

EduEvidence 2025

Efficacy and Effectiveness Certifications: Criteria and Indicators*

Contents

Certification Procedure	2
Case 1: Companies have existing certifications	2
Efficacy	2
Effectiveness	3
Case 2: Companies do not have available certification but have supporting evidence	3
Difference between Efficacy and Effectiveness	4
Qualitative and quantitative studies	5
Indicators	5
Strength of efficacy in quantitative studies	5
Weight of evidence for quantitative studies	5
Weight of evidence for qualitative studies	7
Research support	8
College of Reviewers	8
Bibliography	9

*Updated on September 30th 2025, replacing version from February 2024

Certification Procedure

Companies seeking certification submit supporting documentation, referred to as "evidence," through ICEIE's online platform. ICEIE evaluates this evidence using a 3-point scale, with 0 indicating that the company does not meet certification criteria, and 3 representing the highest level of certification, Gold.

Table 1: Certifications levels and basic requirements

0 (No certification)	1 (Bronze)	2 (Silver)	3 (Gold)
Does not meet the requirements for certification or information provided is missing/irrelevant	Meets the requirements for the criteria for Bronze certification in the given vertical either through previous certification or relevant evidence	Meets the requirements for the criteria for Silver certification in the given vertical either through previous certification or relevant evidence	Meets the requirements for the criteria for Gold certification in the given vertical either through previous certification or relevant evidence

Case 1: Companies have existing certifications

Companies holding current certifications have the option to undergo a streamlined certification process with ICEIE. They submit their existing certifications for evaluation, and ICEIE assesses them against efficacy or effectiveness criteria, assigning them a placement on a scale accordingly.

Efficacy

Table 2: Certifications levels for ESSA, AERO and Queensland standards of evidence (QoE)*

Bronze	Silver	Gold
ESSA Tier IV AERO Level II QoE Low Digital Promise: Research-Based Design Certification ICEI Evidence-Ready completed report Evidence/Impact Audit equivalent to ESSA Tier IV	ESSA Tier III and II AERO Level II and III QoE Moderate	ESSA Tier II and I** AERO Level III and IV** QoE High

*NB: This list is not exhaustive and national, government-approved and officially recognised certifications are being added on an ongoing basis

**Exact level depends on the effect size obtained in the study

Effectiveness

Table 3: Certifications levels based on officially recognised certifications

Bronze	Silver	Gold
Pedagogical certifications with qualified judges or teachers, e.g.: <ul style="list-style-type: none"> • Edtech Impact: Education Alliance Finland and User experience • ISTE Seal of Alignment • Digital Promise: Practitioner-Informed Design: Product Certification • ICEI Pedagogy Ready completed report 	Correlational and observational studies in classrooms aligned with ESSA III and AERO Level II standards with no control or comparison groups OR Feasibility, Usability and Cost-effectiveness evaluation in alignment with ICEI standards	Officially recognised effectiveness ratings by leading expert groups World Bank/UNESCO “Smart Buy” that include cost-effectiveness ratings

*NB: This list is not exhaustive and national, government-approved and officially recognised certifications are being added on an ongoing basis

**Exact level depends on the effect size obtained in the study

Awards

The ICEIE raters take into account officially recognised awards judged by teachers or educational experts and encourage companies to submit evidence of awards such as those from UNICEF, Milken-Penn GSE, SXSW, see [FULL LIST HERE](#)

Case 2: Companies do not have available certification but have supporting evidence

Companies that do not have official certification for their evidence, but have conducted significant internal or external research, can seek validation from ICEIE. The companies need to submit the evidence in a package that aligns with the criteria in this document.

Companies have the option to submit a **self-assessment**, along with supplementary notes explaining their rationale for their certification level and the reasons for their application. While self-assessment does not replace the review by ICEIE's research partner, it can streamline the process and shorten the assessment duration. However, all companies are advised to seek support from one of ICEIE's research partners to present the evidence in a format that facilitates quick evaluation. It's important to note that ICEIE's role is not to validate or develop companies' research evidence but rather to evaluate **the weight of evidence and efficacy**

strength required in the evidence package submission, ultimately determining its eligibility for ICEIE certification.

To accurately assess studies, ICEIE's research partners rely on openly shared criteria and use a thorough peer review process, leading to a consensus for awarding a certificate based on detailed discussions. Qualified PhD-level learning scientists, specializing in either quantitative or qualitative studies, conduct assessments within ICEIE. The criteria that guide ICEIE's research partners when assessing companies' evidence bases, are based on a synthesis of frameworks and consolidated benchmarks (see References). Research partners make recommendations to ICEIE, which then issues the corresponding certificate.

Difference between Efficacy and Effectiveness

ICEIE acknowledges the varying definitions of efficacy and effectiveness across academic disciplines and international clearinghouses, as well as the significance of both teachers' and researchers' testing of educational technology tools in classroom settings. Drawing from Singal et al. (2014), *efficacy* refers to the performance of an intervention under ideal and controlled circumstances, while *effectiveness* pertains to its performance under real-world conditions. ICEIE extends this distinction to EdTech certifications, attributing *efficacy* to products tested under controlled conditions by researchers and *effectiveness* to those tested under less stringent but more ecologically valid conditions by teachers in real-world settings.

This simplification aims to respect the diverse perspectives on quality assurance and evaluations of educational technology that are present internationally (see Kucirkova, 2023b) and among various stakeholders engaged in EdTech evaluations.

Efficacy	Effectiveness
Does it work?	Could it work?
<p>How could it work?</p> <p>To decide whether a company's evidence portfolio qualifies more for an efficacy or effectiveness certification, ICEIE asks: was the product tested empirically with studies conducted with users? How was this testing performed?</p> <p>If YES -> Silver awards and above</p> <p>If Not -> Bronze awards</p>	
Quasi-experimental studies Experimental studies A/B testing	Observational studies Correlational studies Feasibility and Usability studies Cost-effectiveness considerations

Following this distinction, efficacy encompasses empirical studies conducted in classrooms and schools with research teams (these could be internal or external teams). For lowest levels of Bronze, the expectation is that the product was reviewed by qualified researchers to establish its connection to published studies and that the company has a clear research plan.

At higher levels, it relates to correlational studies and for top level to experimental studies- in the case of quantitative research.

Qualitative and quantitative studies

ICEIE prioritizes methodological diversity in recognizing best practices in evidence-building for educational technology (Kucirkova, 2023). This approach incorporates both qualitative and quantitative studies to assess efficacy or effectiveness.

ICEIE distinguishes between *demonstrated* efficacy, evaluated through quantitative studies, and *perceived* efficacy, assessed via qualitative studies. As for indicators of efficacy, for quantitative studies, ICEIE considers internal and external validity, as well as reliability of findings. For qualitative studies, the indicators of credibility, reflexivity, and member validation, are used, as proposed by Kucirkova, Brod and Gaab (2023) and Gough (2007). Note that qualitative and quantitative studies can appear in both Efficacy and Effectiveness categories.

Indicators

Strength of efficacy in quantitative studies

Criteria	Indicator
Efficacy strength overall	Between 20-40% efficacy (Bronze)
	Between 40-60% efficacy (Silver)
	Between 60-80% efficacy (Gold)

Weight of evidence for quantitative studies

NB: Validity and reliability are intertwined concepts, with specific indicators for each varying based on the unique characteristics of the study being conducted.	Bronze	Silver	Gold
Internal validity Internal validity refers to how accurately the observed results reflect the reality within the specific population being studied, ensuring they are not influenced by errors in the research methods.	Sample size on the margin of good statistical power	Sample size sufficient for good statistical power, with well-documented attrition rates	Sample size sufficient for high statistical power, low attrition rate OR attrition well-documented AND standardized treatment conditions with a control group

External validity External validity looks at whether the results of a study can apply to different situations. When studies use random sampling and the sample represents the whole population well, the findings can be applied to the entire population.	Study followed at least some randomization of participants	Appropriate participant randomization for the study design	The study used random or stratified sampling
Reliability Reliability refers to the consistency of results across different raters or observers and when e.g. the same test is administered to the same group of participants at different times.	Plans for replicating the study were documented	Attempts to replicate the study in other contexts were made	Study was Replicated in other contexts

Weight of evidence for qualitative studies

	Bronze	Silver	Gold
Credibility Credibility gauges the accuracy of the study's findings, partly based on the trustworthiness of the researchers and their methods.	Data were collected by at least two different methods and findings compared	Data and findings compared across at least two different methods, and two different analysts	Data/findings compared by at least two different analysts, with at least two different methods, and a representative sub-sample of participants
Member validation Member validation, also called participant or respondent validation, is a method to assess the credibility of findings by sharing data or results with participants to verify their accuracy and alignment with their experiences, commonly cited as one of several validation techniques.	A group of researchers from the same research team interpreted a selection of the data	A group of researchers interpreted the data together, with clear description of how consensus was reached and these were verified with some of the participants	Interpretation of findings was verified with independent researchers and all the participants
Reflexivity Reflexivity involves recognizing researchers' involvement in the research process, where prior experiences, assumptions, and beliefs shape the research. While elements of reflexivity apply to quantitative studies too, attention to theoretical saturation is key in qualitative studies.	Some awareness of the importance of researcher reflexivity noted/ Some components of the guiding theory documents	Mention of researcher reflexivity and evidence of use of reflexivity tools/ Most components of the theory are documented	Evidence of deep researcher reflexivity throughout the study, with supporting evidence/ Theory saturation through data

Research support

In the case when companies lack existing certification or supporting evidence and express a desire to address this situation, they are encouraged to take proactive steps to obtain supporting evidence to meet the necessary criteria. The companies may initiate actions to develop and present documentation that aligns with the certification requirements, demonstrating their commitment to achieving certification or evidencing their capabilities in a given domain. To support this process, ICEIE has a list of growing research partners, who can support interested companies against a fee. ICEIE is not involved in this process but can provide a list of research partners upon request.

College of Reviewers

ICEIE evaluators, organised as Members of the College of Reviewers, are recognized experts in the fields of EdTech efficacy and effectiveness and are committed to providing consistent, transparent and high-quality reviews based on the criteria agreed by ICEIE.

The College of Reviewers are selected and coordinated by the ICEIE's Academic Advisory Board.

Bibliography

Evidence for ESSA: Standards and Procedures, Version 2.1, November 2023, Evidence for ESSA, Available online from: www.evidencefoessa.org

Foster, D., McLemore, C., Olszewski, B., Chaudhry, A., Cooper, E., Forcier, L., & Luckin, R. (2023). EdTech Quality Frameworks and Standards Review. Department for Education, Published 14. December 2023, <https://www.gov.uk/government/publications/edtech-quality-characteristics-frameworks-and-standards-review>

Gough, D. (2007). Weight of Evidence: a framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education*, 22 (2), 213–228. <https://doi.org/10.1080/02671520701296189>

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>

Kucirkova, N. (2023a) Evidence in EdTech: A Global Demand with Local Conditions, *Global Edtech*, <https://global-edtech.com/evidence-in-edtech-a-global-demand-with-local-conditions/>

Kucirkova, N. (2023b) Why EdTech must prioritise research and innovation from the Global South, *World Economic Forum*, <https://www.weforum.org/agenda/2023/05/to-advance-edtech-evidence-we-must-prioritise-research-and-innovation-from-the-global-south/>

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), 35–42.

Kucirkova, N. I., Cermakova Lindroos, A. & Vackova, P. (2024) Consolidated Benchmark for Efficacy and Effectiveness Frameworks in EdTech. Stavanger: University of Stavanger. <https://ebooks.uis.no/index.php/USPS/catalog/book/270>

Singal, A. G., Higgins, P. D., & Waljee, A. K. (2014). A primer on effectiveness and efficacy trials. *Clinical and Translational Gastroenterology*, 5(1), e45.

Streiner, D. L. (2002). The 2 “Es” of research: Efficacy and effectiveness trials. *The Canadian Journal of Psychiatry*, 47(6), 552–556.

UNESCO, 2023. *Global Education Monitoring Report: Technology in Education: A tool on whose terms?*, UNESCO. <https://www.unesco.org/gem-report/en/technology>

The evaluation criteria and academic basis of the indicators and quality criteria in this document were prepared by WiKIT AS, January 2024 and were updated in February 2025 (second iteration) and in September 2025 (third iteration). All rights reserved.