

H4 – Evalueren kun je leren

Conclusies van het hoofdstuk

- Evaluatie is niet zo eenvoudig als het lijkt
- Er zijn verschillende redenen voor evalueren
- Evaluatie gebeurt op verschillende niveau's
- Evaluatie heeft verschillende functies
- Evaluatie kan verschillende actoren betrekken
- Evaluatie kan op elk moment toegepast worden
- Evaluatie kan gebeuren via verschillende tools
- Evaluatie kan gebaseerd worden op verschillende leerdoelen

Begrippenlijst

Authenticiteit	Is de evaluatie een authentieke afbeelding van de verwachtingen?
Betrouwbaarheid	Hebben we gemeten wat we wilden meten? Is de meting consistent?
Certificerende functie	Evaluatieaanpak die leidt tot een certificaat of diploma.
Didactische balans	De samenhang tussen de gekozen leerstof, werkvorm en evaluatieaanpak.
CIPO-model	Oud model voor het beoordelen van kwaliteit op het mesoniveau dat rekening houdt met Context, Input, Proces en Output.
Criteriumgerichte evaluatie	Evaluatievorm waarbij de standaard (criterium) vooraf doordacht wordt vastgelegd.
Formatieve evaluatie	Evaluatie gefocust op het opvolgen van het leren van de leerlingen, met als doel het verdere leerproces te sturen en begeleiden.
Klassenraad	Een raad, samengesteld op het mesoniveau, met een decretaal vastgelegde beslissende functie over de toelating, begeleiding en deliberatie van leerlingen.
Kwaliteitszorg	Evalueren van de kwaliteit van het volledige programma van een school.
Normgerichte evaluatie	Evaluatievorm waarbij de standaard (norm) en het behalen van het leerdoel afhangt van zowel de leerprestatie van de leerling zelf als die van de medeleerlingen.
Peer assessment	Medeleerlingen met een gelijke status toetsen de beheersing van de onderwijssdelen en geven er een waardering aan.
Performance assessment a.k.a. gedragsevaluatie	Evaluatie die complex en authentiek gedrag probeert te beoordelen in reële, praktische situaties.
Performance-indicatoren	Statistische gegevens, verhoudingsgetallen, kosten en alle informatie die opbrengst van een instelling in relatie tot vooropgestelde doelen meet.
PISA	Cognitieve test op leesvaardigheid, wetenschappen en wiskunde, gecombineerd met een achtergrondvragenlijst om de kwaliteit van een onderwijssysteem te beoordelen.
Portfolio-evaluatie	Op basis van een verzameling eigen leerprestaties proberen aantonen dat de leerling een bepaald niveau of leerdoel bereikt heeft.
Predictieve functie	Evaluatieaanpak waaruit men voorspellingen zal doen over verdere schooloriëntatie, slaagkansen en studiekeuze.
Recentheid	Is de evaluatie en het criterium relevant in de hedendaagse context?
Referentiekader Onderwijs (OK)	De huidige basismethode voor het beoordelen van onderwijskwaliteit op het mesoniveau.

Responsabilisering	Scholen moeten op mesoniveau bewijzen dat ze kwaliteitsvolle instructie aanpak hebben en de minimumdoelen bereiken aan het macroniveau.
Selectieve functie	Summatieve evaluatie die leidt tot een formele selectie die bepalend is voor het verdere gevolgde traject in het leerproces.
Self assessment	Leerlingen toetsen zelf hun beheersing van de onderwijsdoelen en geven er zelf een waardering aan.
Summatieve evaluatie	Eindevaluatie gefocust op het beslissen over het al dan niet bereiken van de onderwijsdoelen en de beslissingen over de gevolgen ervan.
Validiteit	Komt de meting overeen met wat we nastreefden?
Vlaamse toetsen	Verplicht af te leggen toetsen in alle Vlaamse scholen die peilen naar de vaardigheden Nederlands en wiskunde bij leerlingen op vier verschillende scharnierzak. De resultaten beoordelen zowel het individuele leerproces van een leerling als de kwaliteit van het onderwijssysteem.
Zelfgerichte evaluatie a.k.a. Ipsative assessment	Evaluatievorm waarbij de standaard bepaald wordt door het vergelijken van de neergelegde prestatie met eerdere leerprestaties.

Het belang van evaluatie

Waarom evalueren we?

- Sterktes en zwaktes van leerlingen vaststellen en hoe we hen als leraar kunnen helpen.
- Aangeven van een standaard
- Nakijken welke onderwijsdoelen de leerlingen bereikt hebben

Evaluatie hangt samen met de keuzes van instructieverantwoordelijken op instructieaanpak.

Constructive alignment – didactische balans: De samenhang tussen de gekozen leerstof, werkvorm en evaluatieaanpak.

Evaluatie als startpunt van een leerproces: Zijn de leerlingen klaar om nieuwe doel aan te pakken?

(Denk: Law of readiness van Thorndike, H2)

Een definitie

Evaluatie is een zeer ruim en contextafhankelijk begrip → een omvattende definitie is lastig

Evaluatie (Thorpe): Het totale proces van verzameling, analyseren en interpreteren van informatie over elk aspect van een instructieactiviteit, met als doel een uitspraak te doen over de efficiëntie of een andere impact.

Wat is evaluatie?

Metten → Evalueren → Beslissen

- Evalueren bestaat uit deelprocessen en deelproducten
 - o Verzamelen van informatie
 - o Analyse en interpretatie van informatie
- Evaluatie bevat een uitspraak over het al dan niet bereiken van de onderwijsdoelen
- Evaluatie verschilt afhankelijk van de context en het niveau
 - o Micro: Leereffecten bij leerlingen onderzoeken
 - o Meso: Monitoringactiviteiten & kwaliteitszorg (voorbereiding op doorlichting inspectie)
 - o Macro: Kwaliteit onderwijssysteem nakijken (vb. centrale toetsen)

Assessment: Het meetproces in evaluatie. (≠ evaluatie: het waarderen van de gemete uitkomst)

Dimensies in evaluatie:

- Aggregatieniveau
- Functie
- Actoren (Wie voert de evaluatie uit?)
- Type leerdoelen (Taxonomie van Bloom)
- Timing (Voor, tijdens of na het leerproces?)
- Technieken & tools

Aggregatieniveau

Microniveau

Evaluatie kan betrekking hebben op ALLE variabelen, processen en actoren in het microniveau.

Focus: Mate waarin de leerdoelen bereikt zijn bij leerlingen

Op welke basis beslissen we dat de leerresultaten voldoen?

Criteriumgerichte evaluatie: Standaard (= criterium) wordt vooraf vastgelegd.

- Keuze criterium is doordacht (bekijk vanuit bekende standaarden voor de leeftijdsgroep)
- Vb. Jongens zijn geslaagd als ze hoogstens 15 seconden over de sprint doen.
- Pluspunt: Zeer objectieve evaluatievorm
- ~ high-stake-assessment
- ➔ Een extreem aantal leerlingen kan (slechts) slagen voor een criterium.

Normgerichte evaluatie: Standaard (= norm) is afhankelijk van de prestatie van de leerlingen.

- Men gaat uit van een gemiddelde (of mediaan, kwartiel, ...)
- Het bereiken van het leerdoel hangt zowel van eigen leerprestatie af als die van de medeleerlingen.
- Nadeel: Norm is sterk afhankelijk van groepssamenstelling.
- Vb. Iedereen die sneller loopt dan het gemiddelde + 1 seconden is geslaagd.
- Vb. Rapporten geven mediaan en klasgemiddelde weer
- ➔ Normgerichte evaluatie levert minder informatie op. (Beheerst de geslaagde student de vaardigheden wel?)

Zelfgerichte evaluatie (ipsative assessment): Standaard is afhankelijk van eerdere leerprestaties van de specifieke leerling.

- Resultaat leerling wordt vergelijken met eerdere leerprestaties
- Zolang de leerling erop vooruitgaat beschouwt men de leerling als geslaagd
- Pluspunt: Deze evaluatievorm houdt rekening met de capaciteiten van leerlingen.
- Is zelfgericht evaluatie wel eerlijk? Worden leerlingen niet ongelijk behandeld?
- Wordt vooral toegepast wanneer weinig leerlingen de vooraf opgelegde criteria niet behalen (vb. beroepsonderwijs)
- ~ low-stake-assessment

Kwaliteit van evaluatie op het microniveau:

- Evaluatie kan ook kijken naar leraren en de kwaliteit van hun toetsing
- Lage kwaliteit toetsen in onderwijs ➔ centrale examens en toetsen
- Inspectie: Scholen streven minimumdoelen zelden 'correct' na

Mesoniveau

Focus: Evalueren van de kwaliteit van het volledige programma van een school (= **kwaliteitszorg**)

Door onderwijsautonomie mogen scholen zelfstandig hun instructieaanpak ontwerpen.

→ **Responsabilisering:** Scholen moeten op mesoniveau bewijzen dat ze kwaliteitsvolle leer- en instructieaanpak hebben en de minimumdoelen bereiken aan het macroniveau.

Spanning responsibility ↔ improvement

Basis voor het beoordelen van onderwijskwaliteit op mesoniveau

= **Referentiekader voor onderwijskwaliteit (OK) (p.312-315):**

- Resultaten en effecten
 - o Leerlingen bereiken minimumdoelen
 - o Stimulatie studievoortgang en leerwinst
 - o Toegankelijkheid leren voor iedereen
 - o Langetermijneffecten
- Kwaliteitsontwikkeling
 - o School evalueert werking vanuit resultaten en effecten bij leerlingen
- Ontwikkeling stimuleren
 - o Doelen (uitdagende, samenhangende doelen; expliciet geformuleerde verwachtingen)
 - o Begeleiding (Gelijke onderwijskansen, studie- en psychische begeleiding, begeleidingen in samenwerking met ouders)
 - o Vormgeving onderwijsleerproces en leeromgeving (positief, divers en stimulerende omgeving)
 - o Opvolging (adequate feedback, sturing leerproces vanuit evaluatie)
- Beleid
 - o Beleid (beleid voor organisatie, samenwerkingen, communicatie,...)
 - o Onderwijskundig beleid (beleid voor leren en onderwijzen)
 - o Personeels- en professionaliseringsbeleid (beleid voor doeltreffend personeel)
 - o Financieel en materieel beleid (uitgewerkt, betaalbaar materiaal beleid)
 - o Veiligheid van leeromgeving

CIPO-model (voorganger OK):

- Context
 - o Bestuur, administratie, SES regio en leerlingen, infrastructuur, financiering, regels
- Input
 - o Kernmerken leerlingen, thuisomgeving en schoolteam
- Proces
- Output
 - o Leerprijzen

Doorlichting: Proces om de 6 jaar waarin de onderwijsinspectie nagaat of een school aan de verwachtingen van het OK voldoet.

- Meer dan alleen een inspecteur die een individuele leraar komt evalueren
- Individuele leerkracht krijgt GEEN persoonlijk rapport, rapport over hele school
- MAAR ook scholen zelf dienen het OK toe te passen om hun kwaliteit te beoordelen.

Ontwikkelingsschalen:

Beneden verwachting, benadert verwachting, volgens verwachting, overstijgt verwachting

→ Scholen motiveren om eigen kwaliteit te blijven ontwikkelen.

Macroniveau

Focus: Kwaliteit van het onderwijssysteem onderzoeken.

We concretiseren het begrip kwaliteit en splitsen op in **performance-indicatoren**.

Internationale aanpak: Verschillende landen werken samen via internationale organisaties en koepels om de kwaliteit van onderwijssystemen in kaart te brengen en te vergelijken.

Internationale performance-indicatorenonderzoeken: PISA, PIRLS & TALIS

Performance-indicatoren: Statistische gegevens, verhoudingsgetallen, kosten en alle informatie die de opbrengst van een instelling in relatie tot de vooropgestelde doelen meet.

- Vormen basis voor de onderzoeken
- Vb. Instroomaantallen, gemiddelde scores, aantal afgeleverde diploma's,...
- Moeten gericht zijn op het verbeteren van een onderwijssysteem

PISA – Programme for International Student Assessment:

- Cognitieve test voor 15 jarigen
 - o Leesvaardigheid, wiskundige geletterdheid & wetenschappelijke geletterdheid
 - o Elk jaar één hoofddomein en twee nevendomeinen
 - o Nadruk: Functionele kennis en vaardigheden
- Achtergrondvragenlijst
 - o Ook in te vullen door ouders en school (→ meer grip op context)
- ➔ Evaluatie van het onderwijssysteem op macroniveau (NIET individuele leerprestaties)
- Scholen krijgen schoolfeedbackrapport met schoolspecifieke resultaten
 - ➔ Helpt de school bij zelfevaluatie op het mesoniveau

Centrale examens:

- 2023: invoering centrale examens: "Vlaamse toetsen"
- Verplicht af te leggen in alle scholen
- Op 4 scharnierenmomenten (4^e leerjaar, 5^e leerjaar, 2^e middelbaar, 6^e middelbaar)
- Focus op Nederlands (spelling, grammatica & begrijpend lezen) en wiskunde
- ➔ Zwart op wit toetsen die beslissing geven over de beheersing van de onderwijsdoelen
- ➔ Kwaliteit van het onderwijssysteem meten
- Voor 2023: Geen centrale examens maar gestandaardiseerde toetsen:
 - o Interdiocesane proeven (KOV) en OVSG-toetsen (GO)
 - o Niet in alle Vlaamse scholen
 - o Resultaten hadden geen directe consequenties op doorstroom

Functie

Formatieve en summatieve evaluatie

Formatieve evaluatie: Evaluatie gefocust op het opvolgen van het leren van de leerlingen, met als doel het verdere leerproces te sturen en begeleiden.

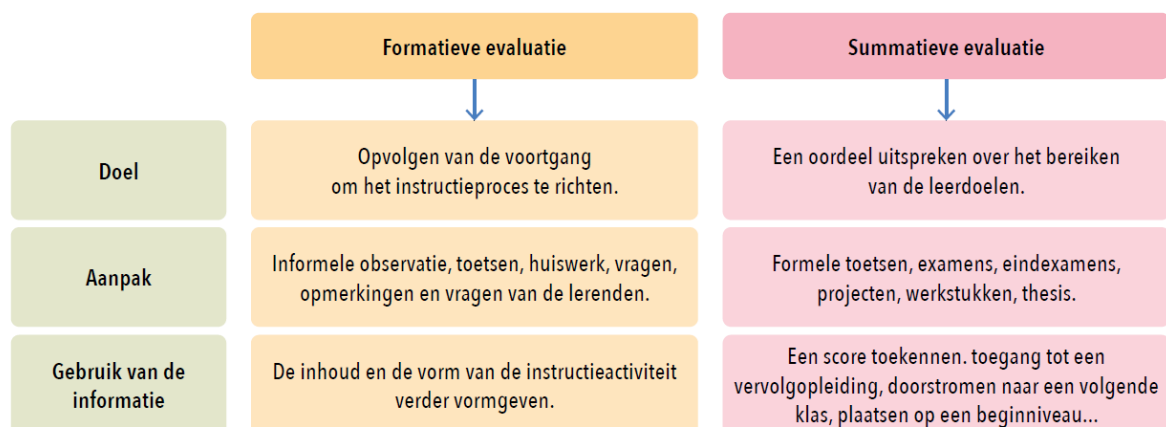
- Tijdens en doorheen het leerproces evalueren en verder helpen
- Als onderdeel van het leerproces, niet als apart onderdeel
- Assessment FOR learning
- Voordeel: Zeer efficiënt, zeker voor zwakkere leerlingen
- Nadeel: Leerlingen spannen zich niet altijd voldoende in voor formatieve toetsing.

Summatieve evaluatie: Evaluatie gefocust op het beslissen over het al dan niet bereiken van de onderwijsdoelen en de beslissingen over de gevolgen ervan.

- Na het leerproces een eindevaluatie doen en gevolgen (doorstroom) ervan bepalen
- Assessment OF learning

MAAR summatief en formatief toetsen kunnen perfect samen voorkomen. (formatief → summatief)

Inzetten op formatief toetsen tijdens leerproces heeft bevorderlijk resultaat bij summatieve toetsing.



Certificerende, selectieve en predictieve functie van evaluatie

Certificerende functie: Evaluatieaanpak die leidt tot een certificaat of diploma.

Selectieve functie: Summatieve evaluatie die leidt tot een formele selectie die bepalend is voor het verdere gevolgd traject in het leerproces.

Predictieve functie: Evaluatieaanpak waaruit men voorspellingen zal doen over verdere schooloriëntatie, slaagkansen en studiekeuze. (vb. CLB, SIMON-test)

Actoren

Meestal staat instructieverantwoordelijke centraal maar evaluatie kan ook uitgevoerd worden door:

- Externe experts
- Andere leerling(en)
- De leerling zelf
- De onderwijsinstelling
- Een toetsingssysteem

Self assessment: Leerlingen toetsen zelf hun beheersing van de onderwijsdoelen en geven er zelf een waardering aan.

- + Zelf je fouten leren inzien en hieruit leren
- + Meer verantwoordelijkheid en betrokkenheid in het eigen leerproces
- + Zelfreflectie en zelfregulatie (= metacognitie in Blooms taxonomie)
- + Leren van academische standaarden en verwachtingen
 - Subjectiviteit (sterke leerlingen schatten zichzelf te laag in, zwakkere te hoog)
 - Onbruikbaar voor evaluatie met score

Peer assessment: Medeleerlingen met een gelijke status toetsen de beheersing van de onderwijssdelen en geven er een waardering aan.

- + Samenwerkend leren
- + Vrij betrouwbare en bruikbare evaluatievorm
- + Academische voorwaarden evalueren bij anderen → leren zelf ook hieruit
 - Niet eenvoudig, vergt voorbereiding
 - Onbruikbaar voor evaluatie met score (vriendjespolitiek, veiligheid)

Peer feedback ↔ peer assessment:

- Peer feedback: Eerder formatieve evaluatie
 - o Feed up, feed back, feed forward
- Peer assessment: Eerder summatieve evaluatie

Beide vormen van assessment sluiten dicht aan bij een constructivistische visie.

Klassenraden:

- Decretaal bepaalde, regulerende functie
- Beslissende rol voor elke individuele leerling op het microniveau
- Opgesteld op het mesoniveau
- Adviezen en beslissingen hebben betrekking op kritische momenten in schoolcarrière
- Oriëntatieattesten: A-attest, B-attest (limitaties), C-attest (overzitten)
- Drie taken:
 - o Toelatingsklassenraad: Kijken of leerlingen die instromen uit andere school of richting geschikt zijn voor toelating.
 - o Begeleidende klassenraad: Opvolging bij tuchtzaken of langdurig ziek zijn
 - o Delibererende klassenraad: Beslissen over het (niet) slagen, bekrachtigen en oriëntatieattesten uitdelen.

Evaluatietechnieken en -tools

Toetsen ontwikkeld door de leraar

Zelf toetsen uitwerken verplicht leraren om een keuze te maken over het **vraagtype** (zie p.348 types).

Tips voor het uitwerken van goede vragen:

- Geen te lange tekst
- Geen verwijzing naar vorige vragen
- Opbouw in complexiteit
- Baken open vragen af
- Wees duidelijk in instructie
- Gebruik verschillende vraagtypes
- Test vraagtypes eerst formatief af (→ minder permanente effecten om te oefenen)
- Laat toets controleren op voorhand en reflecteer zelf achteraf

Rubrics (Zie opdracht en p.350-353)

Gedragsevaluatie

Performance assessment – gedragsevaluatie: Evaluatie die complex en authentiek gedrag probeert te beoordelen in reële, praktische situaties.

- Wordt vooral gebruikt bij stages, professionele opleidingen en bedrijfsopleidingen
- Eerst een reëel voorbeeld, dan in praktijk (poppen → patiënten)
- Rubrics zijn goede tool bij gedragsevaluatie

Portfolio-evaluatie

Portfolio-evaluatie: Op basis van een verzameling eigen leerprestaties proberen aantonen dat de leerling een bepaald niveau of leerdoel bereikt heeft.

Showcaseportfolio-evaluatie: Portfolio-evaluatie waarbij de leerling enkel de beste werken toont.

- Het portfolio dat de leerlingen opbouwen is een soort 'activiteiten-databank'.
- Past bij leergebieden waar continue ontwikkeling van leerdoelen kenmerkend is (vooral bij taal- en kunstvakken)
- Nadruk op het doorgemaakte leerproces
 - o Leerlingen documenteren het doorlopen leerproces zelf in functie van het behalen van competenties en beoordelen zowel hun sterktes als zwaktes
 - o Vb. lesvoorbereidingen, evaluatieverslag mentor, zelfreflecties, observatieverslag

Kritiek op portfolio-evaluatie:

- Administratieve overlast
- Leerlingen verwachten feedback maar krijgen die weinig, niet of te laat
- Te veel zelfafhankelijkheid zaait onzekerheid en vraagt naar meer begeleiding

Kwaliteit van de evaluatieaanpak

Vier criteria voor het beoordelen van de kwaliteit van de evaluatieaanpak:

- Betrouwbaarheid: Hebben we gemeten wat we wilden meten? Is de meting consistent?
 - o Gebruik van zeer concrete criteria → hogere betrouwbaarheid
 - o Stel meer vragen en val terug op voldoende observaties van leergedrag
 - o Maak gebruik van meerdere observatoren en beoordelaars
- Validiteit: Komt de meting overeen met wat we nastreefden?
 - o Wordt alle leerstof voldoende geëvalueerd?
 - o Weerspiegelt de evaluatie de gewenste onderwijsdoelen?
 - o ~~Betrouwbaarheid~~ → validiteit
 - o Tip: externe validiteit – laat je evaluatie bekijken door andere leerkrachten
- Authenticiteit: Is de evaluatie een authentieke afbeelding van de verwachtingen?
 - o Niet vanzelfsprekend voor complexe competenties
 - o Gebruik geen schriftelijke toets voor spreekvaardigheid
- Recentheid: Is de evaluatie en het criterium relevant in de hedendaagse context?
 - o Zeer belangrijk bij domein waarin kennisbasis continu evolueert en verandert
 - o Vakinhoud technologie van vandaag, zal over 20 jaar niet meer relevant zijn

Kritische reflecties op evaluatie

Kritische reflecties op evaluatie:

- Evaluatie is momentopname (onvoldoende om beheersing van doelen te beoordelen)
- Evaluatieaanpak weerspiegelt impliciet een visie op leren en instructie
- Gestandaardiseerde en efficiënte toetsingsmethoden kunnen de andere verdringen
- Gevaar dat niet alle leerdoelen getoetst worden en dus niet bereikt worden
- ➔ Evaluatieaanpak moet zelf door leerkrachten geëvalueerd worden
- ➔ Maak overzicht van toetsaanpakken op school en laat screenen

Kritiek op assessment (vb. formatieve evaluatie):

- Complex proces (hoge werklast voor leerkrachten)
- Zeker niveau van subjectiviteit
- Lastig meetbaar

MAAR we doen het toch wel, want het geeft een standaard aan de leerlingen.

Evaluatie op verschillende niveau's – Kirkpatrick:

- 1) Reactieevaluatie (Hoe vonden de leerlingen het?)
- 2) Leerevaluatie (Is er progressie in het leerproces?)
- 3) Gedragsevaluatie (Uit deze progressie zich in het gedrag? Is de progressie meetbaar?)
- 4) Resultaatsevaluatie (Leidt deze progressie ook tot échte resultaten?)

Voorbeeld: "Hoe vond je de lessenreeks?" → Examen → Gedrag op de stage na examen → Is het onderwijs erop verbeterd?