**Evalueren**

Evaluatie is ‘het vaststellen van de kwaliteit van iets'. De kwaliteit van een te ontwerpen les(senreeks) verbetert aanmerkelijk wanneer gedurende het ontwikkelingsproces de les(senreeks) regelmatig kritisch wordt beoordeeld en op basis daarvan wordt verbeterd. Daarmee is evaluatie een kernactiviteit binnen de ontwikkelcyclus, zoals het ADDIE model hieronder laat zien.



Deze kwaliteitsbepalende kernactiviteit kan uiteenlopende evaluatieactiviteiten omvatten, gericht op wat de vier overige kernactiviteiten van de ontwikkelcyclus opleveren, te weten:

- de uitkomsten van de analyse (overzicht van behoeften, context, beschikbare kennis)

- een ontwerpvoorstel met specificaties voor de uiteindelijke les(senreeks)

- een eerste concrete uitwerking van een (deel van) een les(senreeks)

- een in gebruik genomen les(senreeks)

**Formatieve evaluatie**

Idealiter worden in combinatie met de eerste drie kernactiviteiten van de ontwikkelcyclus evaluatieactiviteiten uitgevoerd, om na te gaan op welke punten de kwaliteit van de les(senreeks) beter kan. Deze evaluatieactiviteiten maken deel uit van de zogeheten *formatieve* evaluatie. Op het moment dat alle kinderziektes zijn opgelost en de les(senreeks) is 'uitgerijpt', kan deze in praktijk worden gebracht.

**Summatieve evaluatie**

Kan de les(senreeks) daadwerkelijk worden uitgevoerd, dan is de evaluatie vooral gericht op het vaststellen van de effectiviteit ervan: leidt de les(senreeks) inderdaad tot de beoogde leeropbrengst bij leerlingen? Een dergelijke evaluatie heet een *summatieve* evaluatie[[1]](#footnote-1). Summatieve evaluatie maakt de ontwikkelcyclus rond. Deze eindevaluatie wijst uit of en in hoeverre de gerealiseerde ontwikkelopbrengst aan de oorspronkelijke verwachtingen voldoet, en wat dit voor consequenties heeft voor een eventuele vervolgontwikkeling. Eindevaluatie is daarmee de opmaat voor de analyse van een volgende ontwikkelcyclus

**Kwaliteitscriteria**

Formatieve en summatieve evaluatie dragen bij tot de totstandkoming van een les(senreeks) waarvan de kwaliteit in orde is. Daarbij zijn vier kwaliteitscriteria in het geding (Nieveen, 1999, 2009):

* relevantie; het ontwerp voorziet in behoefte en is gebaseerd op recente inzichten.
* consistentie; het ontwerp zit logisch in elkaar.
* verwachte bruikbaarheid; inschatting of het ontwerp bruikbaar is voor de situatie waarvoor het is bedoeld.
* verwachte effectiviteit; inschatting of werken met het ontwerp tot gewenste resultaten leidt.
* werkelijke bruikbaarheid; het ontwerp is bruikbaar voor de situatie waarvoor het is bedoeld.
* werkelijke effectiviteit; werken met het ontwerp leidt tot gewenste resultaten.

Als het om de bruikbaarheid en effectiviteit van een les(senreeks) gaat, is het zinvol om onderscheid te maken tussen de *verwachte* en *werkelijke* bruikbaarheid, respectievelijk de *verwachte* en *werkelijke* effectiviteit. De *verwachte* bruikbaarheid van nieuw onderwijsmateriaal kan bijvoorbeeld geëvalueerd worden, door docenten het materiaal te laten bekijken en er een oordeel over te geven. Maar de *werkelijke* bruikbaarheid en effectiviteit van het materiaal komt pas aan het licht, wanneer docenten en leerlingen er daadwerkelijk mee werken en daarop worden geobserveerd of bevraagd. Elk kwaliteitscriterium brengt zo z'n eigen evaluatieactiviteiten met zich mee.

De kwaliteitscriteria bouwen min of meer op elkaar voort. Immers, of een les(senreeks) effectief is, zal mede afhangen van de bruikbaarheid (*is de les(senreeks) wel uitvoerbaar in de klas?*) en van de relevantie en consistentie van de gestelde doelen en aangeboden inhouden. Het ene kwaliteitscriterium roept het andere als het ware op. Dat is de gebruikelijke gang van zaken bij een ontwikkeltraject. In het begin krijgen de relevantie en consistentie van de te ontwikkelen les(senreeks) de volle aandacht. Zodra de eerste concrete uitwerkingen ontstaan, verschuift de interesse naar (verwachte) bruikbaarheid van de les(senreeks). Vervolgens zal de aandacht naar effectiviteit verplaatsen. De focus van de evaluatie is dus in belangrijke mate afhankelijk van het evaluatiemoment

**Evaluatiemomenten met bijbehorende kwaliteitscriteria**

Uitgaande van een ontwikkelcyclus volgens *de* [ADDIE-aanpak](https://slo.nl/zoeken/%4010775/ontwikkelproces/), doen zich daarin in ieder geval vier momenten voor formatieve evaluatie voor. Deze hangen samen met de respectievelijke overgangen van de kernactiviteiten analyseren, ontwerpen, ontwikkelen, implementeren en (opnieuw) analyseren. Onderstaand figuur laat dit zien.

Elk van deze overgangen gaat namelijk gepaard met een concreet tussenproduct, te weten:

* Ontwerpspecificaties; eerste beschrijving van de les(senreeks) met daarin aandacht voor visie, doelen, inhoud, leeractiviteiten en toetsing.
* Globaal ontwerp; eerste uitwerking van de manier waarop visie, doelen, inhoud, leeractiviteiten en toetsing in de les(senreeks) zullen terugkomen.
* Gedeeltelijk uitgewerkte les(senreeks); Een onderdeel van de les(senreeks) is tot op concreet niveau uitgewerkt en zou gebruikt kunnen worden door de doelgroep.
* Volledig uitgewerkte les(senreeks). De les(senreeks) is volledig tot in detail uitgewerkt en kan in de praktijk gebruikt worden.

Deze tussenproducten en de totstandkoming daarvan, kunnen aangrijpingspunt zijn voor formatieve evaluatie. Elk tussenproduct stelt daarbij zijn eigen kwaliteitscriteria, zoals onderstaande tabel laat zien:

| **evaluatie-moment** | **te evalueren ontwikkelopbrengst** | **kwaliteitscriteria** | **mogelijke aandachtspunten:** |
| --- | --- | --- | --- |
| **analyse** | ontwerpspecificaties | relevantie | Heeft de [analyse](https://slo.nl/thema/vakspecifieke-thema/beroepsgericht/leergang-lob/blokken/analyse) een duidelijke en voor de beoogde belanghebbenden (leerlingen, (vak)docenten, school, dan wel samenleving) relevante onderwijsbehoefte(n) aan het licht gebracht? |
|  |  | (inhoudelijke) consistentie | Is de analyse voldoende breed en diepgaand uitgevoerd? Is er bijvoorbeeld werk gemaakt van* doelgroepanalyse,
* behoefteanalyse,
* contextanalyse,
* analyse van de bestaande kennisbasis?

Zijn de bevindingen van dergelijk vooronder-zoek logisch samenhangend verwerkt tot conclusies en aanbevelingen? |
|  |  | (relationele)consistentie | Zijn de beoogde belanghebbenden betrokken bij de totstandkoming van de analyse?Delen zij de conclusies en aanbevelingen? |
|  |  | verwachte bruikbaarheid | Levert de analyse bruikbare aanknopings-punten op om tot een ontwerpvoorstel te komen? |
| **ontwerp** | globaal ontwerp | relevantie | Beantwoordt het ontwerpvoorstel aan de aanbevelingen uit de uitgevoerde onderwijsanalyse? |
|  |  | (inhoudelijke) consistentie | Is het [curriculaire spinnenweb](https://slo.nl/thema/meer/curriculumontwikkeling/instrumenten/spinnenweb/inleiding/) in logische samenhang in het voorstel verwerkt? |
|  |  | (relationele) consistentie | Zijn de beoogde belanghebbenden betrokken bij de totstandkoming van het ontwerpvoorstel? Stemmen zij ermee in? |
|  |  | verwachte bruikbaarheid | Kan het ontwerpvoorstel worden uitgevoerd met de middelen en mogelijkheden die de school daarvoor biedt? Lijken de betrokken ontwikkelaars en beoogde uitvoerenden toereikend bevoegd en voldoende bekwaam? |
|  |  | verwachte effectiviteit | Lijkt het voorstel bij te dragen aan de gewenste situatie? Welke aanwijzingen of bewijzen zijn daarvoor te geven? |
| **ontwikkel** | gedeeltelijk uitgewerkte les(senreeks) | relevantie | Doen de leerdoelen en leerstof van het ontwerp ertoe? Zijn ze, vanuit de leerling, het vak, de school of de samenleving bekeken, zinvol of belangrijk? Zijn leerdoelen en leerstof relevant in het licht van de kerndoelen, eindtermen en/of vakexamenprogramma's, dan wel in het kader van het onderwijsbeleid van de school? |
|  |  | (inhoudelijke) consistentie | Is het curriculaire spinnenweb in logische samenhang in het ontwerp verwerkt? |
|  |  | (relationele)consistentie | Zijn de beoogde belanghebbenden betrokken bij de totstandkoming van het ontwerp? Zijn zij bereid aan de uitvoering ervan mee te werken? |
|  |  | verwachte bruikbaarheid | Is het ontwerp voldoende uitgewerkt om daadwerkelijk te kunnen worden uitgevoerd?Lijkt het ontwerp uitvoerbaar met de mensen, roosters, middelen en voorzieningen die de school daarvoor beschikbaar stelt? Voorziet het ontwerp in alternatieven bij onvoorziene om-standigheden (zoals disfunctionerende ICT)? |
|  |  | verwachte effectiviteit | Valt te verwachten dat leerlingen met dit ontwerp de gestelde leerdoelen zullen behalen? En dat zij efficiënt en effectief zullen leren en werken? |
| **implementatie** | volledig uitgewerkte les(senreeks) | relevantie | Is de essentie van het ontwerp (de *basisvisie*) in de uitvoeringspraktijk overgekomen en als zinvol of belangrijk gewaardeerd? |
|  |  | (inhoudelijke) consistentie | Is de les(senreeks) conform het ontwerp en volgens plan uitgevoerd? |
|  |  | (relationele) consistentie | Zijn de betrokken belanghebbenden tevreden over de uitgevoerde lessenreeks? Zijn zij bereid om de ingezette onderwijsontwik-kelingen voort te zetten? |
|  |  | werkelijke bruikbaarheid | Is de les(senreeks) uitvoerbaar gebleken met de daarvoor beschikbare mensen, roosters, middelen en voorzieningen? |
|  |  | werkelijke effectiviteit | Heeft het leren en werken met dit ontwerp geleid tot de beoogde leerresultaten? Zijn de gestelde leerdoelen door voldoende leerlingen gehaald? |

**Evaluatiemethoden en -activiteiten**

Om tot een goede les(senreeks) te komen is regelmatig evalueren een vereiste. De manier waarop kan variëren; deze hangt samen met het moment van evalueren, de dan beschikbare ontwikkelopbrengsten en de daarvoor geldende evaluatiecriteria en aandachtspunten. Zo zal een evaluatie die nagaat of de betrokken docenten en schoolleiding een ontwerpvoorstel relevant achten, anders in z'n werk gaan dan een evaluatie die insteekt op de vraag of leerlingen zelfstandig uit de voeten kunnen met de opdrachten in een ontworpen les(senreeks). Maar steeds is van belang dat de evaluatie systematisch wordt opgezet en uitgevoerd, zodat de opbrengst van de evaluatie zeggingskracht heeft. Zo is het bijvoorbeeld krachtiger om een collega te vragen aan de hand van een vooraf ontwikkelde lijst met kijkpunten naar uw les te komen kijken, dan hem/haar vijf minuten voor aanvang van de les te vragen of hij/zij even mee wil kijken. In het eerste geval zal de feedback veel gerichter zijn omdat de collega zich heeft kunnen voorbereiden.

In onderstaand schema worden verschillende evaluatiemethoden, evaluatieactiviteiten en daarbij mogelijk te gebruiken evaluatie-instrumenten beschreven en toegelicht (Tessmer, 1993; Nieveen, 2009):

|  |  |  |
| --- | --- | --- |
| **Evaluatiemethode** | **Mogelijke activiteit passende bij deze methode** | **Mogelijk te gebruiken instrumenten** |
| Screening:Als ontwikkelaar(s) aan de hand van een checklist nagaan in hoeverre een les-ontwerp of onderwijsmaterialen voldoen aan het oorspronkelijke ontwerpvoorstel, c.q. de gestelde ontwerpeisen. | Checklist langslopen | [Criterialijst evalueren](https://slo.nl/%405719/criterialijst/?reload=true) [Ontwerpeisen](https://slo.nl/%4011556/ontwerpeisen)[Checklist leerlinggerichtheid](https://slo.nl/%4015437/checklist-leerlinggerichtheid)[Checklist docent onafhan-kelijkheid](https://slo.nl/%4015438/checklist-docentonafhankelijkheid) |
| Focusgroep: Respondenten (bijvoorbeeld vakexperts, docenten) laten reageren op een les-ontwerp of onderwijsmaterialen, en wel aan de hand van een richtlijn met aan-dachtpunten waarop de ontwikkelaar(s) commentaar of suggesties verlangen | Checklist langslopen Interviewen | [Doorloopspel](https://slo.nl/%4015436/doorloopspel) |
| Walkthrough:Als ontwikkelaar(s) een lesontwerp of onderwijsmaterialen samen met respondent(en) doorlopen | Checklist langslopenInterviewenObserverenVragenlijst afnemen | [Criterialijst evalueren](https://slo.nl/%405719/criterialijst/?reload=true) [Ontwerpeisen](https://slo.nl/%4011556/ontwerpeisen)[Checklist leerlinggerichtheid](https://slo.nl/%4015437/checklist-leerlinggerichtheid)[Checklist docent onafhan-kelijkheid](https://slo.nl/%4015438/checklist-docentonafhankelijkheid) |

|  |  |  |
| --- | --- | --- |
| Micro-evaluatie: Een kleine groep leerlingen een gedeelte van de ontwikkelde les(senreeks) of onderwijsmaterialen buiten de beoogde praktijksituatie laten (proef)gebruiken | InterviewenObserverenVragenlijst afnemenToets of leerverslag makenLogboek bijhouden | [Kijkwijzer collegiale lesobservatie](https://slo.nl/%4015439/kijkwijzer-collegiale-lesobservatie)[Docentenlogboek](https://slo.nl/%4015440/docentenlogboek/)[Leerlingen- en docentenenquêtes](https://slo.nl/%4015441/leerlingen-docentenenquete)[Checklist docenthandelen](https://slo.nl/%4015442/checklist-docenthandelen) |
| Try-out:De ontwikkelde les(senreeks) of onderwijsmaterialen binnen de beoogde praktijksituatie door de beoogde doelgroep (leerlingen en docenten) laten gebruiken | InterviewenObserverenVragenlijst afnemenToets of leerverslag makenLogboek bijhouden | [Kijkwijzer collegiale lesobservatie](https://slo.nl/%4015439/kijkwijzer-collegiale-lesobservatie)[Docentenlogboek](https://slo.nl/%4015440/docentenlogboek/)[Leerlingen- en docentenenquêtes](https://slo.nl/%4015441/leerlingen-docentenenquete)[Checklist docenthandelen](https://slo.nl/%4015442/checklist-docenthandelen) [Beoordelen van leerresultaten](https://slo.nl/%4015443/beoordelen-leerresultaten) |

Zoals hierboven als is weergegeven hangt de evaluatiemethode samen met het moment van evalueren, de dan beschikbare ontwikkelopbrengsten en de daarvoor geldende evaluatiecriteria. Het volgende schema visualiseert dit. De keuze voor de evaluatiemethode wordt sterk ingegeven door een combinatie van het ontwikkelstadium waarin het te ontwikkelen product verkeert en het evaluatiecriterium/kwaliteitsaspect dat geëvalueerd wordt.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Kwaliteits-aspect | Ontwikkel-stadia | Ontwerp-specificaties | Globaal ontwerp | Gedeeltelijk in detail uitgewerkt product | Volledig uitgewerkt product | Geïmplementeerd product |
| Relevantie |   | **Screening****Focusgroep** | **Screening****Focusgroep** | **Screening****Focusgroep** | **Screening****Focusgroep** |  |
| Consistentie |   | **Screening****Focusgroep** | **Screening****Focusgroep** | **Screening****Focusgroep** | **Screening****Focusgroep** |
| Bruikbaarheid | *verwachte* | **Screening****Focusgroep** | **Screening****Focusgroep** | **Focusgroep****Walkthrough** | **Focusgroep****Walkthrough** |
|   | *werkelijke* |   |   | **Micro-evaluatie** | **Micro-evaluatie****Try-out** |
| Effectiviteit | *verwachte* | **Screening****Focusgroep** | **Screening****Focusgroep** | **Focusgroep** | **Focusgroep** |
|   | *werkelijke* |   |   | **Micro-evaluatie** | **Micro-evaluatie****Try-out** |  **Survey****(quasi-) experiment** |

**Gebruikte literatuur**

Akker, van den, J., & Thijs, A. (2009). *Leerplan in ontwikkeling*. Enschede: SLO.

Nieveen, N.M. (1999). Prototyping to reach product quality. In J. van den Akker, R. Branch, K. Gustafson, N. Nieveen & Tj. Plomp (Eds.), *Design approaches and tools in education and training*. (pp.125-136). Dordrecht: Kluwer.

Nieveen, N. (2009). Formative evaluation in educational design research. In Tj. Plomp & N. Nieveen (eds.), *An introduction to educational design research* (pp. 89-101). Enschede: SLO.

Scriven, M. (1991). *Evaluation thesaurus (4th ed.).* Newbury Park, CA: Sage.

Tessmer, M. (1993). *Planning and conducting formative evaluation*. London: Kogan Page Limited.

**Voor wie meer wil weten:**

<http://leerplanevaluatie.slo.nl>

1. De bekende evaluatiedeskundige Stake heeft het verschil tussen formatieve en summatieve evaluatie eens als volgt omschreven: "When the cook tastes the soup, that’s formative evaluation, when the guest tastes it, that is summative evaluation" (in Scriven, 1991, p.19). [↑](#footnote-ref-1)