdTech products that are committed to ensuring **fairness, inclusivity, and equal opportunities** for all racial and ethnic groups, from the initial design phase through to the final product. The certification was offered by Digital Promise in partnership with the EdTech Equity Project, but is currently closed for submissions.

CAST UDL product

This new certification is offered in partnership with CAST at Universal Design for Learning (UDL) and draws on a **UDL framework** and **three core values:**
1. The product places a strong emphasis on ensuring that students have easy access and are actively engaged;
2. The product takes into account the interests and motivations of learners and offers various avenues for understanding;
3. The product provides multiple means for learners to express their knowledge and understanding.

For more information see:
https://productcertifications.digitalpromise.org/

ISTE STANDARDS AND ISTE SEAL

The **ISTE Seal** focuses on **five areas targeted by EdTech**:

- Assessment Tools,
- Curriculum,
- Student Resources,
- Educator Resources,
- Assessment Resources.

To earn the ISTE Seal, products are evaluated by a panel of experts.

The standards are aligned to **UNESCO's Sustainable Development Goals** and are grouped by their relevance for students, educators, education leaders, and coaches. Each standard has sub-criteria, which are fully explained and illustrated with videos and examples on the ISTE website.

ISTE Standards are for the use of technology in teaching and learning, published by the **International Society for Technology in Education (ISTE)**, a nonprofit membership association for EdTech educators. The ISTE Standards framework is versatile, catering to students, educators, administrators, coaches, and computer science educators. It aligns with the criteria set forth by the **ISTE Seal,** an accreditation that EdTech companies can seek to attain after they reach a certain maturity stage.

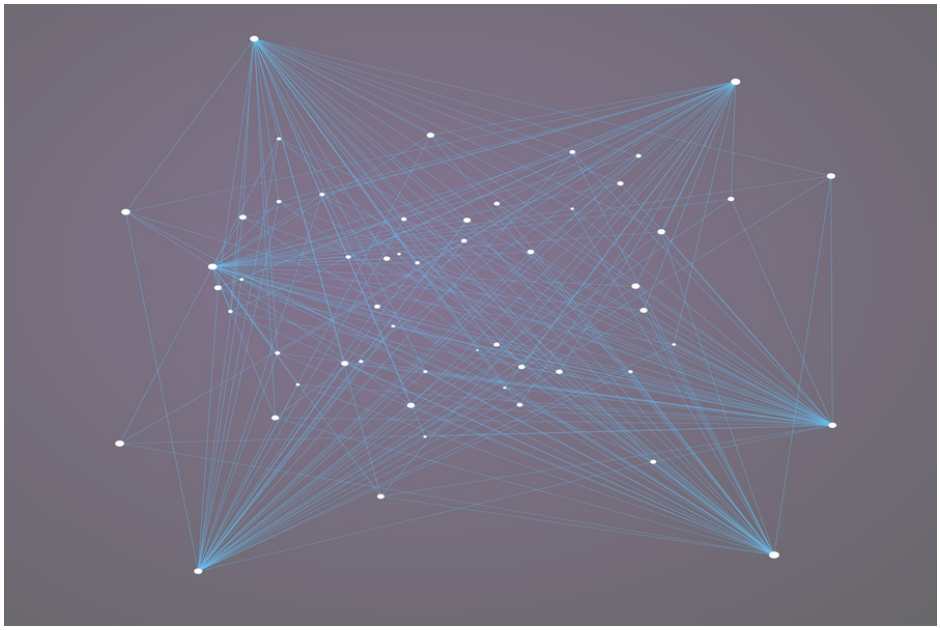
To earn the ISTE Seal, products are evaluated by a panel of experts in education, instruction, and technology. These experts assess the solution's readiness for learning in **three crucial areas**:

- alignment with the ISTE Standards,
- practical usability, and
- digital pedagogical implementation.

For more information see:
https://beta.iste.org/iste-seal

Pedagogical certifications are intended to support teachers in selecting **EdTech that is based on sound pedagogical principles.** Some certifications also pay attention to how the products address holistic child development and whether they are grounded in research. The certifications listed here are not concerned with **data safety or security**, such as the product's compliance with a national standard (e.g., COPPA in USA or GDPR in Europe), nor whether the products are interoperable.

Certifications can be showcased on individual product websites, and in fact, most EdTech companies display these badges alongside industrial design awards, educational awards, and testimonials for their products. Furthermore, certification providers list all certified products on their respective websites, and aggregator platforms, like EdSurge Product Index and EdTech Impact**,** include these certifications as a means to facilitate comparisons and evaluations of the various EdTech solutions accessible to schools.



NATIONAL FRAMEWORKS

While evidence as an imperative for national policy-making is mentioned in many official documents laying out **countries' digitisation strategies**, it is not systematically followed through in all countries.

Countries differ in how they define, incentivise and enable EdTech evaluations. Only a few countries have established clearinghouses to monitor and catalogue evidence-based educational resources (for example, What Works Clearinghouse in the USA or the Education Endowment Foundation in the UK).

Debate persists in the educational field regarding the usefulness of clearinghouses that follow a "medical model" for education. EdTech industry representatives are often concerned that **punitive measures around evidence could standardise and stifle innovation**, in an industry that survives and thrives on continuous, creative, and rapid development. Most agree that **globally shared EdTech quality benchmarks are needed**, but these should be formative and accommodate **diverse evidence needs**, and not follow just one rigid framework of evaluation. Such standards are being discussed at the time of writing, but do not yet exist. In the meantime, individual countries have developed their own standards that they expect EdTech providers to meet.

To illustrate the situation, **we selected four countries** (USA, Australia, India, and the UK) with national education evidence standards that apply to EdTech. These examples of national frameworks of evidence standards are widely recognized and have influenced other nations in formulating their own standards. Additionally, these countries have established a **robust infrastructure for fostering EdTech**, encompassing the conduct of evidence studies, educating teachers on evidence interpretation, and endowing procurement teams with the authority and financial resources to select interventions in accordance with national standards.

In the following sections, we describe these standards in more detail. However, it is important to note that the emphasis placed on each example within this report does not imply that one type holds more prominence than another. The significance of each country's standards hinges on factors such as the EdTech's target market, its prioritisation in meeting teachers' requirements, and its alignment with local government policies.

USA: ESSA STANDARDS

One of the more noteworthy international evidence standards, considering longevity, rigorous implementation, funding incentives, and infrastructure, is the US ESSA Standards of Evidence.

Non-Regulatory Guidance titled '**Using Evidence to Strengthen Education Investments**' was produced by the Department of Education in 2016. Its purpose was to provide guidance to US State and Local educational agencies, individual schools, educators, and partner organisations in how to select and use evidence-based tools and interventions, including but not limited to EdTech. This guidance was adopted and amended from the Elementary and Secondary Education Act of 1965 (ESEA), and the Every Student Succeeds Act of 2015 (ESSA).

The guidance advises educators to consider **evidence behind an intervention or approach** at **four levels,** and makes it clear that "Interventions supported by higher levels of evidence, specifically **strong evidence or moderate evidence,** are **more likely to improve student outcomes** because they have been proven to be effective." In the absence of strong evidence or moderate evidence, schools are advised to consider **promising evidence** and solutions that have a rationale for how they impact children.

Once the evidence supporting an intervention is established, critical monitoring and evaluation decisions must be made as the intervention advances. Educational stakeholders are expected to engage in **reflection and ongoing monitoring of progress.** This entails addressing questions such as whether adjustments are necessary to enhance implementation, or if the intervention is prepared for broader scaling to reach more students or educators.

EdTech companies must consider **the entire cycle**, with a crucial emphasis on comprehending the significance of evidence within the U.S. context. The U.S. government has established and provided training for educators, assessors, and EdTech evidence providers, which align with the defined individual levels of evidence. **These four levels are structured hierarchically**, ranging from promising to strong, each accompanied by **specific criteria.** Comprehensive explanations and guidance documents for these levels are available on the Office of Educational Technology's website, as well as through the What Works Clearinghouse and John Hopkins University.

To attain evidence at **Tier 1 (Strong evidence)** or **Tier 2 (Moderate evidence)** levels, EdTech founders are encouraged **to collaborate with local universities** through the university school research partnerships program. In conjunction with the universities, the school district technology teams should develop an approach tailored to their technology readiness and the desired level of evidence, and conduct a study designed to move them up from **Tier 3 (Promising evidence)** or **Tier 4 (Demonstrates a Rationale)**.

Some of these partnerships are supported by EdTech Evidence research consultancy providers, offering **evidence as a service and conducting ESSA-specific studies** upon request on a paid contract basis. This approach can be useful, given that the stringent and rigorous evaluation methods mandated by the government require qualified and trained researchers. These methods include tasks such as identifying and implementing high-quality outcome measurements, relying on administrative data collected by schools or classroom observations conducted by researchers. The data are then subjected to **appropriate statistical analyses to discern the relationships between outcomes and use of the EdTech solution**.

ESSA TIERS OF EVIDENCE

ESSA Tier 1: Strong Evidence

Strong evidence is acquired through, for example, an experimental study with a **Randomised Controlled Trial** (RCT) design, collaboratively planned and executed by both industry and academia. Such a study would have assessed the **impact of an EdTech-based intervention on student outcomes** by comparing the progress in outcomes between two groups while controlling for the students' prior development in those outcomes before the intervention.

The intervention should involve a **sufficiently large sample** that shares characteristics with the target population. Such an example represents **strong evidence because it is a well-implemented experimental study capable of establishing causal inferences.** Causal inference is drawn through the demonstration of a **statistically significant difference** between the two groups, and the intervention is considered successful when this difference indicates a positive outcome favouring the intervention.

ESSA Tier 2: Moderate Evidence

Moderate evidence is acquired when the EdTech intervention undergoes a **well-conceived and well-executed quasi-experimental study**, resulting in a **statistically significant** and positive (i.e., favourable) impact of the intervention on student outcomes or other pertinent measures. The **participant sample should be substantial,** representing a multi-site approach that aligns with the target population the intervention seeks to serve.

An example study for Tier 2 involves an intervention employing a quasi-experimental design coupled with regression analysis to examine the relationship between student outcomes in the intervention group versus the control group (or intervention group versus control schools). This analysis **should account for potential confounding factors**. This level of evidence is frequently chosen when randomization across classrooms or schools is not feasible, and groups are compared while controlling for pre- and post-intervention effects.

ESSA Tier 3: Promising evidence

Promising evidence can be acquired through a **correlational study** where the relationship between one and another factor is examined, while controlling for the influence of possible bias (e.g. the students' or schools' characteristics).

The intervention is typically voluntary, in that some students in the classroom may use the EdTech solution, while others do not (and the scores are measured in all students). For this Tier of evidence it is also **recommended** - as it is for the other Tiers - **that the research is conducted by an external team of researchers** who partner up with the schools and the EdTech provider. If there is a finding of positive evidence, a check should be made that no negative evidence of the intervention or approach, meeting the same rigorous criteria, exists.

ESSA Tier 4: Demonstrates a Rationale

This **initial level of evidence** is based on the premise that a technology can have a positive impact if it is **based on a logic model** that was developed by research and evaluation of the specific features of the product.

Typically, schools would conduct a needs analysis in their classrooms, and with the help of an **external research team**, develop a logic model outlining the rationale, activities, and relevant outcomes of integrating the intervention into classroom practice. This can be used for **pilot-testing the EdTech** and finding out whether its use is feasible in the classroom. The results of this pilot test, together with the logic model, can be used to plan an empirical study for Tier 3.

For more information see:
https://tech.ed.gov/

AUSTRALIA: AERO STANDARDS

The Australian Education Research Organisation (AERO) is Australia's independent education evidence body, responsible for setting and following up on evidence standards. The Australian Standards of Education criteria are applicable to all varieties of educational evidence, whether it is produced via scholarly research or by educators in their day-to-day activities.

AERO recommends that educators use the standards to **ascertain the strength of the existing evidence** base for a particular approach, including EdTech, in a particular classroom setting. The standards can also be used prospectively, when planning and developing the design of an approach.

AERO provides various **Evidence decision-making tools** to support educators and policymakers in using the standards. These can also be useful for EdTech founders, as they develop their understanding of the evidence criteria.

The standards are arranged in **four levels of confidence** in the evidence basis, from Level 1 (low confidence) to Level 4 (very high confidence), as explained on the following page, with descriptions adjusted to the EdTech context.

AERO STANDARDS

Level 1: Low Confidence

Type of research:

The EdTech's positive impact is explained but it **lacks empirical data**, whether qualitative or quantitative, to support its assertion regarding the approach's effectiveness.

Features of studies:
The studies supporting an EdTech solution offer an **explanation rooted in established theories of learning and development**. They also provide a clear, detailed explanation of how the approach is theorized to generate positive effects, breaking down the process step by step.

Level 2: Medium Confidence

Type of research:

This research illustrates a **correlation between the approach and favourable outcomes**, relying on various study types, including both small-scale (e.g., case studies) and large-scale (e.g., cross-national surveys) investigations. These studies employ a **range of methods**, such as qualitative (e.g., observations and interviews), quantitative (e.g., statistical techniques), or mixed methods. This research does not conclusively establish causation, as there may be other potential explanations for the observed positive effects.

Features of studies:
The studies conducted with the EdTech resource were in the classroom/setting similar to where it is intended to be implemented, reinforcing findings from diverse studies conducted across various settings. It **tracks changes in outcomes** over a period, utilising a sizable sample distributed across **multiple sites**. Employing rigorous methodologies that minimise the likelihood of chance effects, the study compares a group exposed to the approach with a group that has not experienced it. Studies are recent, and conducted by individuals or entities independent of the EdTech developer.

Level 3: High Confidence

Type of research:

The EdTech resource needs to be based on research that utilises **rigorous qualitative, quantitative, or mixed methods research**, to address factors like selection bias, history effects, and maturation effects. In addition, the research should employ outcome measures specifically validated for the study's objectives. This research does not provide causal evidence.

Features of studies:
Same as Level 2, but in addition, the studies take measures to reduce the possibility that effects result solely from the specific attributes of study participants. Furthermore, it explores and / or assesses **contextual factors** that could impact the approach's effectiveness.

Level 4: Very High Confidence

Type of research:

The EdTech resource is based on research that meets all the above criteria, and in addition, summarises findings from **rigorous research via a systematic review or meta-analysis** of studies conducted across various contexts or within contexts akin to the target classroom / use context.

Features of studies:
The studies confirm **consistent findings across diverse contexts**, identify the factors contributing to the approach's success and the prerequisites for its broader implementation, evaluate the approach's effectiveness across various subgroups, provide explanations for any disparities in effectiveness among these subgroups, and **continuously track outcomes for different groups** over time to ensure sustained effectiveness.

For more information see:
https://www.edresearch.edu.au/using-evidence/standards-evidence

INDIA: EDTECH TULNA

India does not have national EdTech standards as such, but the Tulna EdTech standards are widely used. The EdTech Tulna standards establish a set of guidelines for the design of EdTech products. These standards are organised across three dimensions: Content Quality, Pedagogical Alignment, and Technology & Design, with each dimension containing detailed standards.

Tulna is an EdTech evaluation index for product design quality and was developed by the Central Square Foundation with the Indian Institute of Technology in Bombay.

Content quality focuses on the accuracy of the content, its alignment with the national curriculum, and the appropriateness of the content for the targeted grade and intended learners. The standards evaluated for this dimension include:

- Content accuracy
- Correctness and clarity in assessment
- Language comprehensibility
- Alignment to national standards (skill coverage)
- Curriculum alignment
- Inclusivity in learner representation
- Bilingual use (for vernacular products)

Pedagogical alignment evaluates whether the product adopts **learner-centred methods** and integrates learning science theories and pedagogical design principles to craft a **meaningful learning journey.** It assesses the **degree of congruence between the pedagogical strategies and assessments and national educational policies**. The standards evaluated for this dimension include:

- Constructivist approach
- Addressing alternate conceptions
- Content in context
- Learner scaffolding
- Cognitive engagement
- Motivational features
- Logical chunking and connectedness (for products containing audio-visual resources)
- Learning objective alignment
- Pedagogy-assessment method alignment
- Cognitive levels covered
- Feedback quality
- Opportunities for collaboration
- Adaptivity (for PAL products)
- Teacher support
- Facilitating goal-setting (for IAV products)

Technology and design assesses the level of integration between technological features and pedagogy/content to enhance a meaningful learning experience. The standards evaluated for this dimension include:

- Interface design (intuitive use)
- Interface design (assess consequences of an action)
- Learner navigation & pace universal design
- Analytics for learners' progress
- Tools to support problem-solving
- Meaningful interactivity
- Content-type-technology alignment

For more information see:
https://www.edtechtulna.org/standards

UK: NESTA STANDARDS OF EVIDENCE

There are several frameworks and standards of evidence that focus on educational standards and EdTech in the UK, including the What Worked EdTech standards and an accompanying full framework offered by EdTech Impact. National standards are expected to be developed in the near future but were not available at the time of writing.

Nesta Standards of Evidence are **not specific to EdTech** but are focused on innovations more broadly, and as such, often deployed in relation to EdTech. The standards provide a framework for selecting an innovation that "works" and are accompanied with **toolkits** that guide decision-makers on evaluating the **quality of research behind an innovation** and its appropriateness for adoption in a specific context. The **five levels of the NESTA standards** are:

Level 1:
Describe the logic model and theory of change

Level 2:
Gather data showing positive change but not confirming causality

Level 3:
Demonstrate causality using a control or comparison group

Level 4:
Confirm causality with 1 or more independent replication evaluation

Level 5:
Develop manuals, systems, and procedures to ensure consistent replication and positive impact

The Education Endowment Foundation (EEF)'s standards are part of the Teaching and Learning Toolkit, which **systematically summarises evidence regarding specific learning approaches or interventions**. The EEF provides various indicators to help educators gauge the expected impact, cost, and effectiveness of a particular intervention.

Impact is assessed across **12 tiers**, each denoting an additional months' progress gained based on effect size. Interventions receive a padlock rating (**ranging from 1 = very limited evidence, to 5 = very extensive evidence**) depending on the strength of evidence found in systematic research studies that meet specific inclusion criteria. **Interventions with fewer than 11 qualifying research studies are assigned O rating**. The ratings increase with the number of studies conducted; for instance, an intervention would need between 45 and 60 qualifying studies to receive a three-padlock rating, indicating moderate evidence.

Additional quality metrics include independent evaluation if there's substantial unexplained variation among study results, how recently studies were conducted, and whether they were conducted in real-life contexts. **Randomised controlled trial studies are considered the top type of evidence** for establishing causality.

For more information see:

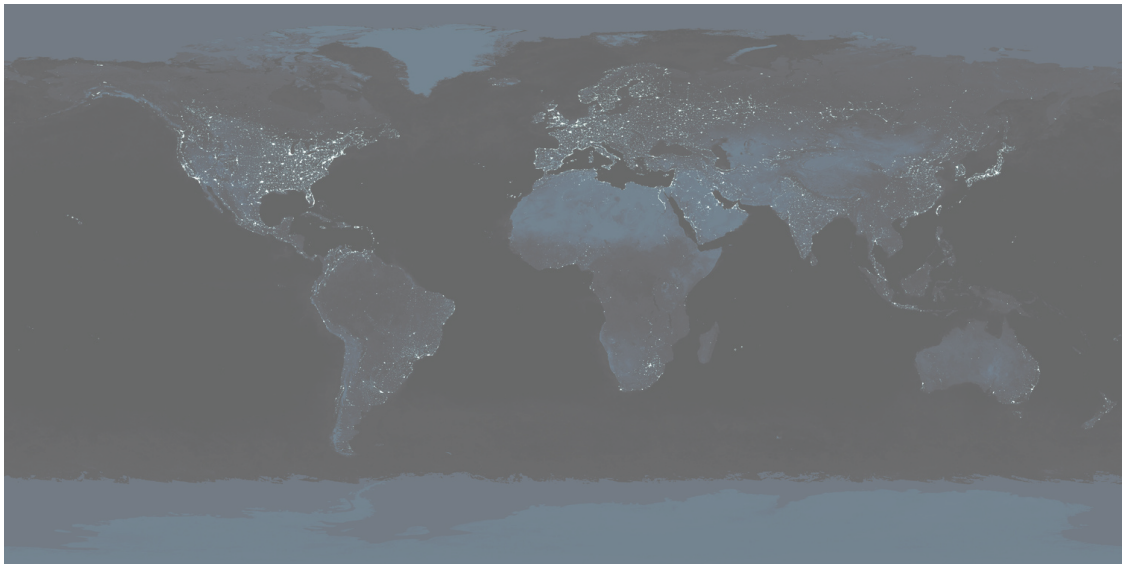
https://interventions.whatworked.education/edtech

https://edtechimpact.com/knowledge-hub/new-framework-launched-to-drive-edtech-quality-worldwide/

https://www.nesta.org.uk/feature/innovation-methods/standards-evidence/

https://educationendowmentfoundation.org.uk/

It remains to be seen whether the current national standards in some countries will be adopted by others, or whether individual countries will develop their own. For **EU countries**, the strategy of many has been to wait-and-see whether there will be a joint EU regulation regarding evidence, as was the case for data use with the GDPR. In the meantime, individual countries rely on their own frameworks and often delegate the task of evaluating and selecting EdTech to individual teachers or schools and districts.



INTERNATIONAL FRAMEWORKS AND STANDARDS

Internationally recognised organisations such as **UNESCO,** the **World Bank**, and the **OECD**, play a **pivotal role in advancing EdTech**. They foster cross-stakeholder collaboration and promote the integration of innovative digital solutions into education systems worldwide. Through research, policy development, and capacity-building initiatives, these organisations strive to **bridge the digital divide**, ensuring equitable access to quality education for all. Additionally, they support the development of guidelines and standards for EdTech implementation to **enhance its effectiveness and impact on learning outcomes**. Typically, the organisations work closely with communities, governments, and businesses, which strengthens their mission of making the use of EdTech more equitable and safe, and to improve EdTech's quality through innovation and alignment with national priorities.

It is important to be aware of the frameworks followed by these organisations, because in many instances **they dictate or guide procurement decisions on a local or national level**. The recommendations provided by these organisations are neither statutory, nor regulatory; but given the gravitas of these organisations, their expertise is influential. EdTech founders should be aware of the quality criteria espoused by these organisations, especially if they have ambitions of cross-national scaling.

The current EdTech funding landscape suggests that the focus on international guidelines can benefit EdTech companies aiming to engage in business-to-government (B2G) relationships, selling directly to governments, as venture capital (VC) and limited partner (LP) funding in EdTech has recently declined. The reports commissioned by these major organisations offer valuable insights into areas that can be leveraged to promote an EdTech agenda. These areas are currently underutilised by EdTech but have been identified by international organisations as pivotal for advancing **not only effective learning, but also equitable and inclusive education**.

THE WORLD BANK

The World Bank maintains a dedicated cohort of experts in the field of educational technology (EdTech), supplemented by personnel tailored to specific countries. These experts are tasked with offering substantial support for the judicious selection and implementation of suitable EdTech solutions at the grassroots level. This initiative is orchestrated under the guidance of the World Bank EdTech Strategy, which makes it imperative that education systems adhere to **five fundamental principles** when undertaking investments in EdTech. The two principles most relevant for EdTech founders are principles 2 and 3:

Five principles of the **World Bank EdTech strategy:**

1. ASK WHY
2. DESIGN FOR SCALE
3. EMPOWER TEACHERS
4. ENGAGE THE ECOSYSTEM
5. DATA DRIVEN

For more information: https://www.worldbank.org/en/topic/edutech

PRINCIPLE 2: DESIGN FOR SCALE

EdTech design must prioritise **flexibility and user-centeredness**, and place e**quity and inclusion** at its core to ensure widespread and lasting impact. Achieving scalability begins with engaging proactively and empathetically with all potential end-users, such as students, teachers, administrators, and parents, as each user group may reveal unique needs. By comprehending and addressing these diverse requirements, EdTech solutions can become more i**nclusive, equitable, and adaptable**, mitigating the current disparities observed in education systems, and fostering sustainable scalability.

PRINCIPLE 3: EMPOWER TEACHERS

Technology should empower teachers by providing access to content, data, and networks, enabling them to concentrate on **individualised student learning**. EdTech doesn't replace teachers, but it can **enhance their role**. Evidence worldwide indicates that as EdTech is effectively integrated, teachers' roles become more central, encompassing diverse responsibilities, from facilitating learning to collaborating with mentors and leading project-based activities. In contexts with teacher shortages or limited capacity, technology can assist learners in overcoming these challenges, while also supporting teachers in areas where they may lack expertise. This teacher-technology partnership empowers educators to deliver personalised learning experiences to students by leveraging a wide range of resources.

SMART BUYS BY THE WORLD BANK AND UNICEF

The Education Evidence Advisory Panel (GEEAP) is an initiative by **the World Bank and UNICEF** intended to bring together **leading international experts** with the aim of ascertaining key interventions based on evidence.

The GEEAP panel produces a catalogue of '**Smart Buys'**, which lists EdTech categorised according to a synthesis of quantitative and qualitative evidence with attention to context, scale, and equity. Whether such catalogues will be used by individual countries as accountability tools, remains to be seen.

GEEAP has recently published a guide aimed at informing governments and stakeholders in low- and middle-income countries on effective strategies for enhancing learning and education outcomes. Coordinated by the **Foreign, Commonwealth & Development Office (FCDO) of the World Bank, UNICEF**, and **USAID**, GEEAP recently updated its 2023 Smart Buys guide for policymakers. This revised guide incorporates additional research and expert consensus, and offers **evidence-based recommendations** regarding EdTech investments. EdTech solutions are categorised as '**Great**', '**Good**', '**Promising but Limited Evidence**', '**Effective but Relatively Expensive**', or '**Bad**' buys, focusing on **cost-effective scalability** for policymakers worldwide. It underscores the interplay between software and hardware in educational technology interventions, emphasising that software can be a worthwhile investment when suitable hardware is already in place, aligning with findings from the UNESCO Gem Report.

SMART BUYS

The following table is adapted from: *2023 Cost-Effective Approaches to Improve Global Learning - What does recent evidence tell us are "Smart Buys" for improving learning in low- and middle-income countries?* (K. Akyeampong, T. Andrabi, A. Banerjee, R. Banerji, S. Dynarski, R. Glennerster, S. Grantham-McGregor, K. Muralidharan, B. Piper, S. Ruto, J. Saavedra, S. Schmelkes, & H. Yoshikawa; London / Washington D.C. / New York; FCDO, the World Bank, UNICEF, & USAID.)

GREAT BUYS	Highly **cost-effective** and supported by a **strong body of evidence** (e.g., targeted instruction by learning level)
GOOD BUYS	**Good evidence** that they are cost-effective (e.g., providing parent-directed early childhood stimulation programs)
PROMISING BUT LIMITED EVIDENCE BUYS	Some rigorous studies that show high levels of effectiveness, but evidence on or **examples of implementation at scale are lacking** (e.g., personalised learning that adapts to the learning level of the child, where hardware is already in schools)
EFFECTIVE BUT RELATIVELY EXPENSIVE BUYS	Relatively expensive but with **evidence of delivering learning outcomes** (e.g., cash transfer as a tool for improving education outcomes)
BAD BUYS	**Repeated evidence** that they do **not positively impact learning** and are not cost-effective (e.g., investing in hardware like laptops,tablets and computers alone)

EDTECH HUB

EdTech Hub is a global research partnership, which aims to "empower people by giving them the evidence they need to make decisions about technology in education". EdTech Hub operates as a **consortium of nine partners:** Bill and Melinda Gates Foundation, Brink, Jigsaw, Open Development and Education, Results for Development, The World Bank, UK Aid, UNICEF, and University of Cambridge. The focus is on **academic research that demonstrates and supports the evidence of EdTech's impact.**

The EdTech Hub conducts research studies in five topic areas:

- Data for Decisions
- Digital Personalised Learning
- Girls' Education & Technology
- Participation & Messaging
- Teacher Continuous Professional Development

These studies rely on quantitative and qualitative research on the ways in which technology can or should be used, the development of new and improved designs through the "sandbox methodology" of innovation, and timely support to governments and policymakers on evidence-based solutions. As a global research partnership, the EdTech Hub aims to address the lack of evidence in EdTech with reliable research and advice, and a suite of supporting research tools (such as survey protocols and observation templates) for EdTech communities to run their own research.

Through collaboration with partners, EdTech Hub is involved day-to-day in ensuring evidence-driven approaches to EdTech, informing the work of UNESCO and other partners. Their work is most recently illustrated in the GEM Report.

UNESCO

As a specialised agency of the United Nations, **The United Nations Educational, Scientific and Cultural Organization (UNESCO)** encompasses several initiatives focused on EdTech. The most recent and relevant for EdTech founders are the activities reported in the **Global Education Monitoring team's report "Technology in education: A tool on whose terms?"** (the 'GEM Report'). The report was released along with a #TechOnOurTerms campaign, which advocates for decision-makers in education technology to **prioritise the needs of learners**, and to conduct assessments to determine the appropriateness, equity, evidence-based nature, and sustainability of applications. The approach offers a guiding framework for policymakers to follow when making such decisions.

In particular, the report emphasises that policymakers and educators should **approach EdTech with questions around education and not technology**. Then, the real question is how can we improve the quality of education. Here, EdTech founders should note the crucial connection between a tool's affordances (for example, the specific parameters of a platform) and the availability of technology infrastructure in the country of implementation. These must be considered, together with teachers' **digital literacy**, and their ability to take on technology for their practice in alignment with the national curriculum. The report emphasises the importance of educational outcomes encompassing an understanding of education in its broader sense, and the impact that an EdTech may have on social, economic, and environmental outcomes.

"The definition of quality in an education system should encompass the system's ability to equip learners to act in ways that help achieve sustainable development in the social, economic and environmental senses"

(GEM Report, 2023, p. 10)

The main premise of the report is that **technology can improve quality in teaching and learning if** it addresses two areas:

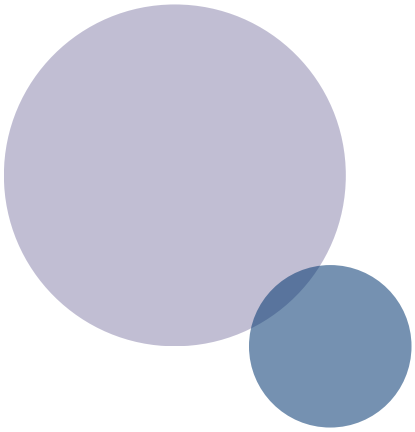
1. The need to decrease the gaps between teachers who have, and teachers who don't have, time to develop and provide best pedagogical approaches and personalised instruction.
2. Students' engagement in learning through varied content, interaction, and collaboration by design.

Educational opportunities are threatened by the **global challenges** of low digital skills, poor communication and collaboration among key stakeholders responsible for high-quality digital learning, and rapid cycles of development in EdTech.

EdTech's ability to make education management more efficient relies on **safe and secure use of data.** However, data safety and security is not guaranteed by current EdTech deployments, and the GEM Report calls on EdTech providers and users to urgently address data safety and **students' privacy** before devices and platforms are placed in children's hands.

While the press coverage of the report's findings has morphed into a recommendation to ban technology, the report itself makes a series of nuanced statements about when EdTech does, and when it does not, support students' and teachers' agency.

It is the pertinent question of "On Whose Terms?" that should be legitimately discussed by all involved in EdTech design and propagation.



UNICEF

UNICEF stands as a pioneering force in public sector innovation. The organisation has demonstrated that by garnering support from governments, forging strategic partnerships, and fostering collective action, innovative solutions can be conceived and disseminated. UNICEF's Global Innovation Strategy aims to identify gaps and potential areas for development, and thereby ignite innovation.

EdTech cuts across several UNICEF units. For example, UNICEF Office of Innovation provides advice to EdTech founders on various levels, including the development of business models that share open-source software and still create value and revenue. In 2021, UNICEF established two innovation hubs in Finland to strategically enhance the development of innovative approaches to EdTech, through, for example, the development of international policy guidance on AI for children and a universal framework of recommendations for selecting effective EdTech.

UNICEF's Venture Fund invests in early- and growth-stage startups to accelerate **open-source technologies** (as of 2023, there were 74 countries where UNICEF has made such investments).

UNICEF's Innocenti digital learning research investigates how EdTech are developed and implemented across individual countries and provides a set of tools to support individual countries to monitor their national EdTech strategies and establish sound monitoring and evaluation routines.

Many reports from offices operating under **UNICEF's digital learning initiative** provide direction for the global adoption of EdTech. For example, the report *Responsible Innovation in Technology for Children Digital technology, play and child well-being* (RITEC), funded by the LEGO Foundation, provides tangible resources for both businesses and governments for prioritising children's well-being in digital design. This report introduces a recently formulated **well-being framework** for children, consisting of eight outcomes centred around the child's needs and wellbeing.

The framework recommends that all technologies aimed at children should adhere to these principles, and that indicators for compliance with these principles should be monitored and considered when adopting new technologies. Research to determine the specific play designs and mechanisms that contribute to particular well-being outcomes is due for publication at the end of 2023.

#RITEC

The eight principles in the RITEC framework are outcomes that are relevant for children's well-being with digital play:

1. Competence: Engaging in digital play experiences has a **positive impact** on children's self-perception of their competence, knowledge, and abilities.
2. Self-actualisation: Children's digital play gives them a **sense of purpose** and enhances their **social engagement** and self-esteem.
3. Empowerment: Participating in digital play should foster children's **autonomy, choice, and agency,** allowing them to feel in control, make decisions, and achieve a sense of mastery.
4. Emotional regulation: Children should be able to utilise the digital environment to regulate their mood, relax, and recharge in preparation for interactions with peers and the world. This involves using digital play experiences to **alleviate stress** and enter **flow states** characterised by **intense focus, deep engagement, and enjoyment** in an activity.
5. Social connection: The digital environment, including play, should enable children to **connect socially** with peers, family, or other significant individuals in their lives while ensuring their **safety** and protection from harm.
6. Safety and security: Children should experience a sense of safety, and actually be safe, while participating in the digital environment and digital play. This includes **protection** from various risks, including but not limited to contact risks, conduct risks, content risks, and contract risks.
7. Creativity: Children's involvement in digital play should **foster their curiosity**, cultivate receptiveness to novel encounters, and enhance their creative skills.
8. Diversity, equity and inclusivity: Digital play experiences should be **varied, fair, and inclusive**, ensuring that children from diverse backgrounds and circumstances can engage. These experiences must also be accessible to children with disabilities, and suitable in terms of age and culture.

These outcomes can be interpreted as **child-centred indicators of high-quality technologies** that prioritise **children's wellbeing.** While the framework doesn't provide specific guidance on the exact play designs or mechanisms that lead to specific well-being outcomes, it serves as a strong foundational resource for emphasising key aspects when designing digital experiences for children. Additionally, it aids in identifying child-centric goals for policy development, legislation, and regulation.

Another recent report concerning EdTech quality developed by the UNICEF Office is the ***Pulse Check On Digital Learning***. This report provides recommendations for financing and developing **ICT Education policies** and centres the question of cost-effectiveness of EdTech alongside the questions of educational effectiveness.

Further examples of EdTech-focused reports, not specifically focussing on effectiveness but discussing questions of impact, include

- UNESCO's *ICT in Education Policy Toolkit*,
- UNESCO's *Guidelines for ICT in education policies and materplans*,
- The World Bank's *Building and sustaining national educational technologies agencies: Lessons, Models and Case Studies From Around the World*
- and specifically dedicated to cost-effectiveness, The World Bank's *Template for costing remote learning encouragement*.

ASIAN DEVELOPMENT BANK

While the Asian Development Bank is not a typical “player” in EdTech evaluations, the QESA quality, effectiveness, scalability, affordability evaluation framework of EdTech is a well-known resource for establishing the extent to which EdTech are of high-quality and scalable. The QESA Framework divides its focus into several subcategories for each of the dimensions. These are:

Quality: Functionality, Pedagogy, User-friendliness
Effectiveness: Evidence, User satisfaction, Impact
Scalability: Infrastructure, Local adaptability, Security and Privacy
Affordability: Investment costs, Operational costs, Sustainability

It is the ‘**Effectiveness’** that is most directly related to the question of evidence of impact. Effectiveness is the **extent to which an EdTech solution supports the intended goals and students’ needs,** and the extent to which this support is **based on rigorous research and evidence of impact**. To rate a product’s effectiveness, evaluators are encouraged to request details concerning a given EdTech’s research foundation, customer satisfaction, and popularity. The evidence might include research papers, case analyses, client endorsements, and user feedback on platforms like the Appstore and EdTech repositories.

The framework is described with use cases, a guide for application, and a freely downloadable Excel-based QESA EdTech Evaluation Tool. More details concerning the Framework are in the report *Reimagine Tech-Inclusive Education: Evidence, Practices, and Road Map* (2023, Asian Development Bank). Before QESA is applied, policy-makers must evaluate the country’s readiness to adopt and scale EdTech. Once that is established, governments are encouraged to choose 3-5 EdTech vendors for evaluation, based on prioritised requirements and product reviews.With these selected vendors, governments should develop use cases and invite the vendors to demonstrate the expected impact according to the framework’s quality criteria.

Funders' and investors' own evaluation frameworks

When it comes to impact, EdTech providers are accountable to learners, but also to their funders and investors who expect reports of impact and impact metrics from their portfolio companies. The extent and focus of impact reporting varies from investor to investor, and depends on how impact-focused specific Venture Capitalists (VC) are. In our experience, the more well-established and more learning-focused a VC is, the more they prioritise and expect their portfolio companies to prioritise impact metrics.

As an example, we include a statement by **Owl Ventures**, a large EdTech VC, on how they approach evidence, effectively distinguish between different types of EdTech, consider the maturity of the EdTech, and accommodate various forms of evidence suitable for each specific EdTech category.

> *We know that different types of research are appropriate to determine effectiveness according to product type and development stage. Therefore, we encourage our portfolio companies to report their impact along a spectrum, from formative research to summative research. Using this spectrum, early-stage companies might showcase their impact through user testimonials and descriptive case studies; conversely, a more-mature company will be able to provide rigorous research data, through participation in causal, comparative, or randomized controlled trial studies. At the core, we want to help companies answer one fundamental question:* ***Does your product have a positive outcome for its intended users?***
>
> (Owl Ventures Education Outcomes Report, 2022)

Given that this report offers only a partial view of the numerous EdTech evaluation frameworks, criteria, and requirements available in the industry, EdTech companies might find it daunting to navigate them all. However, these frameworks offer a language to guide the **development of child-centred, impactful, equitable, and innovative solutions.** As a result, we propose some recommendations for all EdTech companies looking to engage with any or all of these frameworks.

RECOMMENDATIONS FOR EDTECH

01 Prioritise evidence

We recommend that all EdTech providers **prioritise ongoing research integration into their development processes**. It is clear from the various types of reports and certifications that evidence of effectiveness and impact of an EdTech solution is of national and international interest, and EdTech solutions in 2023 need to align their processes with this reality. Considering the similarities across many of the standards described in this report (illustrated in the following rubric), the evidence that is gathered, across progressing levels of rigour, should be carefully and thoughtfully planned.

Evidence ready →		Evidence-based (emerging) →		Evidence-based (strong)
ESSA Standards of evidence	**Tier 4:** Demonstrates a rationale (rigorously developed logic model outlining the rationale, activities, and relevant expected outcomes from the interventions	**Tier 3:** Promising evidence (correlational studies controlling for possible bias, preferably conducted by independent researchers. Not supporting causal evidence)	**Tier 2**: Moderate evidence (well-conceived and well-executed studies finding statistically significant positive impact, accounting for confounding factors, but not fully supporting causal claims)	**Tier 1:** Strong evidence (rigorous intervention studies using best practices for supporting causality)
AERO Standards of Evidence	**Level 1:** Low confidence (explanation of impact but lack of empirical evidence)	**Level 2**: Medium confidence (correlational evidence, but no evidence of causation)	**Level 3**: Hifh confidence (rigorous validated evidence, addressing confounding variables, but not supporting causal claims)	**Level 4**: Very high confidence (rigorous systematic research that supports causal claims)
NESTA Standards of Evidence	**Level 1**: Evidence of logic model and theory of change	**Level 2**: Evidence of data showing positive change but not confirming causality	**Level 3:** Evidence demonstrating causality using a control or comparison group	**Level 4:** Evidence including 1 or more independent replication evaluation	**Level 5:** Evidence including manuals, systems, and procedures to ensure consistent replication and positive impact
World Bank & UNICEF's Smart Buys	**Effective but** relatively expensive buys	**Promising but** limited evidence buys	**Good buys** with good evidence that they are cost effective	**Great buys** with a strong body of evidence supporting high-cost effectiveness

02 Develop adequate metrics

As your company develops a research plan, it needs to be able to align it with the **standards of evidence that matter in your target market**. The plan should include measurements that are based on learning science principles, and that cut across pedagogical and national evidence priorities. To support this process, we recommend formulating a **clear impact statement**, that is then **translated into a set of monitoring and evaluation milestones**. The milestones should be underpinned by measurable metrics. This is part of an "evidence mindset" necessary for effective EdTech Evidence Evaluation Routine (Kucirkova et al. 2023).

03 Keep evidence as an ongoing strategy

Evidence is not a "thing to get" but rather an **ongoing cycle of evaluation and improvement**. To be truly evidence-based in the realm of EdTech therefore means centering research as an ongoing process, characterised by detours and continuous learning. There is no ceiling on how much research is needed to have evidence of what works. Rather, an EdTech's strategy should include evidence as an ongoing iterative process.

04 Invest in the right kind of evidence

Pedagogical certifications are important, but ultimately, it is the level of decentralised decision-making that determines what ends up in children's hands. In the USA, federal funding is tied to ESSA, while in Europe, it is in the hands of local municipalities. In many instances, teachers themselves may decide which app or platform are used in their classrooms. For global operations, we recommend that companies invest in a certification or independent evaluation against their impact strategy, rather than current trends. For example, while a certain certification might appear as essential at a specific point in time, its attainment should be aligned with the EdTech company's **long-term vision for impact**.

In conclusion

The existence of several evaluation frameworks is not a negative thing. Diversity in evidence evaluations can fuel innovation in EdTech, so long as these are guided by principles from the Science of Learning, such as promoting engaged and meaningful learning with social interaction. The existing variation in EdTech evaluations should propel business storytelling tailored to international evidence and national experience of what works. The **2023 EdTech winners** will be those who can tackle this delicate balancing act and **provide convincing narratives of educational impact supported by ethical and rigorous evidence.**

Cited research publications

[1] Kucirkova, N., Brod, G., & Gaab, N. (2023). Applying the science of learning to EdTech evidence evaluations using the EdTech Evidence Evaluation Routine (EVER). npj *Science of Learning,* 8(1), 35.

[2] Hillman, V. (2023). Bringing in the technological, ethical, educational and social-structural for a new education data governance. *Learning, Media and Technology*, 48(1), 122-137.

[3] Kidron, B., Pothong, K., Hooper, L., Livingstone, S., Atabey, A., & Turner, S. (2023). *A Blueprint for Education Data*. Digital Futures Commission, 5Rights Foundation.

[4] Kucirkova, N. I., Livingstone, S., & Radesky, J. S. (2023). Faulty screen time measures hamper national policies: here is a way to address it. *Frontiers in Psychology*, 14.

[5] Hooper, L., Livingstone, S., & Pothong, K. (2022). *Problems with Data Governance in UK Schools: The Cases of Google Classroom and ClassDojo*. Digital Futures Commission, 5Rights Foundation.

[6] Dore, R. A., Shirilla, M., Verdine, B. N., Zimmermann, L., Golinkoff, R. M., & Hirsh-Pasek, K. (2018). Developer meets developmentalist: improving industry–research partnerships in children's educational technology. *Journal of Children and Media*, 12(2), 227-235.

[7] Vanbecelaere, S., Adam, T., Sieber, C., Clark-Wilson, A., Boody Adorno, K., & Haßler, B. (2023). *Towards Systemic EdTech Testbeds: A Global Perspective*. https://discovery.ucl.ac.uk/id/eprint/10167465/

[8] Kucirkova, N. (2022). *Understanding Evidence: A Brief Guide for EdTech Producers*, Report for WiKIT, AS, University of Stavanger. https://doi.org/10.1007/s10643-022-01415-1

[9] Odgers, C., Allen, N., Pfeifer, J., Dahl, R., Nesi, J., Schueller, S., & Williams, J. L. (2022). *Engaging, safe, and evidence-based: What science tells us about how to promote positive development and decrease risk in online spaces*. https://psyarxiv.com/rvn8q/download/?format=pdf

[10] Kucirkova, N. (2023). Debate: Response to "Should academics collaborate with digital companies to improve young people's mental health". *Child and Adolescent Mental Health*, 28(2), 336-337.

About WiKIT

WiKIT supports companies to gather evidence of a proven impact on the educational experience of teachers, students, or both; and enhances the supportive framework upon which outstanding teaching and/or learning is constructed.

WiKIT has contrasted and compared various frameworks to develop the methodology behind our flagship service 'Get Evidence-ready', which supports companies to develop a research plan aligned with various evidence frameworks. WiKIT's researchers are trained on an evaluation methodology that takes into account all the different frameworks and pedagogical certifications, most of which are included in this report.

In creating the WiKIT methodology, we critically reviewed global frameworks and criteria for the pedagogical certifications and the priorities set forth by various international organisations. The results led to the creation of an EdTech Evidence support methodology, which we use when supporting companies to become "evidence-ready".

Recommended citation:

Kucirkova, N., Campbell, J., Lindroos Cermakova, A. (2023): *EdTech Impact Evaluation Frameworks: Summary 2023*, Report for WiKIT AS, DOI: 10.13140/RG.2.2.21563.59681

Note: Any errors or omissions found subsequent to printing will be corrected in the online version available at www.wikit.no

@ info@wikit.no

www.wikit.no