

Toetsing als kans voor leren

Een thematisch overzicht naar het waarom, wat, wanneer en hoe van formatief evalueren

Dominique Sluijsmans | augustus 2020

“We're always looking for the next big thing, when we should be doing the last big thing properly, which is formative assessment” (Dylan Wiliam)

Leerlingen brengen tijdens hun middelbare schooltijd vele uren door met hun medeleerlingen en leraren. Zij ontwikkelen zich tijdens deze periode tot jongvolwassenen, die klaar zijn voor een volgende stap in hun (school)loopbaan. In het onderwijs speelt toetsing voor leraren en leerlingen sinds jaar en dag een belangrijke en sturende rol. Toetsing geeft leerlingen en ouders inzicht in de voortgang van het leerproces, maar is voor scholen ook een middel om de kwaliteit van het onderwijs te bewaken. Negatieve toetservaringen kunnen echter leiden tot een gebrek aan motivatie, zelfvertrouwen en zin in leren (Natriello, 1987). In 2018 constateerde de [Onderwijsraad](#) bovendien een disbalans in de huidige toets- en examenpraktijk tussen het toetsen *van* het leren en het toetsen *voor* het leren. Om tot een evenwichtige balans te komen tussen onderwijzen, leren en toetsen, uiten scholen steeds meer hun behoefte aan kennis over de vormende en lerende functie van toetsing – het *formatief evalueren*. Ook stellen zij de vraag hoe formatief evalueren zich verhoudt tot beslissingen over overgang en diplomering. De informatie en kennis over formatief evalueren is inmiddels echter zo overweldigend, dat het voor leraren lastig is de essentie te doorzien en (als team) de juiste eerste stap te zetten. Het doel van dit thematisch overzicht is (teams van) leraren handvatten te bieden bij de implementatie van formatief evalueren. Voor dit doel is beschikbare en recente kennis gebundeld en focus aangebracht. Belangrijkste conclusies zijn dat formatief evalueren over succeservaringen voor alle leerlingen gaat, het niet kan zonder effectieve didactiek en een visie op leren voorwaardelijk is.

In Figuur 1 is de structuur van deze studie schematisch weergegeven. Het is mogelijk direct naar een specifiek (sub)hoofdstuk, de kernboodschappen of de bronnen te gaan door erop te klikken.



Figuur 1. Schematisch overzicht van de overzichtsstudie

DISCLAIMER:

In deze overzichtsstudie is de cognitieve psychologie (de wetenschap van het leren) en de didactiek (de wetenschap van het goed lesgeven) de basis voor de definitie van formatief evalueren. Om leerlingen tot leren en ontwikkeling te brengen, zijn een veilig pedagogisch klimaat en voortdurende aandacht voor [sociale en emotionele processen](#) echter voorwaardelijk. Voor succesvol formatief evalueren is kennis van deze processen nodig, maar deze worden hier niet expliciet besproken.

1. Wat is formatief evalueren en waarom is het belangrijk voor leren?

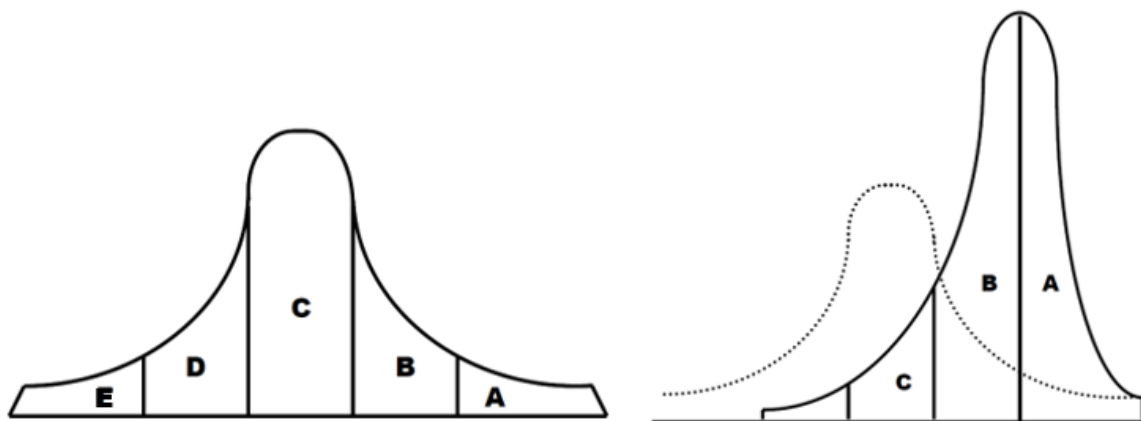
Vele scholen hebben formatief evalueren (hierna aangeduid als FE) definitief opgenomen in hun onderwijswoordenboek. In de praktijk zoekt men echter nog naar een eenduidige interpretatie en vertaling van FE. Zo zijn er meerdere benamingen (formatief toetsen, formatief handelen, formatief evalueren, formatief werken) en leeft er een aantal [misvattingen](#). Voorbeelden van misvattingen zijn dat FE zou gaan om een toets ('*Vandaag krijg je een formatieve toets*') of om het afschaffen van cijfers ('*Het is een formatieve toets, maar telt niet mee dus ik hoef er mijn best niet voor te doen*'). Maar wat is FE dan wel? In dit hoofdstuk wordt het speelveld van FE afgebakend in relatie tot de wetenschap over leren, didactiek en toetsing, met als doel te verhelderen waar formatief precies naar verwijst.

1a. Formatief evalueren is gebaseerd op kennis over leren en goed lesgeven

FE heeft één belangrijk doel: leerlingen ondersteunen in hun leerproces (Popham, 2008). Voor FE vormen de wetenschap over leren (de cognitieve psychologie) en de wetenschap over goed lesgeven (de didactiek; zie [hoofdstuk 2b](#)) de basis. Er is sprake van *leren* als bij leerlingen kennis en vaardigheden zijn verankerd in het langetermijngeheugen (Weinstein & Sumeracki, 2019). In andere woorden, leerlingen hebben de kennis en vaardigheden ook op de langere termijn tot hun beschikking en kunnen deze gebruiken in verschillende situaties - ook buiten de schoolcontext (Soderstrom & Bjork, 2015). Om te kunnen vaststellen of leerlingen hebben geleerd, is één prestatie op één moment direct nadat de leerling hierover les heeft gehad niet voldoende. Dit is wel nog de huidige praktijk: er is bijvoorbeeld een aantal weken les over bepaalde leerstof, dan volgt er een toets en daarna start de behandeling van nieuwe leerstof. Deze manier van toetsen past om drie redenen niet bij een geslaagd leerproces:

1. Leerlingen richten zich in de voorbereiding bij voorkeur op wat op de toets gevraagd wordt en de toetsvorm (bijv. een MC-toets). Zij zullen zich minder richten op het werken aan goede leerstrategieën die nodig zijn om tot langetermijnbeheersing van kennis en vaardigheden te komen (Bjork, Dunlosky & Kornell, 2013). Door deze aanpak zal een toets slechts meten wat een leerling *op dat moment op die toets in die situatie* kan laten zien. Er is echter geen zekerheid of de leerling de stof na een aantal weken nog beheerst en dus sprake is van *leren*. De kans is groot dat opgedane kennis snel weer wordt vergeten. Er is dus herhaling en gespreid leren nodig (zie de [vergeetcurve van Ebbinghaus](#)).
2. Bij toetsen aan het einde van een onderwijsperiode haalt meestal een deel van de leerlingen een onvoldoende. Hieruit blijkt dat een deel van de leerlingen de stof niet heeft begrepen en/of geleerd. Omdat nieuwe stof zich aandient, loopt de leerling kans op achterstand. Voorgaande stof is immers van belang voor het begrijpen van nieuwe stof. Leerlingen die (keer op keer) een onvoldoende halen zullen door het uitblijven van een succeservaring gedemotiveerd raken en hun zin om te leren verliezen (Bowers, 2019). Succeservaringen helpen bij het jezelf te motiveren tot verder leren (Kirschner et al., 2018) en [het geloof in eigen kunnen](#) te versterken (Bandura, 1993).
3. Toetsen voor cijfers zal het vergelijken van leerlingen benadrukken. We zien dit bijvoorbeeld terug in (gestandaardiseerde) toetsen waarbij de norm van de groep bepaalt hoe goed de groep het doet, resulterend in een curve zoals links afgebeeld in Figuur 2. In deze wijze van toetsing zal altijd een deel van de leerlingen zakken, waardoor de toets het karakter krijgt van een selectie-instrument in plaats van een kans voor leren (De Groot, 1966).

FE gericht op succes voor alle leerlingen vraagt om een visie op leren en lesgeven waarin leerlingen de kansen krijgen geboden om tot beheersing van de stof te komen. Bloom (1969) verbindt FE aan het leerlingen helpen tot beheersing te komen (*mastery learning*). Voor dit doel onderzocht hij niet alleen in hoeverre aanpakken die goed werken bij één-op-één onderwijs konden worden vertaald naar het lesgeven aan een groep. Ook keek hij naar welke leerstrategieën goed presterende leerlingen gebruikten ten opzichte van minder goed presterende leerlingen. Hij stelde dat de inzet van een beperkte variatie aan instructiestrategieën grote verschillen in toetsresultaten tot gevolg heeft (zie Figuur 2 links). Verstandiger is te investeren in een breder scala aan instructiestrategieën. Voorbeelden zijn het bewust en herhaaldelijke inzetten van tussentijds checken van begrip en daarop correctieve feedback geven. Als een leerling bijvoorbeeld 80% goed scoort op een eerste toets, kan door FE worden gekeken welke 20% nog niet lukt, waarom dat zo is en hoe zich dat verhoudt tot de volgende stap in het leerproces. Hiermee wordt voorkomen dat een leerling later uitvalt. Door leerlingen tijdig en net zolang te ondersteunen door FE tot een doel is behaald, wordt de kans op positieve resultaten op ‘echte’ toetsen verhoogd en daarmee de motivatie van leerlingen (Willingham et al., 2002). In andere woorden, door FE groeit het aantal leerlingen dat tot beheersing komt (zie Figuur 2).



Figuur 2. Toetsing gericht op brede variatie in resultaten en gebrek aan beheersing (links) versus toetsing gericht op minder variatie in resultaten en blijk van beheersing (rechts) (uit Guskey, 2005)

Om tot leren te komen, moeten leerlingen aan het denken worden gezet. Nadenken is de belangrijkste indicator dat er wordt geleerd (Coe et al., 2020). Leerlingen hebben hiervoor tijd en oefening in verschillende contexten nodig. Een leerproces verloopt grillig, gaat gepaard met veel vallen en opstaan en vraagt volharding. Presenteer leerlingen vooral taken die zij moeilijk vinden, maar niet te moeilijk vinden (*gewenste moeilijkheden*), zodat zij worden uitgedaagd tot nadenken. Ook zullen beginnelingen – leerlingen die nog weinig weten over de beoogde leerdoelen en leerinhouden – meer en andere ondersteuning nodig hebben dan gevorderde leerlingen (Kirschner & Clark, 2006). Voor alle leerlingen geldt dat het zorgen voor een snelle, eerste succeservaring belangrijk is. Een voorbeeld is het geven van een eerste vraag, opdracht of toets waarvan je als leraar weet dat alle leerlingen deze foutloos kunnen maken.

Leerlingen zullen altijd al iets weten en kunnen. Dit zien en waarderen geeft hen een eerste gevoel van succes. Om tot leren te komen is het achterhalen wat leerlingen al weten en kunnen een eerste stap. Vanuit de definitie van leren ligt de primaire focus van FE vervolgens op het tijdig achterhalen wat de leerling nog *niet lukt*, *vergeet* of waar hij *misvattingen* over heeft (Soderstrom & Bjork, 2015). Hierdoor 1) kan de leraar tijdig bepalen waar aanpassing nodig is van de *instructiestrategie*, en 2) weet de leerling wat aangepast moet worden in de *leerstrategie*. FE verwijst dus niet naar een set van ‘losse’ instrumenten, maar naar een zorgvuldig ingericht proces waarin informatie uit doelgerichte activiteiten door leraren en leerlingen wordt gebruikt voor het leerproces (Popham, 2008). Het proces wordt pas formatief, als de verzamelde informatie ook daadwerkelijk bijdraagt aan het leerproces.

Formatief verwijst ook niet naar een eigenschap van een toets, maar naar de eigenschap van de beslissing die een leraar en leerling neemt op basis van de verzamelde informatie (Wiliam, 2011). In het proces van FE kunnen toetsen weliswaar een rol hebben (zie [hoofdstuk 2b](#)), maar de toets zelf maakt een proces nog niet formatief. Dat is ook de reden dat een formatieve toets niet bestaat, maar wel *formatieve beslissingen*.

FE vraagt dus om actie van de leraar én leerling (Sadler, 1989). Het stellen van [hoge verwachtingen](#) en een [mindset](#) gericht op het creëren van kansen voor alle leerlingen op succes zijn belangrijke uitgangspunten voor FE. Hoge verwachtingen leiden ertoe dat leerlingen zich sneller ontwikkelen en meer leren (Rosenthal & Jacobson, 1968), maar zorgt ook voor eerlijke kansen. Onderwijs dat er vanuit gaat dat een deel van de leerlingen per definitie zal of mag zakken op een toets (bijvoorbeeld bij normgerichte gestandaardiseerde toetsen), is ongeschikt voor FE (De Groot, 1966; Popham, 2008).

1b. Formatief evalueren kan alleen op basis van de juiste informatie over het leerproces

Zolang leerlingen nog aan het leren zijn, is informatie over het leerproces van leerlingen bedoeld voor formatieve beslissingen. Formatieve beslissingen zijn aanpassingen in de instructie- en/of leerstrategie in relatie tot beoogde leerdoelen met het doel de leerling vooruit te helpen en tot succes te laten komen. Dat betekent dat de leraar deze informatie op de juiste manier, op het juiste moment met het juiste doel verzamelt. De informatie moet helpen drie vragen te beantwoorden: Waar werkt de leerling naartoe? Waar staat de leerling nu ten opzichte van de beoogde doelen? Welke actie is nodig om de leerling verder te helpen? Door het steeds weer opnieuw ingaan van het proces van FE met deze kernvragen als uitgangspunt, krijgt FE een cyclisch karakter gericht op voortdurend leren (Hattie & Timperley, 2007; zie Figuur 3).



Figuur 3. FE is een cyclisch proces met drie kernvragen.

De leraar kan vijf formatieve strategieën inzetten om tot de juiste informatie te komen ten aanzien van de drie kernvragen uit Figuur 3 (Leahy et al., 2005):

1. *Het verhelderen van leerdoelen en delen van criteria voor succes.* Gebrek aan helderheid over leerdoelen en succescriteria is vaak de reden dat leerlingen niet goed presteren op een 'echte' toets (Black & Wiliam, 1998). Wat betekent bijvoorbeeld een goed betoog schrijven? En waaraan kun je dat zien?
2. *Het voeren van effectieve discussies met (mede)leerlingen en het organiseren van activiteiten die informatie over het leerproces geven.* Kern is helder te krijgen hoe leerlingen denken, wat zijn de redenen dat zij bepaalde stappen nemen? Welke keuzes maken zij en waarom? Op basis hiervan is het mogelijk leerlingen terugkoppeling te geven hoe zij een nieuwe stap kunnen zetten. Met andere woorden: hoe zij hun leerstrategie kunnen aanpassen.
3. *Terugkoppeling geven voor het verdere leerproces.* Het leren geven, ontvangen en kunnen handelen op basis van feedback versterkt de ontwikkeling van *feedback-geletterdheid* (Carless & Boud, 2018). In het voorbeeld van een betoog schrijven gaat het hier om leerlingen betrekken bij

het feedback geven op tussenproducten of deelaspecten van het schrijfproces zodat het leidt tot een vervolgactiviteit.

4. *Het activeren van leerlingen als belangrijke informatiebronnen voor elkaar.* Het geven en ontvangen van peer feedback zorgt voor wederzijdse verantwoordelijkheid en verhoogt tevens de betrokkenheid en prestaties van leerlingen. Ook helpt peer feedback omdat leerlingen elkaars taal goed begrijpen (Kwangsue, & MacArthur, 2010).
5. *Het activeren van leerlingen om verantwoordelijkheid te nemen voor het eigen leerproces.* Met deze strategie wordt informatie verzameld over de mate waarin het leerlingen al lukt hun eigen leerproces te kunnen reguleren (dat zij weten welke stappen zij moeten nemen in het schrijven van een betoog en hoe ze dat kunnen doen).

In [hoofdstuk 2b](#) worden concrete didactische voorbeelden gegeven hoe de vijf formatieve strategieën kunnen worden toegepast in de onderwijspraktijk met het doel het leerproces te versterken. De informatie die voor dit doel nodig is, reikt verder dan de informatie die wordt verkregen op een reguliere toets (Christodoulou, 2016). Een cijfer kan bijvoorbeeld wel inzicht geven in een bepaalde mate van beheersing (bijv. “80% van de sommen heb je goed gemaakt”), maar geeft weinig formatieve waarde. De informatie voor een formatieve beslissing zit in het verhelderen welke 20% van de lesstof *nog niet goed gaat* en *waarom*. Alleen met deze informatie kan er formatief wordt gehandeld, in termen van het aanpassen van de instructie- en de leerstrategie (Popham, 2008). Dit betekent dat bij FE geen cijfers worden gegeven. Cijfers zullen kortetermijnleren, snel vergeten en blokken voor een toets stimuleren (Winstone & Boud, 2020). Nadenken is een vereiste om tot leren te komen. Cijfers kunnen dit denkwerk belemmeren (Leahy et al., 2005). Als leerlingen een cijfer krijgen met feedback, zullen zij de feedback negeren (Butler, 1988). Ook stimuleren cijfers het vergelijken met elkaar.

1c. Formatief evalueren kent een helder speelveld

Conclusie van het voorafgaande is dat FE een cyclisch en interactief proces is waarin zowel de leraar als de leerling hard aan het werk is. Dit proces is gericht op het verzamelen van de juiste informatie over het leerproces. Uitgangspunt van FE zijn heldere doelen. Leerdoelen bepalen de keuze van de activiteiten die informatie leveren voor formatieve beslissingen.

Figuren 4a en 4b laten het speelveld van FE zien (Wiliam, 2006; 2011; 2016). In Figuur 4a staan in de bovenste rij de drie fasen van het FE-proces. In de linkerkolom staan de belangrijkste actoren in het onderwijs: de leraar, de leerling en de medeleerlingen. In de cellen staan de vijf formatieve strategieën. Samen vormen zij het speelveld van FE (Leahy et al., 2005).

	waar werkt de leerling naartoe?	waar staat de leerling nu?	wat is nodig?
leraar	1. verhelderen, delen en begrijpen van leerdoelen en criteria voor succes	2. realiseren van effectieve discussies en activiteiten die bewijs leveren voor leren	3. terugkoppeling geven gericht op verder leren
medeleerling		4. activeren van leerlingen als belangrijke informatiebronnen voor elkaar	
leerling		5. activeren van leerlingen in het stimuleren van verantwoordelijkheid nemen voor het eigen leerproces	

Figuur 4a. Het speelveld van FE: drie kernvragen en vijf formatieve strategieën.

Figuur 4b vertaalt de vijf strategieën naar het handelen van de leraar en de leerlingen. Waar de leerling naar toe werkt - de leerdoelen en succescriteria -, bepaalt het *onderwijsontwerp*. Leerdoelen bevatten de ontwerpcriteria voor de leraar: wat wil ik dat mijn leerlingen kennen en kunnen en wat betekent dat voor mijn onderwijs? De kwaliteit van toetsing zit daarmee vooral in de kwaliteit van de beschrijving van de doelen. Als deze helder zijn, is het eenvoudiger vast te stellen welke informatie moet worden verzameld om iets over het bereiken van de doelen en de weg ernaartoe te kunnen zeggen (Wiliam, 2016). Leerdoelen zijn dus beschreven voordat het onderwijs begint.

Responsief onderwijs verwijst naar wat de *leraar* doet om de juiste respons te ontlokken bij de leerling (Fletcher-Wood, 2018; Gulikers & Baartman, 2017). Daarbij zet de leraar zijn vakkennis, pedagogisch-didactische kennis en kennis over leren in (zie ook [hoofdstuk 3a](#)). Er kan echter pas sprake zijn van FE als ook de leerling hard aan het werk is. Een *actieve leerling* is een leerling die reageert op vragen van de leraar, zodat de leraar kan achterhalen of de leerling nog aan het denken en gemotiveerd is. Ook zijn er zonder een actieve rol van de leerling geen mogelijkheden om vast te stellen wat de leerling nodig heeft. Een actieve leerling is ook een leerling die verantwoordelijkheid neemt voor het aanpassen van de leerstrategie (wat moet ik anders gaan doen om het leerdoel te behalen?) Bij FE staat het zweet op de juiste rug.¹

	waar werkt de leerling naartoe?	waar staat de leerling nu?	wat is nodig?
leraar	ONDERWIJSONTWERP	RESPONSIEF ONDERWIJS	
medeleerling	ACTIEVE LEERLING		
leerling			

Figuur 4b. Responsief onderwijs en een actieve leerling: twee zijden van de FE-medaille.

Kernboodschappen hoofdstuk 1

- Formatief evalueren is gericht op het verkrijgen van een betrouwbaar beeld van het *leerproces* van de leerling en is gebaseerd op wat we weten over leren en goed lesgeven. Momenteel prestaties zijn onvoldoende betrouwbaar om leren te kunnen vaststellen.
- Formatief evalueren is een doelbewust gepland *proces* gericht op het verzamelen van informatie waarbij de leraar van cruciale betekenis is. De informatie helpt de leraar en de leerling bij twee typen formatieve beslissingen: 1) het versterken/aanpassen van de *instructiestrategie* en 2) het versterken/aanpassen van de *leerstrategie* van de leerling. De informatie kan pas ingezet worden voor formatieve doeleinden als deze informatie:
 - Inzicht geeft in wat de leerling nog niet begrijpt of nog niet goed doet (misvattingen blootlegt), zodat de leraar ruimte in kan bouwen om te kunnen inspringen op de misvattingen of problemen die hij constateert.
 - Inzicht geeft welke actie wordt verwacht van de leraar én de leerling.
- Gestandaardiseerde en/of normgerichte toetsen bieden vooral inzicht op schoolniveau, maar zijn minder geschikt voor het leveren van informatie die bruikbaar is voor het bepalen van de instructiestrategie en -leerstrategie. Voorkom of beperk het gebruik van deze toetsen voor het nemen van formatieve beslissingen.
- Formatief evalueren is een doelgericht, voortdurend en cyclisch proces en vereist interactie tussen leraar en leerlingen, beiden zijn hard aan het werk en aan het leren.
- Het formatieve speelveld wordt afgebakend door een helder onderwijsontwerp, een actieve rol van de leraar en een actieve rol van de leerling (zie [Figuur 4a](#)).

¹ Met dank aan Bert Nelissen (voorzitter College van Bestuur Stichting Innovo) voor deze uitspraak.

2. Hoe organiseer je formatief evalueren en tot welke effecten leidt dat?

In het eerste hoofdstuk is FE theoretisch onderbouwd en gekaderd. Het speelveld van FE laat zien dat tijdens drie terugkerende fasen in het leerproces (zie [Figuur 3](#)) leraar en leerlingen actief aan het werk zijn. In dit hoofdstuk staan de ontwerpprincipes voor FE centraal. Er worden voorbeelden gegeven hoe didactiek het proces van FE voedt en hoe FE kan worden gebalanceerd met de summatieve functie van toetsing.

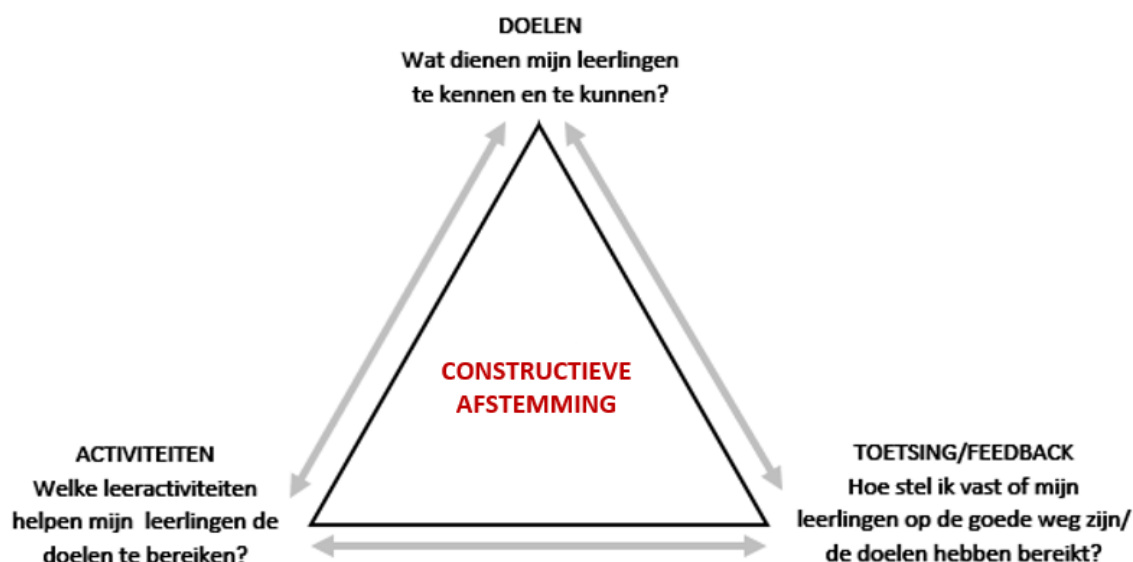
2a. Formatief evalueren is verankerd in een samenhangend curriculum

FE kan op verschillende niveaus worden georganiseerd. Afhankelijk van de tijdsduur heeft FE een bepaalde focus en impact (Wiliam, 2016; zie [Figuur 5](#)).

Vorm	Lengte	Focus	Impact
Lang	Vier weken tot een jaar	Over periodes heen, kwartalen, semesters, jaar	Verbeterde monitoring van leerlingen, een sterk consistent en uitgelijnd curriculum
Medium	Eén tot vier weken	Binnen en tussen onderwijsperiodes	Verhoogde leerlingbetrokkenheid bij toetsing en meer kennis van de leraar over het leerproces
Kort	Van minuut-tot-minuut, van dag-tot-dag	Binnen en tussen lessen	Verhoogde leerlingparticipatie en responsiviteit leraar

Figuur 5. Vorm, lengte, focus en impact van formatief evalueren (Wiliam, 2016; p. 114).

De verschillende vormen FE staan niet los van elkaar. Als er binnen en tussen de lessen geen sprake is van FE, zal er geen sprake zijn van impact van FE op de langere termijn. Om dit te voorkomen, is een doordacht en samenhangend curriculum nodig. Vanuit dit curriculum kan worden bepaald hoe FE op de korte en langere termijn wordt georganiseerd. In een curriculum sluiten de doelen, de leeractiviteiten en de wijze van toetsing naadloos op elkaar aan. Deze aansluiting wordt *constructieve afstemming* genoemd ([constructive alignment](#); Biggs, 1996; zie [Figuur 6](#)).



Figuur 6. Een samenhangend curriculum vraagt om een heldere aansluiting tussen doelen, leeractiviteiten en toetsing/feedback.

FE vormt de brug tussen doelen, leeractiviteiten en toetsing. Deze brugfunctie kan echter pas worden gerealiseerd als er aan twee voorwaarden is voldaan. Een eerste voorwaarde is helderheid welke kennis en vaardigheden leerlingen dienen te leren op *curriculumniveau* (de langetermijndoelen). Een tweede voorwaarde is een gebalanceerd en gefocust ontwerp van het onderwijs dat leerlingen gaat helpen deze kennis en vaardigheden te ontwikkelen en uiteindelijk te beheersen. Een dergelijk ontwerp wordt ook wel een *leerlijn* genoemd (*learning progression*; Popham, 2008). Dit ontwerp dient uiting te geven aan een gebalanceerd, stringent, samenhangend, geschikt, gefocust en belangrijk curriculum (William, 2013). Een leerlijn knipt het leren van grote curriculumdoelen op in kleinere stappen en bijbehorende taken (*backward design*; Wiggins & McTighe, 2006). Deze stappen worden in een bepaalde volgorde genomen en geleerd in het licht van het curriculumdoel. Drie uitgangspunten zijn bij het uitwerken van een leerlijn van belang (Popham, 2008): 1) denk bij het kiezen van de curriculumdoelen meteen na hoe je hierover tijdens het onderwijs informatie kunt gaan verzamelen bij leerlingen. Kies alleen doelen (met onderliggende kennis en vaardigheden) waarover je informatie voor formatieve beslissingen kunt verzamelen, 2) zoek een goede balans tussen de grootte van het curriculumdoel en de condities waaronder deze moeten worden behaald (het heeft bijvoorbeeld geen zin een doel te stellen waarvoor er in de tijd te weinig mogelijkheden zijn om leerlingen te laten oefenen en herhalen) en 3) kijk naar de raakvlakken tussen inhouden. Zo is bijvoorbeeld voor een aantal belangrijke concepten in natuur- en scheikunde wiskundige kennis nodig (Christodoulou, 2016).

Helderheid van doelen en een duidelijke structuur heeft een positief effect op de motivatie van leerlingen te willen leren (Seidel et al., 2005). Om de leerlijn en de doelen voor leerlingen helder te maken is het van belang de doelen in leerlingtaal te beschrijven. Probeer leraartaal te voorkomen (zie Figuur 7). Leerlingen hebben tijd nodig om de taal die voor een leraar als expert vanzelfsprekend is (dit heet ook wel de *kennisvloek*; Camerer et al., 1989) te leren spreken en begrijpen. Leerlingen krijgen vervolgens de kans deze doelen te bestuderen en deze in eigen woorden en/of criteria voor succes te formuleren (door bijvoorbeeld uitgewerkte voorbeelden; zie [hoofdstuk 2b](#)).

Doel in taal leraar	Doel in taal leerling
Leerlingen zijn in staat om in schrift een empirisch, wetenschappelijk verdedigbaar onderzoek te ontwerpen en te beschrijven waarin de relaties tussen twee of meer variabelen kan worden beschreven.	Je wordt gevraagd een onderzoek voor te bereiden waarin je duidelijk maakt hoe de relatie is tussen twee of meer variabelen (zoals tijd of temperatuur). Hoe je het onderzoek gaat ontwerpen en uitvoeren schrijf je op . Je schrijft ook hoe je de procedures die je leert in de lessen over het doen van wetenschappelijk onderzoek hebt verwerkt in jouw onderzoek.

Figuur 7. Het beschrijven van een doel van leraartaal naar leerlingtaal (uit Popham, 2008, p. 76/77).

Een andere manier om helderheid over de doelen en inhouden te geven is leerlingen een schematisch overzicht te presenteren aan de start van een onderwijsperiode (een *advance organiser*; zie p. 38 in [Wijze lessen](#); Surma et al., 2019). In een dergelijk overzicht wordt verhelderd welke stof ze tijdens de lessen kunnen verwachten en waarom, zodat leerlingen ook de verbanden tussen de lessen en de lesinhouden over een langere periode zien.

Leerlijnen zijn opgedeeld in behapbare stappen. Dit maakt het mogelijk om tijdig informatie te verzamelen die de leraar helderheid biedt of leerlingen de stof nog begrijpen (Christodoulou, 2016). Omdat tussentijds begrip zichtbaar wordt, kunnen leerlingen snel tot succeservaringen komen. Deze ervaringen verhogen de motivatie (Kirschner et al., 2018). Informatie op lesniveau helpt de leraar bij het uiteindelijke beslissen of curriculumdoelen zijn behaald (zie ook [hoofdstuk 2c](#)).

2b. Formatief evalueren heeft effectieve didactiek als katalysator

Het speelveld van FE wordt gekaderd door drie kernvragen en vijf strategieën (zie [Figuur 4a](#)). De leerlijnen (*learning progressions*; zie [hoofdstuk 2a](#)) vormen de basis voor het onderwijsontwerp. Binnen het FE-proces doorloopt de leraar steeds vier stappen. De vier stappen zijn (Popham, 2008):

1. Bepalen op welke momenten tijdens de lessen informatie verzameld moet worden om te weten of tijdige aanpassing van de instructiestrategie nodig is. Vragen die de leraar hierbij kan stellen zijn: Waar wil ik dat mijn leerlingen actief mee zijn (over aan het denken zijn)? Wanneer rond ik een specifieke stap in de leerlijn af? Wanneer start ik met nieuwe lesstof? Wat zijn momenten waarvan ik weet dat leerlingen vaak een verkeerde weg inslaan of fouten maken? Juist op die 'scharniermomenten' is het belangrijk informatie over begrip te verzamelen, om vertraging of frustratie op een later moment te voorkomen (Popham, 2008; Wiliam, 2011).
2. Bepalen van de *didactische activiteiten* die nodig zijn de juiste informatie bij leerlingen te verzamelen voor het nemen van formatieve beslissingen (het responsieve onderwijs in [Figuur 4a](#)). Voor dit doel is een gereedschapskist nodig met [didactische handvatten](#).
3. Voordat de leraar de didactische activiteit uitvoert, bepaalt hij *wanneer een aanpassing van de instructiestrategie nodig is*. Wat moet ik van de leerlingen hebben gezien om al dan niet verder te kunnen gaan? Wat moet ik op groepsniveau hebben gezien? Als voorbeeld: als 75% van de groep niet minimaal 80% van de vragen goed heeft, is aanpassing van de instructiestrategie nodig.
4. De *daadwerkelijke aanpassing van de instructiestrategie*. In deze stap bepaalt de leraar wat nodig is om te voorkomen dat de leerling niet meer leert, gedemotiveerd raakt, of een verminderd geloof in eigen kunnen heeft. De aanpassing heeft tot doel de leerling een kans tot leren te bieden.

Analoog aan de stappen van de leraar, is de leerling ook actief (zie [Figuur 8](#)).

	waar werkt de leerling naartoe?	waar staat de leerling nu?	wat is nodig?
leraar	LEERLIJNEN (<i>learning progressions</i>)	1. vaststellen van momenten waarop je informatie wilt verzamelen bij leerlingen 2. bepalen welke didactiek nodig is voor achterhalen juiste informatie (vragen stellen, feedback, toetsing als leerstrategie, ...) 3. bepalen wanneer aanpassing instructie nodig is (welke beheersing wordt verwacht op individueel en groepsniveau) 4. bepalen van aanpassing instructiestrategie op basis van verzamelde informatie	
medeleerling	1. bewust zijn hoe de momenten waarop leraar informatie gaat verzamelen over het leerproces kunnen helpen bij aanpassingen in leerstrategie 2. actief meedoen bij de leeractiviteiten zodat de leraar de juiste informatie kan verzamelen 3. bepalen wanneer aanpassing leerstrategie nodig is (en hoe deze gaat helpen om te voldoen aan de verwachte beheersing) 4. bepalen van aanpassing leerstrategie		
leerling			

Figuur 8. De stappen die de leraar en de leerlingen zetten in het speelveld van FE.

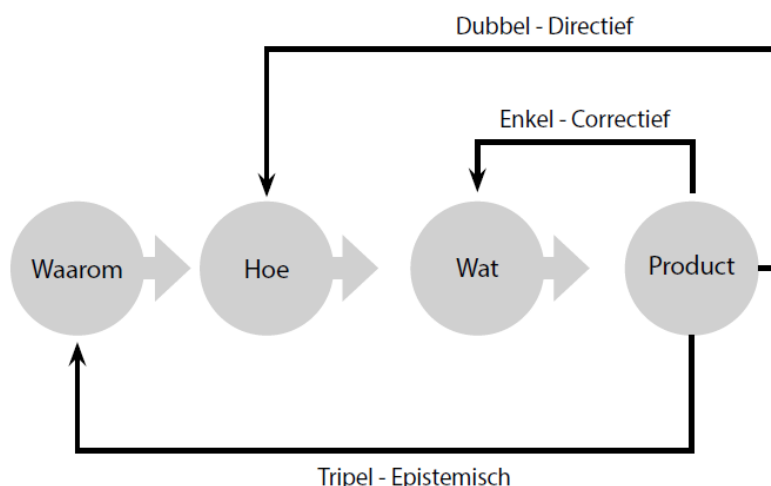
Hoewel de vier stappen voor FE voor de leraar én de leerling belangrijk zijn, vraagt FE vooral inzet van de juiste didactiek op het juiste moment. Didactiek is nodig om het denken en het leerproces van leerlingen zichtbaar te maken. FE is pas mogelijk als de didactische activiteit voldoet aan de volgende criteria:

- De activiteit zet leerlingen aan het denken (zij worden geactiveerd): de leerling kan niet passief blijven en zal iets moeten doen, bijvoorbeeld een vraag beantwoorden.
- De activiteit levert de leraar informatie waarmee hij de instructiestrategie kan aanpassen: uit de antwoorden van leerlingen blijkt dat een aantal leerlingen nog een misvatting heeft over bepaalde stof.
- De activiteit levert de leerlingen informatie waarmee zij hun leerstrategie kunnen aanpassen: de leerling krijgt actiegericht feedback van de leraar die begrijpelijk is en leidt tot een betere aanpak in het oplossen van een bepaalde vraag of opdracht.

Robuust onderzoek uit lerareneffectiviteitsstudies, de cognitieve psychologie en de geheugenpsychologie biedt veel didactische houvast ([Dirkx et al., 2018](#); [Surma et al., 2019](#)). FE kan niet zonder de inzet van effectieve didactiek. De volgende didactische principes zijn belangrijk om FE effectief te laten zijn voor leren:

- *Uitgewerkte voorbeelden.* Door het aanbieden en werken met voorbeelden van de gewenste prestatie kan de leraar laten zien wat wordt verwacht en wat de criteria zijn voor succes. Dit kan bijvoorbeeld door (met een visualiser) hardop sprekend een uitwerking van een bepaalde vraag, som, opdracht aan de leerlingen te presenteren. Hierbij kan de leraar meteen duiden wanneer welke kennis of vaardigheid belangrijk is in relatie tot leerlijn (zie [hoofdstuk 2a](#)). Dat geeft leerlingen inzicht in wat zij gaan leren en waarom.
- *Succescriteria opstellen en bespreken met leerlingen.* Om leerlingen een [neus voor kwaliteit te laten ontwikkelen](#), kan de leraar vragen een aantal voorbeelden van werk te ordenen van zwak naar zeer goed (bijvoorbeeld met de methode van [paarsgewijs vergelijken](#)). Begrip van criteria helpt leerlingen bij het komen tot een leeraanpak en het stellen van haalbare doelen.
- *Activeren van voorkennis.* Leren begint bij wat je al weet (Ausubel, 1960). Om erachter te komen wat de leerling al weet over de te behandelen lesinhoud, kan de leraar op verschillende wijzen de reeds aanwezige kennis van leerlingen achterhalen. Het toetsen met het doel eerdere kennis en ervaring op te halen in combinatie met actiegericht feedback, motiveert leerlingen te blijven leren en het geloof in eigen kunnen te versterken (Abel & Bäuml, 2020). Je kunt voorkennis ook ophalen op [meerdere manieren](#).
- *Het stellen van vragen.* Om te checken of leerlingen nog begrijpen wat je als leraar doet en beoogt, is het stellen van de juiste vragen een krachtig didactisch principe. Er zijn [vele manieren](#) om leerlingen een vraag te stellen. De essentie van een goede, diagnostische vraag is te achterhalen op welk moment in de leerlijn de leerling de lesstof niet meer begrijpt (Christodoulou, 2016), om er vervolgens de instructie- en leerstrategie er op aan te passen. [Meerkeuzevragen](#) zijn zeer geschikt voor formatieve doeleinden, waarbij het belangrijk is antwoorden te formuleren waarvan de leraar weet dat deze vaak door leerlingen worden gekozen, maar incorrect zijn. Een verkeerd antwoord geeft de leraar dan directe informatie waar de leerling nog extra instructie nodig heeft. Voor formatieve beslissingen mag een diagnostische vraag niet te complex zijn en moet deze tijdig in het leerproces worden gesteld. Het risico is dat de leraar niet kan achterhalen waar in het oplossingsproces een mogelijke denkfout wordt gemaakt. [Exit-tickets](#) die leerlingen aan het einde van een les invullen, kunnen de leraar inzicht geven waar de volgende les extra instructie nodig is.
- *Alle-leerlingen-geven-antwoord-aanpakken.* Om alle leerlingen betrokken te houden, is de inzet van bijvoorbeeld [wisbordjes](#) effectief. De leraar presenteert een open of gesloten vraag en elke leerling antwoordt. De leraar krijgt meteen inzicht in de antwoorden, op basis waarvan kan worden besloten of een aanpassing in de instructiestrategie nodig is.

- *Het geven van doelgerichte feedback.* Feedback heeft een zeer groot effect op het leerproces (Shute, 2008). Leerlingen kunnen op [meerdere manieren](#) met feedback aan het denken worden gezet. Om feedback te laten leiden tot een actie bij de leerling, dient de feedback aan te sluiten bij de doelen van de leerlijn. Er kan sprake zijn van *correctieve feedback* (de leraar zegt of het goed of fout is), *directieve feedback* (de leraar zegt waarom het fout en wat de leerling moet doen) of *kennisontwikkende/epistemische feedback* (de leraar stelt leerling vooral waarom-vragen; zie Figuur 9). Voor het leerproces is epistemische feedback het meest effectief (Hattie & Timperley, 2007), maar daar moet de leerling wel al klaar voor zijn. Als een leerling nog niet veel voorkennis heeft, is het beter eerst correctieve of directieve feedback te geven (Brooks et al., 2019).



Figuur 9. Drie vormen van feedback (uit: Op de schouders van reuzen, p. 132)

- *Het geven van actiegerichte feedback.* Bij feedback gaat het er vooral om dat de leerling iets mee kan en gaat doen (Carless & Boud, 2018; Winstone & Boud, 2020). Dit vraagt om feedback die de leerlingen in [actie](#) zet. Actiegerichte feedback krijgt dan het karakter van instructie, waarbij de feedback is ingebed in de taak (Sherrington & Caviglioli, 2020). Voorbeelden van actiegerichte feedback zijn de leerling vragen een (deel van een) eerdere taak nog een keer uit te voeren of een aantal vragen te beantwoorden. Belangrijk is dat de [feedback landt](#) en iets doet met de leerling.
- *Toetsing als leerstrategie.* In plaats van leerlingen te vragen de stof nog een keer te bestuderen, is het (herhaald) maken van een toets effectiever voor het leerproces. Toetsing heeft dan niet de betekenis van 'afrekenen'. De functie van de toets is te achterhalen wat de leerling nog weet, maar vooral wat de leerling niet heeft onthouden. Vergeten is hier dus functioneel! Of zoals Bjork en Bjork (2019) het zeggen: [vergeten is de vriend van leren](#). De toets als een kans voor leren is zinvol bij het regelmatig opruisen van reeds behandelde lesstof (wat weten de leerlingen nog?), bij de start van nieuwe lesstof (voorkennisactivatie; wat weten ze al van?) en bij het langetermijnleren versterken (*elke dag een toets*; Leeming, 2002). Toetsen wordt zo een [leerversterker](#) in plaats van een leerverslapper. Ook kunnen leerlingen elkaar helpen bij het toetsen door eerst zelf te antwoorden, dan uit te wisselen in tweetallen en vervolgens antwoorden plenair te bespreken met de leraar (*think-pair-share*). Het regelmatig inzetten van een toets als leerversterker stimuleert bovendien het gespreid leren en het zorgen dat leerlingen de stof minder snel vergeten (de kennis wordt onderhouden).
- *Dialogen en peer feedback.* Leerlingen kunnen in het proces van FE ook veel betekenen voor elkaar. Peer feedback werkt goed bij het reageren op elkaars antwoorden en het elkaar ondersteunen bij het werken aan taken. Peer feedback heeft niet alleen leerwaarde voor de feedbackontvanger, maar vooral ook voor de [feedbackgever](#). Ook informele gesprekken tijdens de les tussen leerlingen en leraar genereren voor de leraar rijke informatie (Ruiz-Primo, 2011).
- *Het leren zelfstandig te leren.* Het ontwikkelen van zelfregulatievaardigheden is een belangrijk doel van FE (Black & Wiliam, 1998; Panadero et al., 2018). Leerlingen zijn maar een deel van hun tijd op

school. Van hen wordt verwacht dat zij ook thuis actief zijn en zichzelf aan het werk en denken kunnen zetten. Leren je eigen leren te plannen, monitoren en te evalueren gaat echter niet vanzelf. Leerlingen hebben hierbij hulp nodig. [Voorbeelden](#) zijn het modelleren van een leerstrategie (Hoe kun je jezelf vragen stellen over de stof? Hoe kun je een goede samenvatting schrijven?), het ondersteunen bij het oefenen van de leerstrategie en het vragen naar de resultaten van het werken met de leerstrategie (Education Endowment Foundation, 2018). Het maken van [flashcards](#) helpt leerlingen actief met kernvragen en -concepten bezig te zijn en zichzelf op gespreide momenten te overhoren. Alhoewel zelfregulatie bij jonge leerlingen te ontwikkelen is, zullen beginners andere ondersteuning nodig hebben dan gevorderde leerlingen (Brooks et al., 2019). Ook het maken van goede flashcards vraagt aanvankelijke instructie en modellering.

Met het overzicht van de bovenstaande didactische principes, kan het speelveld van FE zoals afgebeeld in [Figuur 4a](#), [Figuur 4b](#) en [Figuur 8](#) worden aangescherpt (zie [Figuur 10](#)).

	waar werkt de leerling naartoe?	waar staat de leerling nu?	wat is nodig?
leraar	- schematisch overzicht leerlijn presenteren en uitleggen (<i>advance organiser</i>) - uitgewerkte voorbeelden laten zien en leerlingen uitleggen - leerlingen met uitgewerkte voorbeelden succescriteria laten opstellen	- stellen van diagnostische vragen - het geven van doel- en actiegericht feedback - uitgewerkte voorbeelden geven met processtappen	
medeleerling		- informele dialogen - stellen van vragen aan elkaar - peer feedback - toetsing als leerstrategie en uitwisselen van antwoorden (<i>think-pair-share</i>)	
leerling		- toetsvragen maken over de lesstof - jezelf overhoren - flashcards	

Figuur 10. Didactiek als katalysator voor FE

Wanneer welk didactisch principe zal werken, is afhankelijk van de context en doelgroep. De literatuur over FE laat dat ook zien: veel principes hebben een positief effect op het leerproces, motivatie en zelfregulatie (Surma et al., 2019), maar of een leraar op basis van de informatie de juiste formatieve beslissingen neemt hangt af van meerdere factoren (zoals de kennis en vaardigheden van leraar en leerling, de mate van ondersteuning in FE door de school en de helderheid van de doelen).

2c. Formatief evalueren dient in balans te zijn met summatieve toetsing

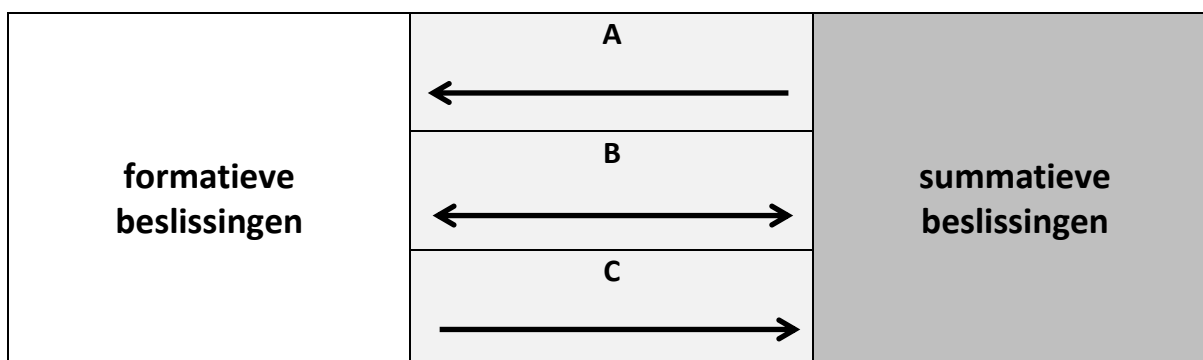
Bij FE staan de doelen van de leerlijnen (de curriculumdoelen) centraal: Waarover moet een beslissing worden genomen en welke informatie hebben we daarvoor nodig? Bij FE gaat het vooral om beslissingen ten gunste van het leerproces en het ervaren van succes. Leerlingen weten dat er geen zwaarwegende beslissingen meer zullen worden genomen op basis van een of een paar losse toetsmomenten, maar dat zij wel actief mee moeten doen om tot beheersing en uiteindelijk een diploma te komen. Bij FE mogen en moeten leerlingen fouten durven maken, deze zijn nodig om te weten wat nodig is voor het verdere leerproces. Dit betekent dat op meerdere momenten tijdens de les, een lessenreeks en een schooljaar zorgvuldig wordt vastgesteld of leerlingen nog actief zijn, aan

het denken zijn en zin hebben in leren. Hiervoor gebruiken leraren en leerlingen informatie die door de inzet van didactische activiteiten is verzameld (zie [Figuur 10](#)).

Om tot zorgvuldige, summatieve beslissingen over beheersing van curriculumdoelen, overgang of diplomering te komen, zijn cijfers op basis van momentprestaties waarop leerlingen zich kort van tevoren hebben voorbereid niet geschikt. Er is een goede balans nodig tussen formatieve en summatieve beslissingen, om recht te doen aan wat een leerling heeft laten zien. Summatieve beslissingen dienen te zijn gebaseerd op een onderbouwd leerdossier van de leerling waaruit beheersing van de curriculumdoelen blijkt. Ook is het belangrijk dat met name zwaarwegende beslissingen niet worden beïnvloed door vooroordelen van de leraar over de leerling (De Groot, 1966). Uit onderzoek naar rapportenvergaderingen blijkt dat leraren zelf aangeven dat summatieve beslissingen niet altijd eerlijk, gestructureerd of inzichtelijk worden genomen (Sleenhof et al., 2019).

Linda Allal (in Laveault & Allal, 2016) schetst drie scenario's die scholen kunnen helpen een gezonde balans tussen formatieve en summatieve beslissingen te realiseren (zie [Figuur 11](#)).

In scenario A neemt de leraar een summatieve toets af op basis van de doelen van de leerlijn. Op deze toets laat de leerling zien in hoeverre de verwachte kennis en vaardigheden binnen (een deel van) de leerlijn wordt beheerst. De toetsresultaten worden vervolgens gebruikt voor formatieve beslissingen. Dit scenario is minder geschikt voor FE en leren, omdat de toets waarschijnlijk te weinig informatie zal geven welke hulp een leerling precies nodig heeft. Ook is het risico dat deze toetsen te laat in het leerproces worden afgenomen.



Figuur 11. Naar een balans tussen formatieve en summatieve beslissingen (bewerking van Laveault & Allal, 2016).

In scenario B worden formatieve en summatieve beslissingen met elkaar verweven. Dit scenario zien we terug bij onderwijs waar leerlingen een portfolio bijhouden over hun groei ten aanzien van curriculumdoelen. Er wordt in het hoger onderwijs veel ervaring opgedaan met deze vorm van evalueren, het zogenaamde [programmatisch toetsen](#) (Van der Vleuten et al., 2015). Ook dit scenario heeft implicaties voor FE. Het risico is dat de formatieve en summatieve beslissingen vermengen en er daardoor geen transparantie is waarvoor welk moment van informatieverzameling is bedoeld.

In scenario C weegt de leraar informatie die is verzameld tijdens het onderwijs mee in het komen tot een summatieve beslissing. Dit kan bijvoorbeeld als een leerling tijdens het onderwijs meerdere keren heeft laten zien dat de kennis en vaardigheden worden beheerst en de curriculumdoelen zijn bereikt. In dit scenario wordt de formatieve beslissing duidelijk van de summatieve beslissing gescheiden: formatief is gericht op de leerlingen iets laten leren, summatief is gericht op de vraag of er geleerd is. Vanuit de theorie over leren is in scenario C het meest sprake van FE, omdat het gericht is op veel oefenen, herhalen, gespreid leren en het voortdurend onderhouden van eerdere kennis en vaardigheden (het vergeten wordt vertraagd). Eerst het leerproces, dan het toetsen, Niet andersom. Toetsing is bedoeld om het curriculum te dienen, niet om het curriculum te sturen.

Voor elk scenario geldt de aanbeveling om summatieve beslissingen te koppelen aan het uiteindelijke curriculumdoel en niet aan de beheersing van alle specifieke onderliggende kennis en vaardigheden. Wanneer summatieve beslissingen worden gekoppeld aan kleine stappen, is het risico dat kennis weer wordt vergeten en er geen sprake is van een langetermijneffect. Wanneer summatieve beslissingen worden gekoppeld aan curriculumdoelen, betekent dit dat het maar op een beperkt aantal momenten nodig is een cijfer te geven. Zo wordt voorkomen dat het leerproces vroegtijdig stopt. Ook blijven het leerproces en de summatieve toetsing van elkaar gescheiden (Wiliam, 2011; Winstone & Boud, 2020). Een formatief moment mag geen afname van een mini-versie van een summatieve toets zijn. FE is gericht op de kleine stappen die nodig zijn om een curriculumdoel te behalen (zie [hoofdstuk 2a](#)). De summatieve beslissing staat voor een gedeelde betekenis over een bepaalde waarde buiten de context van een specifieke school (Christodoulou, 2016). Dit betekent dat een cijfer 8 dat is behaald op een examen bij school X, van eenzelfde waarde is als een cijfer 8 behaald op school Y of Z. Dit kan echter alleen als er ook een gedeelde betekenis is over het curriculum en de doelen. De toetsing is er om het curriculum te dienen, niet om het curriculum te sturen.

Kernboodschappen hoofdstuk 2

- ☐ Constructieve afstemming (de aansluiting tussen doelen, leeractiviteiten en toetsing; *constructive alignment*) en het ontwerpen van leerlijnen (*learning progressions*) zijn het fundament van een samenhangend les- en curriculumontwerp dat alle leerlingen in staat stelt tot leren en beheersing te komen (mits zij actief meedoen!).
- ☐ Het bereiken van curriculumdoelen is moeilijk en vraagt om het opdelen in haalbare stappen waarin kennis en vaardigheden op opbouwende wijze kunnen worden ontwikkeld. Haalbare stappen zijn nodig om tijdig mogelijk afhaken te voorkomen, maar ook om leerlingen tussentijds succes te laten ervaren.
- ☐ Instructie is niet hetzelfde als FE, maar het *middel* om leerlingen de curriculumdoelen te laten bereiken. FE is het proces dat ervoor zorgt dat de instructie effectiever wordt voor het onderwijzen door de leraar en het leerproces van de leerling.
- ☐ Didactische principes als voorkennisactivatie, toetsing als leerstrategie, het stellen van vragen en feedback halen de informatie naar de oppervlakte die leraren nodig hebben voor het bepalen van de instructiestrategie. Ze helpen leerlingen bij het bepalen van hun leerstrategie.
- ☐ Om het proces van formatief evalueren van waarde te laten zijn voor het leerproces, kunnen summatieve beslissingen en het geven van cijfers worden uitgesteld tot het moment dat alle kansen om tot beheersing te komen zijn benut. Dit komt de betrouwbaarheid en validiteit van summatieve beslissingen ten goede. Voorkom te veel vermenging van formatieve en summatieve beslissingen.

3. Wat vraagt formatief van (teams van) leraren?

In de voorgaande twee hoofdstukken is het speelveld van FE afgebakend en is onderbouwd wat de organisatie van FE betekent in de concrete onderwijspraktijk voor de leraar en de leerling. Omdat FE vraagt om een herbezinning op de huidige onderwijs- en toetspraktijk, is inzicht nodig wat FE betekent voor professionalisering van leraren en de wijze van implementeren. In dit derde hoofdstuk komt aan bod welke kennis en vaardigheden FE vraagt van (teams van) leraren, welke cultuur nodig is voor FE en welke stappen zinvol zijn in de weg naar structurele implementatie van FE in de school.

3a. Formatief evalueren vraagt een set aan kennis en vaardigheden

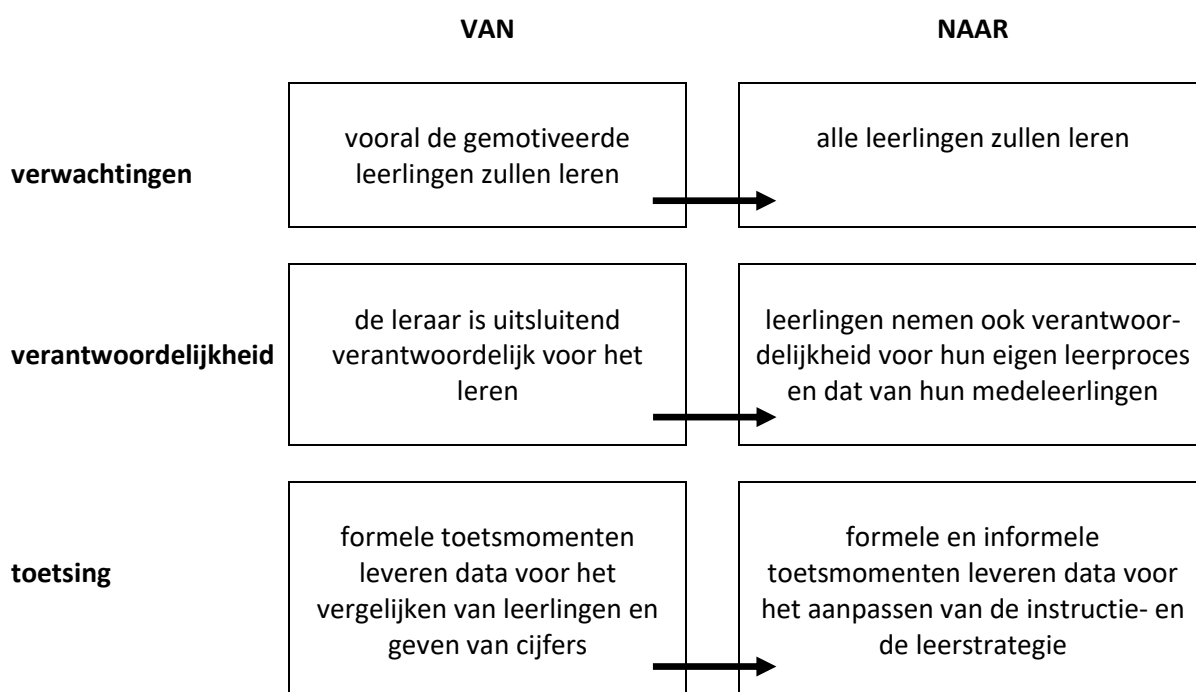
De kern van FE is om door inzet van didactiek tot de juiste informatie te komen over het leerproces van de leerlingen. Met deze informatie kan worden bepaald welke aanpassingen nodig zijn in de instructiestrategie (door de leraar) en de leerstrategie (door de leerling). Eerdere reviewstudies over FE en onderzoek naar didactische bekwaamheid en toetsbekwaamheid (*assessment literacy*) beschrijven wat een leraar moet kennen en kunnen om FE in te zetten zoals bedoeld (Gulikers & Baartman, 2017; Coe et al., 2020; Fletcher-Wood, 2018; Heitink et al., 2020; Panadero et al., 2018; Sadler, 1989; Schildkamp et al., 2014; Sluijsmans et al., 2013; Surma et al., 2019; Wiliam, 2016). Op basis van deze studies kan worden geconcludeerd dat leraren die FE succesvol inzetten in staat zijn om:

- Een veilig psychologisch en pedagogisch klimaat te creëren voor leerlingen in de tienerleeftijd met aandacht voor goed klassenmanagement. In een dergelijk klimaat worden verwachtingen verhelderd over de wijze van werken en leren en is er ruimte voor het maken van fouten zonder zwaarwegende consequenties. Het leerproces staat voorop, niet het presteren.
- Hun diepgaande kennis van de vakinhoud en de vakdidactiek te vertalen naar wat leerlingen moeten kennen en kunnen na de les, lessenreeks en op curriculumniveau (de leerlijn, zie [hoofdstuk 2a](#)). Maar ook in welke volgorde zij dat leerlingen gaan leren en op welke wijze.
- Hun kennis over *effectieve didactische principes* te vertalen naar activiteiten die informatie opleveren over het onderwijs en het leerproces. Deze activiteiten hebben een gemeenschappelijk doel: leerlingen aan het denken zetten en dit denken zichtbaar maken. Er wordt zoveel informatie verzameld als nodig is om formatieve en summatieve beslissingen te kunnen nemen. Specifieke vaardigheden van deze responsieve leraren zijn (Fletcher-Wood, 2018):
 - o focus kunnen aanbrengen in een les in relatie tot een specifiek leerdoel;
 - o leerlingen duidelijk maken hoe succes eruit ziet;
 - o tijdens de les erachter komen wat leerlingen denken en er vervolgens hun instructiestrategie erop aan te passen;
 - o leerlingen helpen bij het verbeteren van hun werk en hun leerstrategie;
 - o aan het einde van de les evalueren of leerlingen hebben geleerd (met bijvoorbeeld exit-tickets) en er vervolgens hun instructiestrategie erop aanpassen.
- Onderwijstechnologie op het juiste moment voor het juiste doel in te zetten. De inzet van digitale didactiek ondersteunt het verzamelen en interpreteren van informatie, maar kan ook ondersteunen bij de opbouw van een rijk leerdossier van en door leerlingen. Wilfred Rubens heeft een mooie koppeling gelegd tussen de twaalf bouwstenen voor effectieve didactiek ([Surma et al., 2019](#)) en de mogelijkheden van de [inzet van ICT](#).
- Leerlingen te activeren te leren van elkaar, bijvoorbeeld door vormen van peer feedback of het gezamenlijk bestuderen van voorbeelden.
- Leerlingen zelfstandig te leren leren en hun metacognitieve en zelfregulatievaardigheden te ontwikkelen.
- Vast te stellen wanneer, waarom, op basis van welke informatie een summatieve beslissing over curriculumdoelen kan en mag worden genomen.

Kernboodschap in het onderzoek naar de kennis en vaardigheden voor FE is dat de focus niet zozeer moet liggen op de praktische principes van FE (bijvoorbeeld het werken met wisbordjes of exit-tickets). Het succes van FE zit primair in het *begrijpen van het doel van deze principes* in relatie tot kennis over leren. Daarom is voortdurende en gezamenlijke professionalisering op schoolniveau essentieel (zie hoofdstuk 3b en 3c).

3b. Formatief evalueren vraagt om een lerende cultuur in de school

De implementatie van FE is niet makkelijk en vraagt om veel oefening en tijd. Startende leraren kunnen bijvoorbeeld wel al goed zien waar leerlingen moeite mee hebben, maar hebben vlieguren nodig om te achterhalen wat hier de oorzaken van zijn en hoe je dat achterhaalt (Bailey & Drummond, 2006). De uitdaging van FE is immers niet zozeer het interpreteren van informatie van leerlingen, maar het komen tot de informatie die leerlingen en leraren helpt tot juiste vervolgstappen te komen. Het werken aan FE vraagt om een cultuur waarin het team bereid is tijd te nemen om de weg van een presteer- naar een leercultuur te bewandelen (zie Figuur 12; Popham, 2008, p. 95):



Figuur 12. De weg naar een formatieve cultuur

Om deze beweging in gang te zetten is een lerende en professionele cultuur nodig. Kernelementen binnen een lerende cultuur zijn het gezamenlijk doelen stellen, het gezamenlijk onderwijs ontwerpen, bij elkaar in de keuken kijken en intercollegiale consultatie. De methodiek van [Stichting LeerKRACHT](#) is een goed voorbeeld van deze lerende manier van werken. In deze methodiek werken teams van leraren volgens een vast ritme aan gestelde doelen en succescriteria. [Onderzoek](#) naar de effecten van deze methodiek laat zien dat deze wijze van samenwerken de schoolcultuur en het enthousiasme van het team versterkt, maar ook tijd nodig heeft om tot volle bloei te komen (De Jong et al., 2020).

Een lerende of professionele cultuur kan ook een *formatieve cultuur* worden genoemd: door het leren en werken met elkaar als team wordt het team maar ook de school gevormd, en bouwen team en school steeds meer kennis op wat in welke context werkt. Analoog aan [Figuur 4a](#) kan ook het speelveld van formatief werken op het niveau van de school worden gekaderd (zie Figuur 13).

	waar werken we naartoe als team en als individuele leraren?	waar staan we nu als team en als individuele leraren?	wat is nodig als team en als individuele leraren?
schoolleiding	1. verhelderen, begrijpen en delen van doelen en criteria voor succes	2. realiseren van effectieve discussies en activiteiten die bewijs leveren voor teamleren	3. terugkoppeling geven gericht op verder teamleren
collega's		4. activeren van collega's als belangrijke informatiebronnen voor elkaar	
leraar		5. activeren van leraren in het stimuleren van verantwoordelijkheid nemen voor het eigen leren	

Figuur 13. Het speelveld van formatief werken als team (Wiliam, 2016; p.235).

3c. Formatief evalueren vraagt om een weloverwogen implementatie

Om tot structurele implementatie te komen van FE gericht op het (leren) leren van de leerling, is een eerste belangrijke stap dat er een gedeeld beeld is wat FE is (en vooral ook wat niet) (Popham, 2008). Zonder deze kennis bestaat het risico dat misvattingen ontstaan, er geen sprake is van de juiste informatieverzameling en dus onjuiste formatieve beslissingen worden genomen. Een tweede belangrijke stap is het leerproces van leraren te versterken. Op basis van jarenlange ervaring met professionalisering in FE, concludeerde Wiliam dat vijf factoren belangrijk zijn voor het leerproces van leraren (p. 165):

1. *Keuze*: Leraren krijgen de kans zelf te kiezen in welk aspect binnen het speelveld van FE zij zich het eerste willen bekwamen. Dit geeft hen een gevoel van controle en autonomie.
2. *Flexibiliteit*: Leraren krijgen de kans bepaalde formatieve principes flexibel aan te passen aan de context van hun vak of doel (zonder daarbij het doel van FE uit het oog te verliezen). Dus in plaats van wisbordjes kan ook een andere aanpak worden gekozen om tot antwoorden van alle leerlingen te komen.
3. *Kleine stappen*: Leraren nemen haalbare stappen die op korte termijn te realiseren zijn. Het gaat niet om perfectie, het onderwijs loopt immers gewoon door. Het veranderen van bestaande routines in een bestaande onderwijspraktijk is complex. Kleine stappen vergroten de kans op succes (en daarmee de motivatie). En elke stap levert – net als bij leerlingen – informatie over het leerproces op.
4. *Aanspreekbaarheid*: Leraren schrijven hun eigen leerplan, doelen en succescriteria. Een dergelijk plan helpt focus te houden, maar ook aanspreekbaar te zijn.
5. *Ondersteuning*: Om de ontwikkeling van FE te stimuleren, zijn vier vormen van ondersteuning belangrijk: helderheid over verwachtingen ten aanzien van onderwijsverbetering; focus houden op de zaken die werken voor leerlingen; tijd en ruimte voor innovatie; en verantwoord risico mogen nemen.

Om het leerproces van leraren en de vijf factoren daarvoor te realiseren, is reflectie op het schoolklimaat nodig om te bepalen of er voldoende voedingsbodem is voor structurele onderwijskundige veranderingen en verbeteringen. Bennett (2017) kwam tot de conclusie dat scholen die hiertoe in staat zijn [acht kenmerken](#) hebben:

- Betrokken en zeer zichtbare schoolleiders, met ambitieuze doelen, ondersteund door een sterk team.
- Goed gecommuniceerde, realistische, gedetailleerde verwachtingen die worden begrepen door alle betrokkenen in de school.
- Consistente werkprocessen.
- Een helder begrip van de cultuur: *'Dit is hoe wij handelen, en dit zijn onze gedeelde kernwaarden'*.
- Een hoge mate van commitment van het team aan de visie en strategie van de school.
- Een hoge mate van ondersteuning tussen leider en team, bijvoorbeeld in voldoende scholing.
- Aandacht voor detail en zorgvuldigheid in het uitvoeren en realiseren van het beleid en de strategie.
- Hoge verwachtingen naar leraren en leerlingen, vanuit de gedachte dat alle leerlingen even belangrijk zijn.

Uit deze kenmerken is af te leiden dat de rol van de schoolleider cruciaal is om een effectieve en succesvolle school te creëren (Galenkamp & Schut, 2018). De essentie van effectief leiderschap is leraren helpen te stoppen met het doen van goede dingen, om hen zo tijd te geven voor het doen van nog betere dingen (Wiliam, 2016). Het gemiddelde effect van onderwijskundig leiderschap op leerlingprestaties is vier keer sterker dan leiderschap dat alleen gericht op verandering en niet op de inhoud ervan (Robinson et al., 2008). Onderwijskundig leiderschap zorgt voor een omgeving waarin leraren zich voortdurend kunnen professionaliseren en als team kunnen samenwerken op de formatieve wijze zoals afgebeeld in Figuur 13. Een dergelijke omgeving heet ook wel een professionele leergemeenschap (PLG).

Een mooi voorbeeld hoe de implementatie van FE weloverwogen kan worden geïmplementeerd is het *Keeping Learning on Track* (KLT) programma ontwikkeld door o.a. Dylan Wiliam in samenwerking met het ETS (Educational Testing Service, 2010). KLT is een meerjarig programma waarin teams van leraren worden geprofessionaliseerd in het vormgeven van FE in de dagelijkse praktijk (zie voor voorbeelden Thum et al., 2015; Wylie & Lyon, 2015). Een eerste belangrijke stap in het programma is leraren een stevige kennisbasis over de theorie van FE en de formatieve strategieën te geven. Vervolgens gaan 10-12 leraren die met FE in hun praktijk aan de slag willen voor de duur van twee jaar een PLG vormen (Wiliam, 2016). Deze leraren hebben een vergelijkbaar takenpakket (bijv. zelfde leerjaar/zelfde domein).

De leraren gaan in de PLG als volgt te werk:

- Vanuit de schoolleiding wordt gegarandeerd dat de leden van de PLG maandelijks een sessie van 75-120 minuten kunnen organiseren tijdens of na contact- c.q. lesuren. Ook zorgt de schoolleiding voor de noodzakelijke ontheffingen van beleidsactiviteiten.
- De eerste bijeenkomsten staan in het teken van kennisontwikkeling over FE. Daarna kiest elke leraar één activiteit waarmee hij zal oefenen, bijvoorbeeld een van de didactische principes uit [Figuur 10](#) (zie [hoofdstuk 2b](#)).
- Elke leraar verantwoordt deze keuze in relatie tot de doelen van FE: hoe gaat de gekozen activiteit in jouw context en voor jouw leerdoelen bijdragen aan de juiste informatie die nodig is voor formatieve beslissingen? Wanneer is sprake van succes? (zie: Waar werken we naartoe; [Figuur 13](#)).
- Tussen de maandelijkse bijeenkomsten wordt geoefend en wordt tijd gereserveerd om bij elkaar te kijken in de klas (peer feedback) of andere samenwerkingsactiviteiten (Waar staan we nu; [Figuur 13](#)).
- Tijdens de PLG-bijeenkomsten bevragen leraren elkaar op ervaringen: Wat ging goed? Zijn succescriteria behaald? Zo nee, wat kan de reden zijn? Het doel van de PLG is te bepalen wat (een team van) leraren nodig heeft om de gestelde doelen voor FE te bereiken (Wat is nodig; [Figuur 13](#)). Van niet succesvolle ervaringen wordt geleerd, succesvolle ervaringen worden gevierd, gedeeld en verankerd in de nieuwe manier van werken.

Kernboodschappen hoofdstuk 3

- Investeer in de ontwikkeling van de (vak)didactische kennis en vaardigheden van leraren die voor FE nodig zijn, en geef hen voldoende tijd te oefenen en feedback te organiseren bij collega's.
- Investeer in onderwijskundig leiderschap als hefboom voor het creëren van een cultuur gericht op het leren van zowel leraren als leerlingen.
- Denk en werk vanuit een gedeelde visie op leren en toetsing, waarbij formatief evalueren wordt verankerd in de didactiek en het ontwerp van het curriculum (de leerlijnen).
- Organiseer professionele leergemeenschappen (PLG) die met een vast ritme en vaste opzet op een formatieve wijze werken aan de inzet en evaluatie van evidence-informed interventies.
- Volhard in het werken met FE, zodat het gangbare, op vergelijking gebaseerde en cijfergerichte onderwijs kan transformeren naar leer- en beheersingsgericht onderwijs, waar de focus ligt op het verbeteren van het onderwijzen en welbevinden van de leraar en het verbeteren van het leerproces en het welbevinden van de leerling.

De effecten van formatief evalueren samengevat

FE heeft tot doel leerlingen aan het denken te zetten en hen te ondersteunen bij het versterken van hun leerstrategieën. FE helpt de leraar bij het tijdig inzichtelijk krijgen of leerlingen de lesstof nog begrijpen en helpt hem bij het aanpassen van de instructiestrategie. FE vormt de brug tussen het actieve leerproces van de leerling en het responsief onderwijzen door de leraar. Uit eerder onderzoek naar FE is gebleken dat het lastig is effecten van FE eenduidig te beschrijven. Onderzoek naar feedback laat bijvoorbeeld zien dat er uiteenlopende resultaten (van heel positief tot heel negatief) te zien zijn (zie bijvoorbeeld Shute, 2008; Hattie & Timperley, 2007). Ook bemoeilijkt het gebrek aan een gedeelde definitie van FE in de literatuur, de wisselende methodologische kwaliteit van studies en de grote variatie in FE-interventies het moeilijk om eenduidige uitspraken te doen over de effecten van FE (Sluismans et al., 2013). Zoals bij iedere onderwijskundige interventie, zal (de grootte van) het effect van FE worden bepaald door de context en condities waarbinnen FE wordt georganiseerd. Wél is vanuit de wetenschap over leren en didactiek veel bekend over welke factoren bijdragen aan het leerproces van de leerling - in termen van langetermijnleren, motivatie, welbevinden - en de kwaliteit van de leraar/school - in termen van het behalen van de onderwijsdoelen, versterken van de kennisbasis over leren en didactiek, teamleren en een gezond pedagogisch klimaat. Voor de **leerling** zal FE effectief zijn als zij actief worden betrokken bij het begrijpen en opstellen van succescriteria, zij tijd krijgen om na te denken en feedback te verwerken en zij succeservaringen opdoen. Door FE bouwt de leerling niet alleen een sterke kennisbasis op, maar ontwikkelen zij vaardigheden waarmee zij – ook als de leraar er niet is – hun eigen leerproces kunnen reguleren. FE biedt leerlingen de kans in haalbare stappen te leren en te werken, waardoor het geloof in eigen kunnen wordt versterkt en daarmee ook de motivatie te willen blijven leren. Voor de **leraar** zal FE effectief zijn als hij kennis heeft over leren en binnen de eigen context en vanuit (pedagogische) waarden de juiste didactiek weet in te zetten. Het op de juiste wijze inzetten van didactische principes als voorkennisactivatie, het werken met uitgewerkte voorbeelden, het organiseren van feedback en het stellen van denkgerichte vragen, zal de leraar tijdig van informatie voorzien waar leerlingen staan en wat zij nodig hebben. Het effect van FE zal worden bepaald door het op een juiste wijze verzamelen en interpreteren van deze informatie ten behoeve van een formatieve beslissing die de leerling verder helpt. FE zal bovendien een positief effect hebben op de ontwerpvaardigheid van leraren, aangezien een samenhangend curriculum nodig is om FE te kunnen organiseren. Eerst helderheid creëren over de doelen en het curriculum, dan pas het proces van FE inrichten. De effecten van FE voor **de school** worden vergoot wanneer teams van leraren werken vanuit een gemeenschappelijke visie op leren en didactiek, en zij vanuit sterk onderwijskundig leiderschap de ruimte krijgen met FE te werken en van elkaar te leren.

Referenties

- Abel, M., & Bäuml, K. T. (2020). Would you like to learn more? Retrieval practice plus feedback can increase motivation to keep on studying. *Cognition*, 201. <https://doi.org/10.1016/j.cognition.2020.104316>
- Ausubel, D. P. (1960). [The use of advance organizers in the learning and retention of meaningful verbal material](#). *Journal of Educational Psychology*, 51, 267-272.
- Bailey, A. L., & Drummond, K. V. (2006). Who is at risk and why? Teachers' reasons for concern and their understanding and assessment of early literacy. *Educational Assessment*, 11(3-4), 149-178.
- Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, 28, 117-148.
- Bennett, T. (2017). [Creating a Culture: How school leaders can optimise behaviour](#). Department of Education.
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32, 347-364.
- Bjork, R., & Bjork, E., (2019). Forgetting as the friend of learning: implications for teaching and self-regulated learning. *Advances in Physiology Education*, 43, 164-167.
- Bjork, R. A., Dunlosky, J., Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, 64(1), 417-444.
- Black, P., & William, D. (1998). [Assessment and classroom learning](#). *Assessment in Education: Principle, Policy, & Practice*, 5(1), 7-74.
- Bloom, B. S. (1969). Some theoretical issues relating to educational evaluation. In R. W. Tyler (Ed.), *Educational evaluation: new roles, new means: the 63rd yearbook of the National Society for the Study of Education* (part II) (Vol. 69(2), pp. 26-50). Chicago, IL: University of Chicago Press.
- Bowers, A. J. (2019). Report card grades and educational outcomes. In Guskey, T., Brookhart, S. (Eds.) *What we know about grading: What works, what doesn't, and what's next*, (p.32-56). Alexandria, VA: Association for Supervision and Curriculum Development.
- Brooks, C., Carroll, A., Gillies, R. M., Hattie, J. (2019). [A matrix of feedback for learning](#). *Australian Journal of Teacher Education*, 44, 4, 14-32.
- Butler, R. (1988). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology*, 79(4), 474-482.
- Camerer, C., Loewenstein, G., & Weber, M. (1989). The curse of knowledge in economic settings: An experimental analysis. *Journal of Political Economy*, 97, 1232-1254.
- Carless, D., & Boud, D. (2018) [The development of student feedback literacy: enabling uptake of feedback](#). *Assessment & Evaluation in Higher Education*, 43, 8, 1315-1325
- Christodoulou, D. (2016). *Making good progress. The future of Assessment for Learning*. Oxford: University Press.
- Coe, R., Rauch, C. J., Kime, S., & Singleton, D. (2020). [Great teaching toolkit: Evidence review](#). Evidence Based Education in partnership with Cambridge Assessment International Education.
- Dirkx, K., Joosten-ten Brinke, D., Camp, G. (2018). [Ontwerprichtlijnen voor formatief toetsen vanuit de geheugenpsychologie](#). Den Haag: NRO.
- Education Endowment Foundation (2018). [Metacognition and self-regulated learning. Guidance report](#). London: Education Endowment Foundation.
- Educational Testing Service (2009). *Research rationale for the Keeping Learning on Track program*. Princeton: ETS.
- Fletcher-Wood, H. (2018). *Responsive teaching. Cognitive science and formative assessment in practice*. London: Routledge.
- Galenkamp, H., & Schut, J. (2018). *Handboek professionele schoolcultuur. Focus op koers en gedrag*. Huizen: Uitgeverij Pica.
- Groot, A. D. de (1966). *Vijven en zessen. Cijfers en beslissingen: het selectieproces in ons onderwijs*. Groningen: Wolters Noordhoff.
- Gulikers, J., & Baartman, L. (2017). [Doelgericht professionaliseren: formatieve toetspraktijken met effect! Wat DOET de docent in de klas?](#) Den Haag: NRO.

- Guskey, T. R. (2005). [*Formative classroom assessment and Benjamin S. Bloom: Theory, research, and implications*](#). Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81-112.
- Heitink, M., van der Kleij, F., Schildkamp, K., Kippers, W., & Veldkamp, B. (2020). [*Formative assessment: A systematic review of critical teacher prerequisites for classroom practice*](#). *International Journal of Educational Research*, 103. <https://doi.org/10.1016/j.ijer.2020.101602>
- Jong, A., de, Knijff, R., Lockhorst, D., Bulder, E., Klein, T., & van Tartwijk, J. (2020). [*Wat is de kracht van leerKRACHT? Oplegger bij 2e infographic van effectonderzoek*](#). Utrecht: Universiteit Utrecht/Oberon.
- Kirschner, P. A., Claessens, L., & Raaijmakers, S. (2018). [*Op de schouders van reuzen. Inspirerende inzichten uit de cognitieve psychologie voor leraren*](#). Meppel: Ten Brink Uitgevers.
- Kirschner, P. A., & Clark, R. E. (2006). [*Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching*](#). *Educational Psychologist*, 41(2), 75-86..
- Kwangsue, C., & MacArthur, C. (2010). Student revision with peer and expert reviewing. *Learning and Instruction*, 20, 4, 328-338.
- Laveault, D., & Allal, L. (2016). *Assessment for learning: meeting the challenge of implementation*. Dordrecht: Springer.
- Leahy, S., Lyon, C., Thompson, M., & Wiliam, D. (2005). Classroom assessment that keeps learning on track minute-by-minute, day-by-day. *Educational Leadership*, 63(3), 19-24.
- Leeming, F. C. (2002). The exam-a-day procedure improves performance in psychology classes. *Teaching of Psychology*, 29(3), 210-212.
- Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist*, 22(2), 155-175.
- Onderwijsraad (2018). [*Toets wijzer*](#). Den Haag: Onderwijsraad.
- Panadero, E., Andrade, H., Brookhart, S. (2018). Fusing self-regulated learning and formative assessment: a roadmap of where we are, how we got here, and where we are going. *The Australian Educational Researcher*, 45, 13-31.
- Popham, W. J. (2008). *Transformative assessment*. Alexandria, VA: ASCD.
- Robinson, V. M. J., Lloyd, C. & Rowe, K. (2008). [*The impact of leadership on student outcomes. An analysis of the differential effects of leadership types*](#). *Educational Administration Quarterly*, 44(5), 635-674.
- Rosenthal, R. & Jacobson, L. (1968). *Pygmalion in the classroom*. New York: Holt, Rinehart & Winston.
- Ruiz-Primo, M. A. (2011). Informal formative assessment: The role of instructional dialogues in assessing students' learning. *Studies in Educational Evaluation*, 37, 15-24.
- Sadler, R. D. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.
- Schildkamp, K., Heitink, M., van der Kleij, F., Hoogland, I., Dijkstra, A., Kippers, W., & Veldkamp, B. (2014). [*Voorwaarden voor effectieve formatieve toetsing. Een praktische review*](#). Den Haag: NRO.
- Seidel, T., Rimmel, R., & Prenzel, M. (2005). Clarity and coherence of lesson goals as a scaffold for student learning. *Learning and Instruction*, 15, 539-556.
- Sherrington, T., & Caviglioli, O. (2020). *Teaching Walkthru's. Five-step guides to instructional coaching*. Melton, Woodbridge: John Catt Educational Ltd.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153-189.
- Sleenhof, J. P. W., Koopman, M., Thurlings, M. C. G., & Beijard, D. (2019). An exploratory study into teachers' beliefs and experiences about allocating students. *Teaching and Teacher Education*, 80, 94-105.
- Sluismans, D. M. A., Joosten-ten Brinke, D., & Van der Vleuten, C. P. M. (2013). [*Toetsen met leerwaarde. Een reviewstudie naar effectieve kenmerken van formatief toetsen*](#). Den Haag: NWO.
- Sluismans, D. M. A., & Kneyber, R. (2016). [*Toetsrevolutie. Naar een feedbackcultuur in het voortgezet onderwijs*](#). Culemborg: Uitgeverij Phronese.
- Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative review. *Perspectives on Psychological Science*, 10(2) 176-199.
- Surma, T., Vanhoyweghen, K., Sluismans, D. M. A., Camp, G., Muijs, D., & Kirschner, P. (2019). [*Wijze lessen. Twaalf didactische bouwstenen voor effectieve didactiek*](#). Meppel: Ten Brink Uitgevers.

- Thum, Y. M., Tarasawa, B., Hegedus, A., Yun, X., & Bowe, B. (2015). *Keeping Learning on Track. A case-study of formative assessment practice and its impact on learning in meridian school district*. Portland, OR: Northwest Evaluation Association.
- Vleuten, C. P. M van der., Schuwirth, L., Driessen, E., Govaerts, M. J. B., & Heeneman, S. (2015). *Twelve tips for programmatic assessment*. *Medical Teacher*, 37, 641-664.
- Weinstein, Y., & Sumeracki, M. (2019). *Understanding how we learn. A visual guide*. Milton Park, Abingdon, Oxon: Routledge.
- Wiggins, G., & McTighe, J. (2006). *Understanding by design* (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Wiliam, D. (2006). *Formative assessment: Getting the focus right*. *Educational Assessment*, 11(3), 283-289.
- Wiliam, D. (2011). *Embedded Formative Assessment*. Bloomington: Solution Tree Press.
- Wiliam, D. (2013). *Principled curriculum design*. London: SSAT Limited.
- Wiliam, D. (2016). *Leadership for teacher learning*. Bristol: Learning Science Ltd.
- Wiliam, D., Lee, C., Harrison, C., & Black, P. J. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy & Practice*, 11(1), 49–65.
- Willingham, W. W., Pollack, J. M., & Lewis, C. (2002). Grades and test scores: Accounting for observed differences. *Journal of Educational Measurement*, 39(1), 1–37.
- Winstone, N. E., & Boud, D. (2020). The need to disentangle assessment and feedback in higher education. *Studies in Higher Education*. <https://doi.org/10.1080/03075079.2020.1779687>
- Wylie, C. E., & Lyon, C. J. (2015) The fidelity of formative assessment implementation: issues of breadth and quality. *Assessment in Education: Principles, Policy & Practice*, 22(1), 140-160.