

Evidence-informed Evaluatie van EdTech (3E).

Een Nederlands raamwerk
voor continue beoordeling van
de effectiviteit van EdTech

Evidence-informed Evaluatie van EdTech (3E).

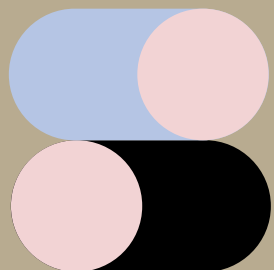
Een Nederlands raamwerk voor continue beoordeling van de effectiviteit van EdTech

Dit raamwerk is ontworpen voor Npuls door Dr. Manika Garg (De Haagse Hogeschool, projectleider en eindredacteur) en Dr. Theo Bakker (De Haagse Hogeschool, reviewer). We bedanken alle experts die we hebben geïnterviewd, evenals collega's van Npuls en SURF en de adviesraadsleden die het document hebben gelezen voor hun input. Het originele document is geschreven in het Engels.

Dit werk is een Nederlandse vertaling van Garg, M., & Bakker, T. C. (2025). Evidence-informed Evaluation of EdTech (3E): A Dutch Framework for Continuous Assessment of EdTech Effectiveness (pp. 1-15). Npuls. doi.org/10.5281/zenodo.15070789



De Creative Commons Naamsvermelding-GelijkDelen 4.0-internationaal licentie is van toepassing op deze publicatie. Als je dit werk gebruikt, vermeld dan de volgende referentie: Garg, M., Bakker, T.C. (2025). Evidence-informed Evaluatie van EdTech (3E): Een Nederlands raamwerk voor continue beoordeling van de effectiviteit van EdTech. (pp. 1-15). Npuls. doi.org/10.5281/zenodo.15524637



Inleiding

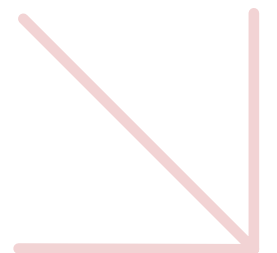
Onderwijstechnologie (EdTech) verwijst naar digitale hulpmiddelen die zijn ontworpen om het onderwijs te bevorderen door het verbeteren van onderwijs- en leerervaringen en het stroomlijnen van administratieve taken¹. Deze tools spelen meer en meer een centrale rol in het onderwijs en transformeren de manier waarop studenten leren, docenten lesgeven en instellingen middelen beheren².

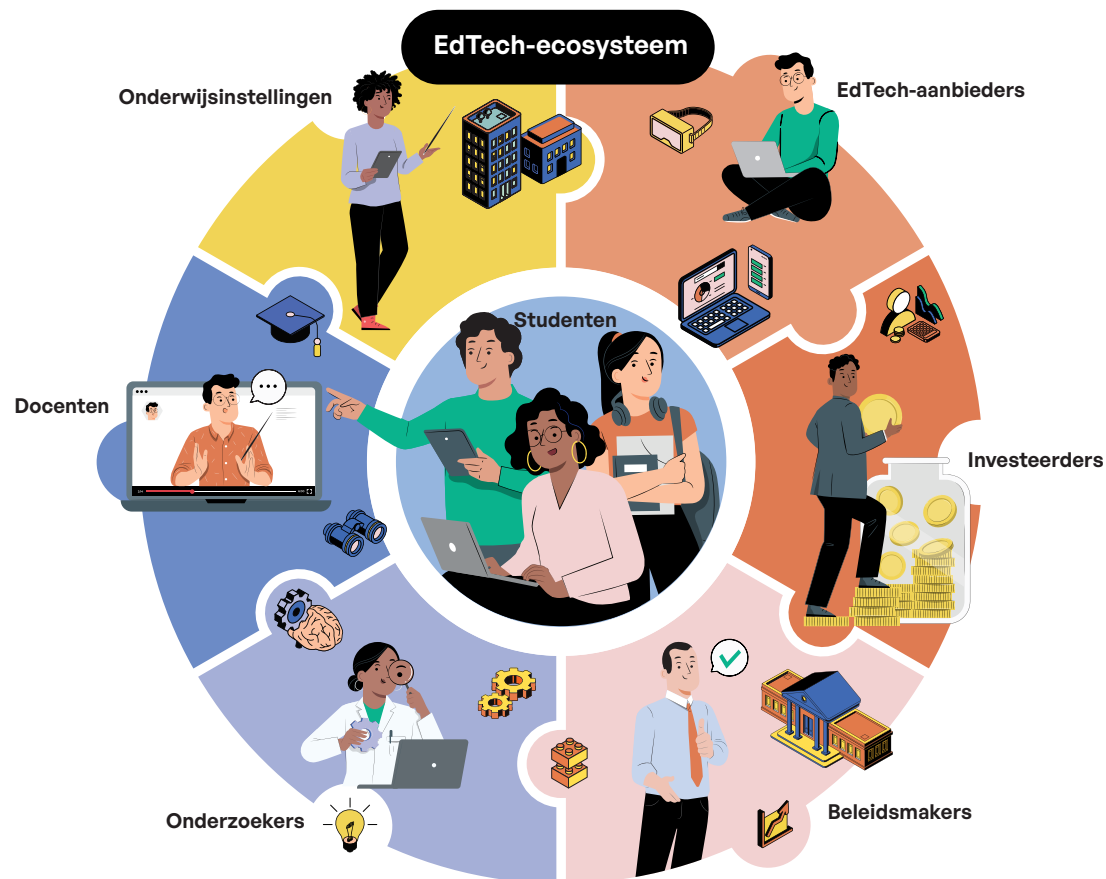
“We zouden zonder onderwijstechnologie uit de voeten kunnen, maar we zouden niet het beste onderwijs kunnen bieden.”

Silvester Draaijer,

Programmamanager bij de VU Onderwijswerkplaats,
Vrije Universiteit Amsterdam

Een EdTech ecosysteem omvat verschillende stakeholders: studenten, docenten, instellingen, EdTech-aanbieders, investeerders, beleidsmakers en onderzoekers³. Elke stakeholder speelt een cruciale rol: studenten en docenten maken gebruik van de technologie, instellingen stimuleren de adoptie, EdTech-aanbieders en investeerders stimuleren innovatie, beleidsmakers stellen regelgeving op en onderzoekers analyseren de behoefte, het ontwerp en de effectiviteit van tools om toekomstige ontwikkelingen vorm te kunnen geven.





Figuur 1. Het EdTech ecosysteem

Hoewel EdTech het potentieel heeft om het leren te verbeteren, hangt de effectiviteit ervan af van de juiste aansluiting op de onderwijscontext⁴. Wanneer EdTech in de verkeerde context wordt ingezet, kan dit geen enkele verbetering opleveren of zelfs een negatieve invloed hebben op onderwijsprocessen, doordat het leidt tot meer verwarring en minder betrokkenheid⁵. Het komt vaak voor dat instellingen EdTech aanschaffen op basis van een hype (omdat het populair is) of overtuiging (omdat anderen het gebruiken) in plaats van op gedegen evidence^a van hun effectiviteit^{6,7}. Omdat instellingen geen evidence vereisen, voelen EdTech-aanbieders niet de noodzaak om hun producten grondig te evalueren. Hierdoor ontstaat een vicieuze cirkel waarin tools worden gebruikt zonder duidelijk evidence van hun effectiviteit. Het gebrek aan evaluatie zorgt er ook voor dat er veel EdTech-tools van lage kwaliteit op de markt zijn, die

^a In deze uitgave gebruiken we de Engelse term 'evidence' in plaats van 'bewijs'.

weliswaar de schermtijd verhogen maar het leren niet verbeteren^{8,9}. Deze tools lijken misschien succesvol omdat studenten er meer tijd aan besteden, maar actief gebruik van technologie leidt niet altijd tot betekenisvol leren¹⁰.

De huidige situatie onderstreept de noodzaak van een op feiten gebaseerde evaluatie van EdTech-tools om hun effectiviteit te beoordelen en te bepalen in welke contexten ze het beste werken. Hoewel evaluatie op zichzelf inzicht kan geven in gebruiksvriendelijkheid en implementatie, zorgt evidence ervoor dat evaluaties gebaseerd zijn op meetbare impact in plaats van subjectieve indrukken of gebruikstatistieken.

Om in deze behoefte te voorzien stellen we een Evidence-informed Evaluatie van EdTech (3E) raamwerk voor, ontworpen voor de Nederlandse onderwijscontext. Het 3E-raamwerk is gebaseerd op uitgebreid literatuuronderzoek, wereldwijde best practices en interviews met experts. Vergelijkbare initiatieven in andere landen laten al veelbelovende resultaten zien in het verbeteren van besluitvorming over EdTech (zie meer in hoofdstuk 2).

Het Nederlandse 3E-raamwerk biedt een empirisch onderbouwde aanpak^b om de effectiviteit van EdTech te evalueren. Een belangrijk kenmerk van het raamwerk is de nadruk op continue evaluatie in plaats van statische, eenmalige toetsing. Dit zorgt ervoor dat EdTech-tools in de loop van de tijd worden verbeterd. Het raamwerk is een stap naar een effectievere inzet van technologie in het onderwijs, omdat het de besluitvorming op basis van evidence – bij het ontwikkelen, inkopen en verbeteren van EdTech – begrijpelijk en toepasbaar maakt.



Het hoofddoel van het Nederlandse 3E-raamwerk is niet om evaluatie voor alle tools verplicht te stellen, maar om een cultuur te bevorderen waarin evidence-informed keuzes zijn ingebed in alle EdTech-gerelateerde processen.

Het raamwerk dient als een praktische handleiding voor het genereren, interpreteren en gebruiken van evidence. Het biedt een toegankelijk en bruikbaar proces om evidence-informed keuzes te maken over het ontwikkelen, inkopen en verbeteren van EdTech.

^b Een gebrek aan evidence betekent niet dat een tool niet effectief is. Het kan gewoon betekenen dat de tool nog niet formeel is getest of zich nog in een vroeg ontwikkelingsstadium bevindt. Als er geen evidence beschikbaar is, moedigt het raamwerk aan om kritisch naar keuzes te kijken en – waar nodig – zelf evidence te verzamelen.

^c Bij het meten van effectiviteit kijken we naar de context waarin een tool het beste werkt. We erkennen dat de effectiviteit afhangt van wie de tool gebruikt, hoe deze wordt gebruikt en onder welke omstandigheden.

Het raamwerk is bedoeld om zinvolle discussies binnen de Nederlandse onderwijsgemeenschap mogelijk te maken over Evidence-informed Evaluatie EdTech en om beter om te gaan met de complexiteit van het gebruik van technologie in het onderwijs.

Voor EdTech-aanbieders kan het raamwerk een leidraad zijn voor het genereren van empirisch onderbouwde inzichten over hun tools die hen in staat stellen om hun tools te verfijnen. Voor instellingen biedt het raamwerk een leidraad voor het interpreteren en toepassen van evidence, wat hen kan helpen bij beslissingen over de inkoop van tools die aansluiten bij hun specifieke onderwijsbehoeften.

Leeswijzer

Het Nederlandse 3E-raamwerk is in de eerste plaats bedoeld voor onderwijsprofessionals (zoals docenten en informatiemanagers), EdTech-bedrijven (waaronder aanbieders en investeerders), beleidsmakers en onderzoekers. Het biedt praktische handvatten om de effectiviteit van EdTech te beoordelen en waardevolle inzichten om weloverwogen keuzes te maken over het gebruik van technologie in het onderwijs.

Dit document is onderverdeeld in vijf hoofdstukken. **Hoofdstuk 1** introduceert het begrip ‘evidence’ en legt uit waarom dit belangrijk is bij het evalueren van EdTech. **Hoofdstuk 2** verkent het huidige landschap van EdTech-evaluatie en vergelijkt wereldwijde praktijken met de Nederlandse situatie. **Hoofdstuk 3** presenteert het Nederlandse 3E-raamwerk. **Hoofdstuk 4** biedt een praktische routekaart voor onderwijsinstellingen en EdTech-aanbieders om het raamwerk effectief toe te passen. **Hoofdstuk 5** sluit af met de conclusies. Een extra bijlage biedt uitleg over het toepassen van de onderzoeksmethoden die in het raamwerk worden genoemd.

We illustreren de inzichten aan de hand van interviews met experts en voorbeelden uit de praktijk. Een overzicht van de geïnterviewde experts vind je aan het eind van dit document.

Heb je na het lezen van dit document suggesties? Laat het ons weten via edtech@npuls.nl of op de community pagina van de Pilotheb EdTech. Dan kunnen we je input meenemen in een volgende versie van deze uitgave.

1. Het begrijpen van Evidence in EdTech-evaluatie

Wat is evidence?

Evidence kan worden gedefinieerd als een feit, teken of object dat iemand doet geloven dat iets waar is¹. In de context van EdTech-evaluatie verwijst het naar data, onderzoek en gevalideerde inzichten die aantonen dat een specifieke tool aan de beoogde doelstellingen voldoet binnen een bepaalde context.



Verschillende soorten evidence leggen in verschillende mate gewicht in de schaal, afhankelijk van de nauwkeurigheid waarmee ze zijn verzameld en de betrouwbaarheid van bevindingen. Het kan afkomstig zijn van academisch onderzoek of intern uitgevoerde evaluaties. Beide vormen kunnen waardevol en belangrijk zijn om te bepalen of een EdTech-tool effectief is.

Waarom hebben we evidence nodig?

Evidence is essentieel voor het evalueren van de effectiviteit van EdTech, omdat het ervoor zorgt dat evaluaties gebaseerd zijn op meetbare impact in plaats van indrukken of populariteit². Het levert inzichten op die met data onderbouwd zijn en laat zien of een tool daadwerkelijk bijdraagt aan betere leerresultaten. Evidence geeft geen garanties voor succes in elke situatie, maar biedt wel waardevolle inzichten op basis van bestaand onderzoek en praktijkervaringen.

Een cultuur van evidence-informed evaluatie ondersteunt continue verbetering. Het helpt om veranderende onderwijsbehoeften te herkennen en ervoor te zorgen dat EdTech-tools zich daaraan aanpassen. Omdat middelen vaak beperkt zijn, moeten instellingen en aanbieders prioriteiten stellen. Evidence helpt om te bepalen welke tools écht bijdragen aan het onderwijs en zorgt ervoor dat investeringen leiden tot betekenisvolle opbrengsten in het onderwijs. Het stelt onderwijsinstellingen in staat om onderbouwde keuzes te maken over het selecteren, bijstellen of uitfasen van EdTech-tools.

Daarnaast bevordert evidence ook transparantie en verantwoordelijkheid bij EdTech-aanbieders. Het stimuleert onderbouwde claims en ontmoedigt misleidende marketing. Beleidsmakers en subsidieverstrekkers vertrouwen vaak op evidence om regels op te stellen en middelen toe te wijzen aan tools die meetbare verbeteringen laten zien. Door evidence te verankeren in besluitvorming wordt de inzet van technologie in het onderwijs waarschijnlijk effectiever, duurzamer en impactvoller.



Evidence als kompas en kaart

Bij het evalueren van EdTech werkt evidence als een combinatie van een kompas en een kaart. Het kompas geeft richting – het helpt stakeholders gefocust te blijven op hun doelen, zoals het verbeteren van leerresultaten of het aanpakken van specifieke onderwijsvraagstukken. De kaart geeft een duidelijk beeld van het landschap – het laat zien welke tools waar hebben gewerkt, en onder welke omstandigheden.



Net zoals een kompas en een kaart onmisbaar zijn om een bestemming te bereiken, vervult evidence beide functies: het wijst naar de gewenste resultaten en brengt het landschap in kaart, zodat je bewust kunt kiezen voor EdTech-tools en deze kunt ontwikkelen op basis van werkelijke onderwijsbehoeften.

Voor instellingen betekent deze gecombineerde gids dat zij tools kunnen selecteren die echt aansluiten bij wat er nodig is. Voor aanbieders helpt het bij het ontwerpen en bijstellen van tools op basis van bewezen effect. In die zin is evidence niet alleen richting of informatie – het is allebei: weten waar je naartoe wilt én begrijpen hoe je daar komt.

2. Evaluatie van EdTech: internationale praktijken versus Nederland

Veel landen wereldwijd erkennen het belang van Evidence-informed Evaluatie van EdTech en hebben hiervoor verschillende certificeringen, raamwerken en standaarden ingevoerd^{13,14}.

“Als de vraag naar evidence toeneemt door nationale standaarden, zullen meer EdTech-bedrijven investeren in evaluaties.”

Michael Forshaw, CEO van EdTech Impact

Certificeringen in het onderwijs zoals Digital Promise in de Verenigde Staten (VS), en de International Society for Technology in Education normen (ISTE) bieden gestructureerde rubrics waarmee docenten EdTech kunnen evalueren. Daarnaast hebben verschillende landen nationale kaders opgezet voor het evalueren van EdTech, zoals de Every Student Succeeds Act (ESSA) in de VS en de Australian Education Research Organisation (AERO) in Australië. Internationale organisaties – waaronder de Wereldbank, de Organisatie van de Verenigde Naties voor Onderwijs, Wetenschap en Cultuur (UNESCO), het Internationale Kindernoodfonds van de Verenigde Naties (UNICEF) en de Aziatische Ontwikkelingsbank (ADB) – hebben ook wereldwijde normen opgesteld voor de evaluatie van EdTech. In veel gevallen ontwikkelen investeerders en financieringsinstanties evaluatiekaders, zoals International Certification Of Evidence Of Impact In Education (ICEIE), om de impact van EdTech-instrumenten te meten voordat ze financiële beslissingen nemen¹⁵.

Hoewel grootschalige, longitudinale studies naar evaluatiekaders voor EdTech zich nog aan het ontwikkelen zijn, wijst de groeiende verzameling praktijkgerichte en beleidsgestuurde voorbeelden erop dat het hebben van een raamwerk leidt tot betere keuzes en de algehele kwaliteit van EdTech in de praktijk verhoogt.

Kaders zoals de ESSA Tiers of Evidence in de VS en het EdTech Impact Quality Framework in het Verenigd Koninkrijk hebben de inkooppraktijken al beïnvloed. Het What Works Clearinghouse van het Amerikaanse Ministerie van Onderwijs moedigt het gebruik van gevalideerde interventies aan. Deze kaders hebben de verwachtingen in de markt veranderd en aanbieders aangemoedigd om onderzoek te doen en hun producten te verbeteren op basis van gebruikers-

feedback en aangetoonde impact. Het 5Es Framework (EduEvidence) heeft al meerdere internationale aanbieders ertoe aangezet om te starten met formele evaluatieprocessen. Daarnaast hebben deze kaders bijgedragen aan het creëren van een gemeenschappelijk taal tussen stakeholders. Zo ontdekt Educate Ventures (VK) dat het gebruik van een gemeenschappelijk kader instellingen en startups hielp sneller hun doelen op elkaar af te stemmen.



In tegenstelling tot andere landen kent Nederland geen gestandaardiseerd raamwerk voor de evaluatie van EdTech. Instellingen en docenten hebben de autonomie om EdTech-tools te kiezen op basis van hun eigen behoeften, waarbij vaak factoren zoals functionaliteit, kosten, privacy en compliance zwaarder wegen dan evidence van effectiviteit. Daardoor blijven gesprekken over evidence tussen instellingen en aanbieders beperkt.

“Beslissingen over inkoop aan Nederlandse universiteiten missen een gestandaardiseerde aanpak en zijn vaak gebaseerd op functionaliteit, kosten, privacy en compliance in plaats van bewezen effectiviteit.”

Bob Mooijenkind, Onderwijsadviseur, De Haagse Hogeschool

Hoewel sommige EdTech-bedrijven in Nederland het belang inzien van het evalueren van hun producten, stuiten zij vaak op uitdagingen in de samenwerking met onderzoekers en onderwijsinstellingen bij gebrek aan een gestructureerd ecosysteem dat partnerschappen tussen bedrijven en onderzoekers stimuleert¹⁶.

“Startups en kleine EdTech-aanbieders ondervinden aanzienlijke belemmeringen bij het uitvoeren van evaluaties door beperkte middelen en een gebrek aan goede testomgevingen.”

Erwin van Vliet, Universitair hoofddocent en opleidingsdirecteur bachelor Psychobiologie, Universiteit van Amsterdam

Er is duidelijk behoefte aan een gemeenschappelijke standaard of ecosysteem die evidence-informed besluitvorming in EdTech in Nederland ondersteunen. Het Nederlandse 3E-raamwerk is een stap in die richting. Het bouwt voort op internationale best practices en speelt in op de specifieke behoeften van de Nederlandse onderwijscontext.

3. Het Nederlandse 3E-raamwerk

Het Nederlandse 3E-raamwerk is een praktische gids die is ontworpen om de effectiviteit van EdTech-tools te evalueren op basis van de kwaliteit van de evidence die impact van deze tools op het onderwijs ondersteunt. Wanneer we spreken over het evalueren van tools, bedoelen we de evaluatie van één of meerdere specifieke functionaliteiten binnen de tool. Een tool kan verschillende functionaliteiten bieden en elke functionaliteit kan afzonderlijke overweging vereisen bij de evaluatie. Dit betekent ook dat een tool tegelijkertijd op verschillende functionaliteiten kan worden geëvalueerd.

Het Nederlandse 3E-raamwerk deelt evidence in op drie niveaus: Brons, Zilver en Goud^d

Deze niveaus geven aan hoeveel vertrouwen we kunnen hebben in de effectiviteit van een EdTech-tool. Brons-niveau tools hebben enige ondersteunende evidence maar vereisen verder onderzoek. Zilver-niveau tools tonen sterker evidence van impact. En Goud-niveau tools voldoen aan de hoogste standaarden van grondige studies die succes aantonen in meerdere contexten. Elk niveau bouwt voort op het vorige, waardoor tools kunnen doorgroeien naarmate er meer evidence beschikbaar komt.

Elk niveau van evidence vertegenwoordigt een verschillende mate van vertrouwen in het vermogen van een tool om de beoogde onderwijsresultaten te bereiken. Hoe sterker de evidence die een tool ondersteunt, hoe groter de zekerheid dat deze daadwerkelijk leidt tot betekenisvolle verbeteringen in het onderwijs.

Om geclassificeerd te worden op een specifiek niveau van evidence, moet een tool voldoen aan de onderzoeksstandaarden die voor dat niveau zijn gedefinieerd. Waar classificatie op Brons-niveau mogelijk gebaseerd is op niet-empirische bronnen zoals logische modellen of expertadvies, vereisen de niveaus Zilver en Goud ten minste één goed uitgevoerde studie die

^d De niveaus van evidence in het Nederlandse 3E-raamwerk zijn geïnspireerd op het 5Es Framework, geïntroduceerd door professor Natalia Kucirkova in 2023.

een positieve impact aantoon. Met 'positieve impact' wordt hier een meetbare verbetering bedoeld in het beoogde resultaat van de tool – dit kunnen academische prestaties, betrokkenheid of andere onderwijsdoelen zijn, afhankelijk van de context en het doel van de tool.

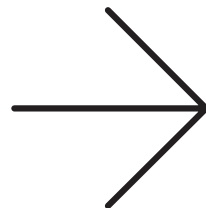
▶ Een belangrijk kenmerk van dit raamwerk is dat het uitsluitend positieve evidence in overweging neemt.

In plaats van zich te richten op tools waarvoor geen onderzoek beschikbaar is of die mogelijk een negatief effect hebben, legt het de nadruk op tools die getest zijn en positieve resultaten hebben laten zien. Maar dit betekent niet dat tools zonder evidence worden afgewezen; ze hebben simpelweg aanvullend onderzoek nodig voordat ze met vertrouwen kunnen worden aanbevolen. Het doel is om de kwaliteit van EdTech-tools te bevorderen door evidence-informed processen aan te moedigen in plaats van tools alleen maar te valideren. Het idee om uitsluitend positieve evidence bij evaluatie te gebruiken wordt sterk aanbevolen door EduEvidence en de bijbehorende vijf rapporten¹⁷.

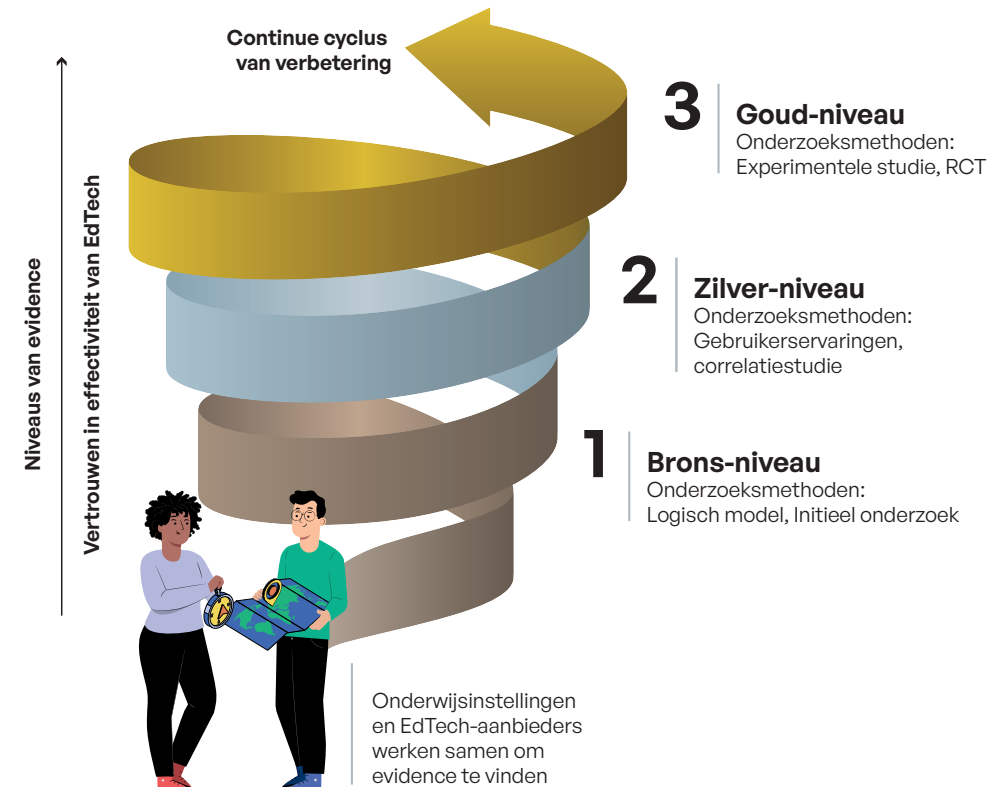
Een ander belangrijk kenmerk van het raamwerk is dat het een cyclus van continue verbetering doorloopt. Dit betekent dat zelfs wanneer een tool het hoogste niveau van evidence (Goud) bereikt, het evaluatieproces niet stopt. Er wordt voortdurend nieuwe evidence verzameld, zodat bestaande tools kunnen worden aangepast en verbeterd op basis van nieuw onderzoek, veranderende onderwijsbehoeften en technologische ontwikkelingen.

▶ De continue verbetercyclus zorgt ervoor dat tools relevant, effectief en afgestemd op de praktijk in het klaslokaal blijven.

In het volgende hoofdstuk worden de drie niveaus van evidence en de bijbehorende onderzoeksmethoden nader toegelicht. In de bijlage vind je meer informatie over de onderzoeksmethoden.



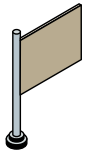
Evidence-informed Evaluatie van Edtech (3E) raamwerk



Figuur 2. Het Nederlandse 3E-raamwerk

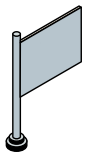
Niveaus van evidence

Het **brons-niveau** vertegenwoordigt de basisstandaard van evidence en is gebaseerd op niet-empirisch onderzoek zonder meetbare data. De bevindingen zijn afgeleid van theoretische modellen in plaats van toetsing in de praktijk. Ze geven aan of de tool is gebaseerd op solide onderwijstheorieën en goed gerepliceerde onderzoeksstudies. Onderzoeksmethoden op dit niveau omvatten het opstellen van een logisch model, het bouwen van een theory of change, literatuuronderzoek en consultaties met experts om de effectiviteit van een EdTech-tool te veronderstellen.



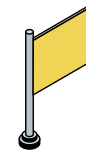
Voorbeeld van onderzoek op het Brons-niveau: Een startup ontwikkelt een AI-gestuurde lees-assistent om beginnende lezers te helpen hun leesvaardigheid te verbeteren. In dit stadium is de tool nog niet formeel geëvalueerd, maar wordt wel ondersteund door sterke onderwijstheorieën en goed gerepliceerd onderzoek naar AI-gestuurd taalonderwijs. De ontwikkelaars stellen een logisch model op ze uiteenzetten hoe de tool het begrijpend lezen zou kunnen verbeteren, en raadplegen geterdheidsexperts en docenten voor kwalitatieve feedback. Deze tool vereist nog verder onderzoek en tests voordat die met vertrouwen kan worden aanbevolen, maar heeft potentie op basis van de theoretische onderbouwing en raadpleging van experts.

Het **zilver-niveau** omvat experimentele studies zonder een controlegroep of vergelijkingsgroep, wat betekent dat causaliteit niet definitief kan worden vastgesteld. Onderzoeksmethoden op dit niveau zijn onder andere correlatieve studies, gebruikersonderzoek, analyses van contributie en kleinschalige casestudies. Ook al kunnen bevindingen wijzen op een positieve samenhang tussen het gebruik van EdTech en relevante uitkomsten, missen ze de wetenschappelijke degelijkheid die nodig is om causaliteit aan te tonen. Resultaten, missen ze de nauwkeurigheid die nodig is om een causaal verband te bevestigen.



Voorbeeld van onderzoek op het Zilver-niveau: Een universiteit integreert een adaptieve leermodule om studiemateriaal te personaliseren voor studenten. De universiteit voert een correlatieve studie uit waarbij het gebruik van de tool en de cijfers van studenten gedurende een semester worden geanalyseerd. De resultaten laten zien dat studenten die de tool vaker gebruiken, beter presteren. Omdat de studie echter geen controlegroep bevat, kan geen causaliteit worden vastgesteld. Hoewel de tool veelbelovende verbanden laat zien met verbeterd leren, zijn meer gecontroleerde studies nodig om het directe effect op leernutkomsten te bevestigen.

Het **Goud-niveau** vertegenwoordigt de hoogste standaard van evidence. Dit niveau vereist experimentele studies met een controle- of vergelijkingsgroep om sterke causale verbanden vast te stellen. Onderzoeksmethoden zijn onder andere gerandomiseerde gecontroleerde onderzoeken (Randomized Control Trials: RCT's), quasi-experimentele studies en case-control studies. Dit niveau garandeert dat de bevindingen statistisch significant zijn en de effectiviteit worden geëvalueerd binnen verschillende subgroepen, waarbij eventuele verschillen worden verklaard.



Voorbeeld van onderzoek op het Goud-niveau: Een Ministerie van Onderwijs test een AI-gestuurd bijlessysteem op 50 scholen. Studenten worden willekeurig toegewezen aan een groep die de AI-bijles gebruikt of aan een groep die traditioneel onderwijs krijgt. Na zes maanden blijkt uit de toetsresultaten dat de AI-groep een statistisch significante verbetering laat zien in probleemoplossende vaardigheden. Omdat dit een RCT is, levert de studie sterke causale evidence dat de AI-bijles daadwerkelijk het leren van studenten verbetert, maar vervolgonderzoek is nodig om te waarborgen dat de impact consistent blijft in verschillende onderwijscontexten.

De volgende tabel is een evidence bibliotheek die samenvat welk type onderzoek vereist is voor elk niveau van evidence, hoe deze evidence kan worden gegenereerd en wat het geloofwaardig maakt.

Tabel 1. Evidence bibliotheek

Niveau van evidence	Onderzoeksbenadering	Methode	Beschrijving	Factoren die de geloofwaardigheid vergroten
Goud	Experimentele studies met controle- of vergelijkingsgroepen, waarbij causaliteit wordt vastgesteld	Randomized Controlled Trials (RCT)	• Maakt gebruik van willekeurige toewijzing van deelnemers aan groepen (controle en vergelijking) • Kwantitatieve en kwalitatieve analyse • Gecontroleerde omgeving • Hoge interne validiteit • Sterke causale inferentie	Grote steekproeven, consistentie over verschillende locaties, replicatie
		Experimentele studie	• Vergelijkt vooraf bepaalde of zelfgekozen groepen, zonder willekeurige toewijzing. • Methoden omvatten quasi-experimentele studies, case-control studies • Kwantitatieve en kwalitatieve analyse • Praktijksituatie	Goed gematchte groepen, replicatie

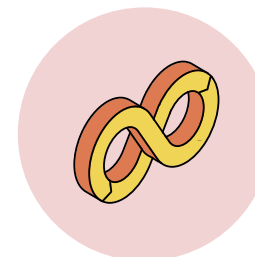
Niveau van evidence	Onderzoeks-benadering	Methode	Beschrijving	Factoren die de geloofwaardigheid vergroten
Zilver	Experimentele studies zonder controle- of vergelijkings-groepen	Correlatie-studie	- Onderzoekt relaties tussen variabelen maar stelt geen causaliteit vast - Methoden omvatten kleinschalige casestudies, observatieonderzoek, etnografie, contributie-analyse - Kwantitatieve maatstaven (bijv. betrokkenheid, retentiepercentages) - Kwalitatieve analyse	Sterke statistische analyse, grote steekproeven, meerdere data-bronnen, recent onderzoek uitgevoerd door een onafhankelijke partij (niet de aanbieder van de tool) in een praktijk-situatie in diverse omgevingen
		Gebruikers-onderzoek	- Beoordeelt tevredenheid, bruikbaarheid en percepties van de tool via gebruikers-feedback - Methoden omvatten enquêtes, interviews, persoonlijke anekdotes, feedback uit appstores, participatief ontwerp, aanbevelingen van klanten	Grote en diverse gebruikersgroep, feedback van meerdere belanghebbenden, hoge responspercentages
Niveau van evidence	Onderzoeks-benadering	Methode	Beschrijving	Factoren die de geloofwaardigheid vergroten
Brons	Niet-empirisch onderzoek zonder meetbare evidence	Initieel onderzoek	- Baseert zich op secundaire bronnen om een onderbouwing te formuleren - Methoden omvatten literatuurstudies, raadplegingen met domein-experts, belanghebbenden en gelijkgestemde organisaties	Peer-reviewed en recente studies, ondersteund door meerdere bronnen
		Logisch model	- Beschrijft input, activiteiten en output die verondersteld worden te leiden tot impact - Methoden omvatten het opstellen van een Theory of Change en conceptuele kaders	Goed onderbouwde redenering, stap voor stap verantwoording
Geen badge	Geen evidence ter ondersteuning van effectiviteit			

Continue verbetercyclus

Een belangrijk kenmerk van het Nederlandse 3E-raamwerk is de continue verbetercyclus in tegenstelling tot statische, eenmalige certificeringsmodellen in andere kaders. Zelfs wanneer een EdTech-tool het hoogste niveau van evidence bereikt, stopt evaluatie niet. Nieuwe evidence wordt regelmatig verzameld en geactualiseerd om te waarborgen dat de tool relevant blijft naarmate onderwijsbehoeften, technologieën en gebruikerscontexten zich verder ontwikkelen. Dit voorkomt dat verouderde oplossingen onbeperkt als effectief worden beschouwd en stimuleert innovatie en aanpassingsvermogen.

“Je moet niet pas aan het eind evalueren, maar tijdens de hele ontwikkelcyclus. Dat voorkomt mislukkingen in een laat stadium en zorgt dat de tool blijft aansluiten bij de werkelijke behoeften.”

Ewoud de Kok, Oprichter en CEO, FeedbackFruits



Er zijn verschillende manieren om continue verbetering te bevorderen — bijvoorbeeld door een tijdslimiet of vervaldatum te koppelen aan de bruikbaarheid van specifieke evidence. Net zoals beleid of technologieën na verloop van tijd geüpdatet moeten worden, geldt dat ook voor de data en het onderzoek waarmee EdTech-tools worden geëvalueerd. Als een tool bijvoorbeeld drie jaar geleden is geëvalueerd in een andere context of met verouderde technologie, is die evidence mogelijk niet meer representatief voor de prestaties van de tool vandaag de dag. Door een tijdslimiet of vervaldatum vast te leggen, kunnen we oude evidence bewust uifaseren en instellingen en EdTech-aanbieders stimuleren om de tool regelmatig opnieuw te beoordelen.

Evidence portfolio

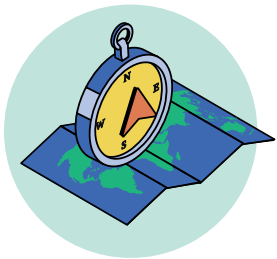
Het evidence portfolio is een belangrijk resultaat van het Nederlandse 3E-raamwerk. Het fungeert als een levend document waarin instellingen en EdTech-aanbieders systematisch onderzoeksbevindingen over de effectiviteit van een tool verzamelen, ordenen en actualiseren.

“Door deze richtlijnen in je EdTech-startup te verwerken, kun je gaandeweg de documentatie over de impact van je technologie opbouwen. Een model dat ik erg waardeer is dat van een levend evidence portfolio.”

Natalia Kucirkova,

Medeoprichter en directeur, International Centre for EdTech Impact
(geciteerd in een artikel¹⁸ gepubliceerd in Startups Magazine)

Het evidence portfolio weerspiegelt de toepassing van het raamwerk door vast te leggen hoe een EdTech-tool zich ontwikkelt over de verschillende evidence niveaus (Brons, Zilver en Goud). Het biedt een volledig overzicht van de effectiviteit van een tool, gebaseerd op praktijk-data, onderzoeksstudies en feedback van gebruikers. Omdat het Nederlandse 3E-raamwerk werkt met een continue verbetercyclus, is het evidence portfolio geen statisch document — het wordt regelmatig geüpdatet wanneer er nieuw onderzoek beschikbaar komt.



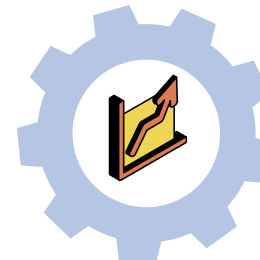
Door een portfolio bij te houden, beschikken onderwijsinstellingen over een betrouwbare bron voor datagedreven inkoopbeslissingen, terwijl EdTech-aanbieders hiermee hun impact kunnen aantonen, hun producten verfijnen en hun geloofwaardigheid bewijzen.

4. Stappen voor het gebruik van het raamwerk

Dit hoofdstuk licht toe hoe onderwijsinstellingen en EdTech-aanbieders het raamwerk in de praktijk kunnen toepassen. Het proces volgt drie hoofdfasen: **Evidence verzamelen – toepassen – actualiseren.**

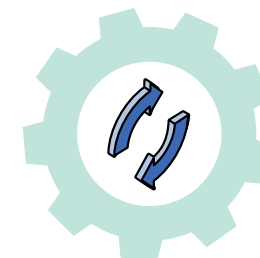


1. Evidence verzamelen – Evidence wordt gegenereerd via publiek-private samenwerkingen, waarin instellingen en EdTech-aanbieders in co-creatie samenwerken om onderzoek te doen en betrouwbare evidence te produceren over de effectiviteit van EdTech.



2. Evidence toepassen – Zodra de evidence beschikbaar is, wordt deze ingezet bij besluitvorming.

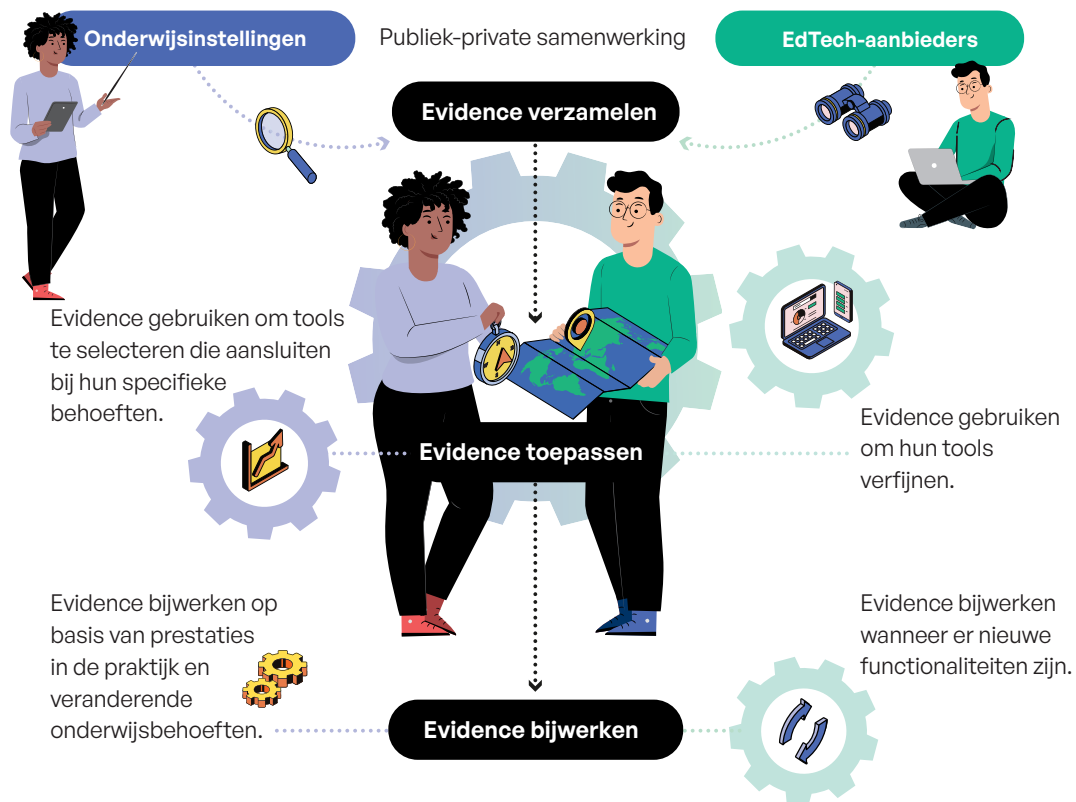
- **Instellingen** gebruiken de evidence om tools te selecteren die aansluiten bij hun specifieke onderwijsbehoeften.
- **EdTech-aanbieders** gebruiken de evidence om verbeterpunten te identificeren en hun tools gericht door te ontwikkelen om de impact daarvan te vergroten.



3. Evidence actualiseren – Het verzamelen van evidence is een continu proces dat meebeweegt met veranderende behoeften.

- **Instellingen** actualiseren evidence op basis van prestaties in de praktijk en hun veranderende prioriteiten in het onderwijs.
- **EdTech-aanbieders** actualiseren evidence telkens wanneer zij verbeteringen doorvoeren of nieuwe functionaliteiten introduceren, zodat hun tools relevant en effectief blijven.

Stappen om het Nederlandse 3E-raamwerk te gebruiken



Figuur 3. Stappen om het Nederlandse 3E-raamwerk te gebruiken

5. Conclusie

Het Nederlandse 3E-raamwerk stimuleert een cultuur waarin beslissingen over EdTech worden genomen op basis van evidence in plaats van aannames. Het creëert een natuurlijke cyclus: betere evidence leidt tot betere tools, en betere tools genereren op hun beurt weer nieuwe evidence. Dit is gunstig voor zowel instellingen, die zo beter passende tools kunnen selecteren, – als voor aanbieders, die waardevolle inzichten krijgen om hun producten te verbeteren.

Het raamwerk helpt om de eerste stap richting evidence te zetten. De hiërarchische opbouw is niet bedoeld om tools in te delen op hun waarde, maar biedt een overzichtelijk, stapsgewijs proces om naar effectiviteit toe te werken en effectief te blijven. Door aanbieders en instellingen langs een duidelijke route te begeleiden, verlaagt het de drempel om met evidence te werken volgens een toegankelijker en beter uitvoerbaar proces.

Toch blijven er belangrijke vragen om samen te beantwoorden: Is dit raamwerk bruikbaar in Nederland? Zal het een formele vereiste worden of een best practice? Wat is ervoor nodig om het te laten werken – qua governance, onderzoekscapaciteit en draagvlak binnen instellingen? De antwoorden hangen af van de manier waarop instellingen en de EdTech-sector met het raamwerk aan de slag gaan.

Het Nederlandse 3E-raamwerk legt een stevige basis voor een diepere samenwerking tussen wetenschap, onderwijsinstellingen en de EdTech-sector. Wanneer alle betrokkenen samen evidence genereren, interpreteren en toepassen, verschuift de inzet van EdTech van *trial-and-error* naar een strategisch en duurzaam proces. Uiteindelijk maakt dit de weg vrij voor een effectieve en toekomstbestendige integratie van onderwijstechnologie.

► Het Nederlandse 3E-raamwerk is je kompas én routekaart in het complexe landschap van EdTech: het wijst de weg naar effectieve EdTech en toont hoe je daar kunt komen – via gedeelde, evidence-informed praktijken.

Bijlage

Deze bijlage legt de onderzoeksmethoden uit die bij elk niveau van evidence van het raamwerk horen, met voorbeelden van hoe deze in de praktijk kunnen worden toegepast.

1. Logisch model

Een logisch model laat zien hoe een interventie naar verwachting tot bepaalde uitkomsten leidt. Het beschrijft de belangrijkste onderdelen, zoals input, activiteiten, output en beoogde impact. Dit helpt eindgebruikers om de **theory of change** begrijpen achter een EdTech-tool. Om een sterk logisch model te ontwikkelen, beantwoord je de volgende vragen:

- Wat doet het product?
- Welke impact wil het bereiken?
- Hoe wordt die impact gerealiseerd?

Het opstellen van een **theory of action** helpt om het denkproces te verduidelijken en oorzaak-gevolgrelaties inzichtelijk te maken. De volgorde is: *Als (een gebruikersactie plaatsvindt) ... Dan (verandert er iets) ... Zodat (een observeerbare impact optreedt).*

2. Initieel onderzoek

Initieel onderzoek helpt een basis te leggen voor evaluatie door gebruik te maken van bestaande kennis en de inzichten van experts. Belangrijke benaderingen zijn:

- **Literatuuronderzoek:** Gebruik maken van academisch onderzoek, rapporten en systematische reviews om bias te vermijden en betrouwbaarheid te waarborgen.
- **Raadplegen van experts:** Experts betrekken om onderzoeksvragen aan te scherpen, methodes te toetsen en af te stemmen op best practices.
- **Samenwerken met organisaties:** Aansluiting zoeken bij gelijkgestemde instellingen, beroepsverenigingen en onderzoeksnetwerken.

3. Gebruikerservaring

Gebruikerservaring verzamelt data over hoe gebruikers omgaan met EdTech-tools en beoordeelt gebruiksvriendelijkheid en tevredenheid. Het beoordeelt de bruikbaarheid en tevredenheid via verschillende methoden, en zorgt ervoor dat de tools effectief en gebruiksvriendelijk zijn. Belangrijke benaderingen zijn:

- **Verzamelen van gebruikersfeedback** via persoonlijke ervaringen, docentenbeoordelingen, studentenpercepties, aanbevelingen van klanten, grootschalige enquêtes en appstore-reviews.
- **Participatief ontwerp en co-design**, waarbij gebruikers actief worden betrokken bij het ontwikkelen en verfijnen van het product.

4. Correlationeel onderzoek

Correlationele studies onderzoeken de relatie tussen twee of meer variabelen, maar tonen geen oorzakelijk verband aan. Ze helpen om betekenisvolle trends en verbanden te vast te stellen die als basis kunnen dienen voor verdere experimentele validatie van EdTech-tools. Hierbij worden kwantitatieve gegevens gebruikt, zoals analyses van studentenbetrokkenheid en retentie (bijvoorbeeld het aantal actieve gebruikers per maand, week of dag). Kwalitatieve analyse biedt daarnaast diepgaand inzicht via gebruikersfeedback, gedragsobservaties en contextuele gegevens. Toegepaste methoden zijn onder andere:

- **Observationele studies** – Het monitoren van gebruikersinteracties in realistische situaties om patronen in gebruik en betrokkenheid te ontdekken.
- **Contributie analyse** – Het onderzoeken van hoe specifieke functies of interventies bijdragen aan gebruikersgedrag of leeruitkomsten.
- **Etnografie** – Participerend onderzoek waarbij direct contact met gebruikers in hun leer-omgeving centraal staat, om inzicht te krijgen in de context van EdTech-gebruik.

5. Experimenteel onderzoek

Experimentele studies worden uitgevoerd in praktijksituaties, waarbij vooraf bepaalde groepen met elkaar worden vergeleken om de impact van een interventie te meten. Hoewel er vaak geen sprake is van willekeurige toewijzing, bieden deze studies gestructureerde omstandigheden om oorzaak-gevolgrelaties te onderzoeken. Zowel kwantitatieve als kwalitatieve analyses worden ingezet om de uitkomsten te evalueren. Methodes zijn onder andere:

- **Quasi-experimentele studies** – Onderzoek waarbij geen willekeurige toewijzing mogelijk is, maar statistische technieken worden gebruikt om verschillen tussen groepen te corrigeren.
- **Case-control studies** – Retrospectief onderzoek waarin groepen met en zonder een specifieke uitkomst worden vergeleken om bepalende factoren te achterhalen.

6. Gerandomiseerde gecontroleerde testen (RCT's)

RCT's zijn experimentele studies waarbij deelnemers willekeurig worden toegewezen aan een interventie- of controlegroep, om het causale effect van een interventie te meten. Deze methode minimaliseert bias en zorgt voor een hoge interne validiteit bij het beoordelen van de effectiviteit van EdTech. Vormen van RCT's zijn onder andere:

- **Stapsgewijze wig designs** – Interventies worden gefaseerd ingevoerd bij verschillende groepen.
- **Micro-RCT's** – Kleine proeven waarbij een interventie op beperkte schaal wordt getest voordat bredere implementatie plaatsvindt.

Dankwoord

Wij willen graag onze dank betuigen aan de leden van onze adviesraad: Dr. Natalia Kucirkova (medeoprichter en directeur van het International Centre for EdTech Impact), Dr. Ludo Juurlink, (directeur van het Leiden Learning and Innovation Centre, Universiteit Leiden) en Dr. Kees Mastenbroek (niet-uitvoerend directeur van What Worked Education) voor hun waardevolle begeleiding en ondersteuning.

Daarnaast willen we de volgende experts hartelijk bedanken voor het delen van hun inzichten tijdens de interviews die voor deze uitgave zijn afgenomen:

1. Dr. Silvester Draaijer, MSc., Programmamanager VU Onderwijswerkplaats (Vrije Universiteit Amsterdam)
2. Bob Mooijkenkind, MA, Onderwijsadviseur (De Haagse Hogeschool)
3. Dr. Kelly Beekman, MSc., Lector (Fontys Hogescholen)
4. Dr. Ludo van Meeuwen, MSc., Programmamanager Onderwijsinnovaties en Technologie (Technische Universiteit Eindhoven)
5. Floor Visser, MSc., Informatiemanager (Hogeschool van Amsterdam)
6. Ewoud de Kok, MSc., Oprichter en CEO (FeedbackFruits)
7. Jitske van Os, BBA., Algemeen Directeur (Dutch EdTech)
8. Dr. Erwin van Vliet, Universitair Hoofddocent en Opleidingsdirecteur Bachelor Psychobiologie (Universiteit van Amsterdam)
9. Michael Forshaw, CEO (EdTech Impact)

Bronvermelding

1. Kucirkova, N. *Understanding Evidence: A Brief Guide for EdTech Producers*. (2022).
2. Shengjergji, S. *et al. Environmental Impact of EdTech: The Hidden Costs of Digital Learning*. ebooks.uis.no/index.php/USPS/catalog/book/285 (2024) doi:10.31265/USPS.285.
3. Schoors, W. *An Analysis of the European EdTech Ecosystem*. (2023).
4. McLaren, B. M., Richey, J. E., Nguyen, H. & Hou, X. How instructional context can impact learning with educational technology: Lessons from a study with a digital learning game. *Comput Educ* 178, 104366 (2022).
5. Learning EdTech Impact Funds (LEIF). *Jacobs Foundation* jacobsfoundation.org/activity/leif-learning-edtech-impact-funds.
6. Kucirkova, N. Why EdTech must prioritise research and innovation from the Global South. *World Economic Forum* weforum.org/stories/2023/05/to-advance-edtech-evidence-we-must-prioritise-research-and-innovation-from-the-global-south/ (2023).
7. Digital education content guidelines for schools in the works. *European Commission* education.ec.europa.eu/news/digital-education-content-guidelines-for-schools-in-the-works (2024).
8. *Global Education Monitoring Report 2023: Technology in Education: A Tool on Whose Terms?* (GEM Report UNESCO, 2023). doi:10.54676/UZQV8501.
9. 81% of users on edtech platforms face refund, trust issues: Survey. *The Economic Times* economictimes.indiatimes.com/tech/technology/81-of-edtech-platforms-users-face-refund-trust-issues-survey/articleshow/101260478.cms?from=mdr (2023).
10. Kucirkova, N. I., Livingstone, S. & Radesky, J. S. Faulty screen time measures hamper national policies: here is a way to address it. *Front Psychol* 14, (2023).
11. Breckon, J. *Using Research Evidence for Success: A Practice Guide*. (2016).
12. Kucirkova, N., Brod, G. & Gaab, N. Applying the science of learning to EdTech evidence evaluations using the EdTech Evidence Evaluation Routine (EVER). *NPJ Sci Learn* 8, 35 (2023).
13. Kucirkova, N. I. & Cermakova, A. L. *Consolidated Benchmark for Efficacy and Effectiveness Frameworks in EdTech*. ebooks.uis.no/index.php/USPS/catalog/book/270 (2024) doi:10.31265/USPS.270.
14. Foster, D. *et al. EdTech Quality Frameworks and Standards Review*. gov.uk/government/publications (2023).
15. Cermakova, A. L., Clary, A. & Havinga, B. Evaluation Mechanisms: A Way to Increase Trust in Educational Technologies? *European Edtech alliance* static1.squarespace.com/static/65cf66be42b8b06a33c319d8/t/66ebef748b5b5e11c6eb035a/1726738295019/Deep+Dive+I_Evaluation+Mechanisms_EdTech+Strategy+Lab_2024.pdf (2024).
16. Dore, R. A. *et al.* Developer meets developmentalist: improving industry–research partnerships in children’s educational technology. *J Child Media* 12, 227–235 (2018).
17. The 5Es of EdTech Impact. *University of Stavanger* ebooks.uis.no/index.php/USPS/catalog/series/edtech.
18. Kucirkova, N. EdTech Startups Make Evidence Their North Star. *Startups Magazine* startupsmagazine.co.uk/article-edtech-startups-make-evidence-their-north-star.



**Onderwijs
bewegen.**