



# **EVALUATIE & DIVERSITEIT IN HET BASIS- EN SECUNDAIR ONDERWIJS**

## Eindrapport

**Joke Ysenbaert, Piet Van Avermaet & Mieke Van Houtte**



# **EVALUATIE & DIVERSITEIT IN HET BASIS- EN SECUNDAIR ONDERWIJS**

## **Eindrapport**

**Joke Ysenbaert, Piet Van Avermaet & Mieke Van Houtte**

**Promotor: Prof. Dr. Piet Van Avermaet**

**Copromotoren: Prof. Dr. Kris Van den Branden & Prof. Dr. Mieke  
Van Houtte**

Research paper SONO/2020.OL1.6/2

Gent, juni 2020

Het Steunpunt Onderwijsonderzoek is een samenwerkingsverband van UGent, KU Leuven, VUB, UA en ArteveldeHogeschool.



Gelieve naar deze publicatie te verwijzen als volgt:

Ysenbaert, J., Van Avermaet, P., & Van Houtte, M. (2020). Evaluatie & diversiteit in het basis- en secundair onderwijs: Eindrapport.

Voor meer informatie over deze publicatie [Joke.Ysenbaert@UGent.be](mailto:Joke.Ysenbaert@UGent.be)

Deze publicatie kwam tot stand met de steun van de Vlaamse Gemeenschap, Ministerie voor Onderwijs en Vorming.

In deze publicatie wordt de mening van de auteur weergegeven en niet die van de Vlaamse overheid. De Vlaamse overheid is niet aansprakelijk voor het gebruik dat kan worden gemaakt van de opgenomen gegevens.

© 2020 STEUNPUNT ONDERWIJSONDERZOEK

p.a. Coördinatie Steunpunt Onderwijsonderzoek  
UGent - Vakgroep Onderwijskunde  
Henri Dunantlaan 2, BE 9000 Gent

Deze publicatie is ook beschikbaar via [www.steunpuntsono.be](http://www.steunpuntsono.be)

# Voorwoord

Dit rapport is het sluitstuk van een onderzoek dat wordt uitgevoerd binnen het Steunpunt voor Onderwijsonderzoek (SONO). SONO heeft als doel om gedurende de periode 2016-2020 beleidsvoorbereidend onderzoek te voeren rond verschillende thema's die door de Vlaamse Overheid als prioritair en relevant beschouwd worden. Eén van de onderzoekslijnen (OL. 1.6) focust op het evaluatiebeleid van scholen en de evaluatiepraktijken van leerkrachten in het basis- en secundair onderwijs. Hierbij wordt nagegaan in welke mate het evaluatiebeleid en de - praktijken van scholen rekening houden met diversiteit onder leerlingen.

Binnen deze onderzoekslijn werden reeds drie deelrapporten opgeleverd: de literatuurstudie (Ysenbaert, Van Avermaet, & Van Houtte, 2017), het rapport met de resultaten van het casestudieonderzoek (Ysenbaert, Van Avermaet, & Van Houtte, 2018) en een rapport met de bevindingen van het surveyonderzoek (Ysenbaert, Van Avermaet, & Van Houtte, n.d.).

Het voorliggend document vormt het eindrapport waarin de bevindingen van de drie deelstudies getrianguleerd worden. In het eerste hoofdstuk wordt de doelstelling en het opzet van het onderzoek kort geschetst. Het tweede hoofdstuk geeft de fasering weer van het onderzoek en de methodologie die hierbij gehanteerd werd. In een derde hoofdstuk worden de resultaten toegelicht die voortvloeien uit de triangulatie van de drie deelstudies. Dit hoofdstuk wordt afgerond met een samenvatting aan de hand van de centrale onderzoeksvragen. In de discussie, hoofdstuk IV, worden de bevindingen teruggekoppeld naar de opzet van dit onderzoek. Het discussieluik vormt de aanzet tot het formuleren van de beleidsaanbevelingen die na te lezen zijn in het vijfde en laatste hoofdstuk van dit rapport.

Voorafgaand aan de verschillende hoofdstukken, geven we eerst een beleidssamenvatting.

# Beleidssamenvatting

De evolutie naar een steeds diverser wordende leerlingenpopulatie doet zich niet enkel voor in de Vlaamse context, maar is een wereldwijd fenomeen (Banks & al., 2007). Dit uit zich in doelstellingen van internationale samenwerkingsverbanden. Zo stelt het Verdrag van de Verenigde Naties (2006), inclusief onderwijs als doel voorop om het recht op kwaliteitsvol en toegankelijk onderwijs wereldwijd te garanderen. Dit betekent dat België sinds de ratificatie in 2009 moet kunnen aantonen op welke wijze ingezet wordt op het realiseren van zo kwaliteitsvol mogelijk onderwijs voor alle leerlingen. De definiëring van inclusief onderwijs is daarenboven door de jaren heen getransformeerd (Operti, Walker, & Zhang, 2013). In de huidige benadering van het VN-verdrag ligt de klemtoon op het zodanig inrichten van het onderwijssysteem opdat álle leerlingen bereikt worden, ongeacht hun specifieke achtergrond of eigenschap(en) waarmee ze aan een (doel)groep gelinkt worden. Het evaluatiebeleid van scholen en de evaluatiepraktijken van leerkrachten spelen een cruciale rol in het realiseren van kwaliteitsvol onderwijs voor álle leerlingen, zonder hierbij eng te focussen op ‘risico-leerlingen’ of leerlingen met specifieke onderwijsbehoeften. Dit onderzoek bestudeert dan ook het evaluatiebeleid van scholen en de evaluatiepraktijken van leerkrachten in het licht van deze (inter)nationale ontwikkelingen richting inclusief onderwijs.

Voor de aanpak werd geopteerd voor een mixed-method design zodat de uiteindelijke combinatie van data complementair is in sterktes en elkaar niet overlapt in zwaktes (Johnson & Turner, 2003). Het mixed method-design omvat drie fases die elkaar opeenvolgend in de periode september 2016 tot en met juni 2020. Op basis van de literatuurstudie werd een referentiekader ontwikkeld van waaruit het huidige evaluatiebeleid van scholen en de bijhorende praktijken verklaard kunnen worden. Via het casestudieonderzoek werden zes basisscholen en zes secundaire scholen bestudeerd om zicht te krijgen op hun evaluatiebeleid en evaluatiepraktijken. Via interviews en observaties werd nagegaan welke hierbij de bepalende factoren zijn en welke opvattingen betrokken actoren hebben ten aanzien van evaluatie. Het surveyonderzoek, waaraan 56 scholen deelnamen, laat tenslotte toe om te bepalen of de tendensen die vastgesteld werden in het casestudieonderzoek al dan niet veralgemeend kunnen worden.

Op basis van de resultaten wordt vastgesteld dat het evaluatiebeleid en de evaluatiepraktijken in zowel basis- als secundair onderwijs in Vlaanderen kansen onbenut en/of onderbenut laat met betrekking tot het maximaliseren van leer- en ontwikkelingskansen van álle leerlingen. Om te beginnen worden de evaluatiepraktijken van leerkrachten in zowel basis- als secundair onderwijs gekenmerkt door een sterke scheiding tussen evaluatie en de algemene onderwijspraktijk. Een sterke scheiding tussen evaluatie- en instructieactiviteiten belemmert de ondersteunende functie van evaluatie aangezien de uitkomsten van de evaluatie nauwelijks gebruikt worden om het leerproces en/of het onderwijsproces bij te sturen. De nadruk ligt hierbij sterk op ‘evalueren van het leren’ in plaats van ‘evalueren om te leren’ waardoor de leer- en ontwikkelingskansen van leerlingen eerder ingeperkt worden. Bovendien zullen leerkrachten wiens evaluatiepraktijk aanleunt bij het gescheiden model over het algemeen minder inzetten op differentiatie en zullen leerlingen weinig betrokken worden bij het evaluatieproces. Leerlingen zijn echter vragende partij voor een meer geïntegreerde evaluatiepraktijk omdat dit hen

beter ondersteunt en vooruit helpt in hun leerproces. Zo hebben ze bijvoorbeeld een expliciete voorkeur voor observatie door de leerkracht omdat dit hen meer directe feedback oplevert.

Dit onderzoek stelde bovendien vast dat waar transitiemomenten vaker aan de orde zijn, er vaker geëvalueerd wordt volgens het gescheiden model. Dit is bijvoorbeeld meer het geval in het secundair onderwijs in vergelijking met het lager onderwijs en binnen het secundair onderwijs wordt in het bso dan weer vaker geëvalueerd volgens het geïntegreerde model. We verklaren dit aan de hand van het waterval-denken dat bepalend is bij het nemen van beslissingen over studiekeuze. Binnen het bso is er bijvoorbeeld nog maar weinig sprake van 'transitie' aangezien dit het 'laagste' niveau vormt binnen het waterval-denken. De oriëntering van leerlingen wordt sterk gestuurd door het waterval-denken waarbij eerder de cognitieve prestaties een rol spelen in plaats van de talenten en interesses van leerlingen. Dit zorgt er niet alleen voor dat de algemene evaluatiepraktijk meer gekenmerkt wordt door een gescheiden evaluatiepraktijk ten nadele van een meer geïntegreerde evaluatiepraktijk, maar ook dat studiekeuze gestuurd wordt door deficit-denken. Oriëntering komt op deze manier pas aan de orde als *het niet meer lukt in de huidige onderwijsvorm*. Een positieve, talentgerichte benadering waarbij op zoek gegaan wordt naar het potentieel van de leerling, komt daarentegen beter tegemoet aan het creëren van maximale leer- en ontwikkelingskansen.

De vaststelling dat de ondersteunende functie van evaluatie eerder beperkte aandacht krijgt en dat de functie van het nemen van beslissingen eerder bepaald wordt door de structuur van het onderwijs dan door de talenten van de leerling betekent dat er heel wat kansen onbenut blijven om maximale leer- en ontwikkelingskansen te realiseren voor alle leerlingen. We kunnen dit verklaren door een zwak evaluatiebeleid op schoolniveau. Dit onderzoek stelt vast dat evaluatie heel vaak een zaak is van de individuele leerkracht en dat er op schoolniveau nauwelijks sprake is van een krachtig evaluatiebeleid, van gedeelde opvattingen omtrent evaluatie of gedeelde evaluatiepraktijken van leerkrachten binnen eenzelfde school. Hoewel scholen op dit vlak over een ruime autonomie beschikken, blijkt men deze autonomie niet aan te wenden om een krachtig evaluatiebeleid uit te bouwen. Deze onderbenutte of gemiste kansen zijn het gevolg van een niet-bewustzijn van en/of een gebrekkige aandacht voor het feit dat evaluatie eigenlijk uitermate essentieel is voor het realiseren van maximale leer- en ontwikkelingskansen voor alle leerlingen.

In wat volgt, worden beleidsaanbevelingen geformuleerd die bijdragen tot de realisatie van een krachtig evaluatiebeleid dat de leer- en ontwikkelingskansen van alle leerlingen maximaliseert. In overeenstemming met de literatuur, die stelt dat evaluatiepraktijken het resultaat zijn van een complexe wisselwerking tussen factoren op verschillende niveaus (Ercikan & Solano-Flores, 2016), voelen we ons genooddaakt om van start te gaan met één algemene, overkoepelende aanbeveling. Deze richt zich tot alle niveaus en betrokken actoren die (on)rechtstreeks invloed uitoefenen op het evaluatiebeleid van scholen en de evaluatiepraktijken van leerkrachten. Deze aanbeveling vormt als het ware het basisprincipe van waaruit de verschillende actoren best handelen bij het opzetten en ontwikkelen van initiatieven die gerelateerd zijn aan evaluatie.

Daarnaast formuleren we specifieke aanbevelingen ten aanzien van de overheid, de pedagogische begeleidingsdiensten, de lerarenopleidingen, de schoolleiders en de leerkrachten. We willen benadrukken dat de aanbevelingen voor leerkrachten bewust als laatste in de rij werden opgenomen. In navolging van de onderzoeksbevindingen willen we benadrukken dat de grootste uitdagingen zich situeren op de overige beleidsniveaus. De oplossing om tegemoet te komen aan het probleem dat

evaluatie nu te veel een zaak van de individuele leerkracht is, zal dan ook vanuit die andere beleidsniveaus geïnitieerd moeten worden.

**0. Algemene aanbeveling: Stel maximale leer- en ontwikkelingskansen centraal bij elk initiatief dat aan evaluatie raakt.**

Het staat buiten kijf dat een diverse leerlingenpopulatie vandaag de dag realiteit is in Vlaamse scholen en klassen. Dit uit zich bovendien in doelstellingen van internationale samenwerkingsverbanden, zoals bijvoorbeeld het Verdrag van de Verenigde Naties (2006) dat inclusief onderwijs als doel voorop stelt. Dit verdrag werd door België geratificeerd in 2009 wat betekent dat België sindsdien moet kunnen aantonen op welke wijze werk wordt gemaakt van het realiseren van zo kwaliteitsvol mogelijk onderwijs voor alle leerlingen. De meest recente definiëring van inclusief onderwijs (Opertti et al., 2013) legt bovendien de nadruk op een systeembenadering. Deze benadering doelt op het zodanig inrichten van het onderwijssysteem zodat alle leerlingen bereikt worden, ongeacht hun specifieke achtergrond of kenmerken. De idee van een gecompartmenteerde doelgroepenbenadering waarbij specifieke maatregelen nodig zijn voor specifieke (groepen van) leerlingen wordt hier opgeheven doordat het systeem gelijke kansen op kwaliteitsvol onderwijs creëert voor alle leerlingen, ook voor deze leerlingen die momenteel niet 'gelabeld' worden als leerlingen met specifieke onderwijsbehoeften.

Leerlingenevaluatie biedt heel wat mogelijkheden om eenieders leer- en ontwikkelingskansen te maximaliseren. De resultaten van dit onderzoek tonen helaas dat het huidige evaluatiebeleid en de evaluatiepraktijken de kansen op kwaliteitsvol onderwijs voor alle leerlingen eerder dreigt te belemmeren dan te bevorderen. We besluiten dat het gebruikelijke concept van evaluatie én de uitwerking ervan in de praktijk grondige bijstellingen vereist. We schuiven daarom volgend principe naar voren: evaluatie moet in de eerste plaats het leren en ontwikkelen van alle leerlingen dienen. Dit uitgangspunt maximaliseert de leer- en ontwikkelingskansen van elke leerling. Het is cruciaal dat elk van de drie hieronder beschreven functies van evaluatie vanuit dit idee vorm en invulling krijgen. Zoniet, dreigt het streven om via evaluatie kansen te bevorderen alsnog te resulteren in het limiteren van kansen. Zo verdwijnen bijvoorbeeld dagdagelijkse brede vormen van evaluatie in het niets als uiteindelijk de studiekeuze bepaald wordt door het resultaat op één examen van wiskunde.

We lichten enkele aandachtspunten bij elk van deze functies toe.

1. Informerende functie

- Evaluatie moet erop gericht zijn om het profiel van de leerling zo breed mogelijk in kaart te brengen. Specifiek houdt dit in dat er naast de cognitieve aspecten, ook aandacht gaat naar affectieve, motorische, motivationele en metacognitieve aspecten.
- Het brede profiel van de leerling besteedt niet alleen aandacht aan de actuele stand van zaken in het leerproces en de ontwikkeling, maar biedt ook informatie over de gerealiseerde groei in relatie tot vooropgestelde doelen.
- De informatie wordt verkregen op basis van een brede waaier aan evaluatie-instrumenten en evaluatie-vaardigheden. We stellen een weloverwogen mix voor van observaties, leerlingcontacten, zelf- en peerevaluatie, opdrachten, interacties, vraaggesprekken, toetsen, examens en gestandaardiseerde evaluatievormen. Binnen deze ruime waaier wordt best gewicht toegekend aan elke evaluatievorm en -tool.

## 2. Ondersteunende functie

- Evalueren en leren zijn onlosmakelijk met elkaar verboden. Idealiter maakt deze functie bijgevolg integraal deel uit van de dagdagelijkse didactiek door evaluatieactiviteiten sterk in te bedden in de algemene onderwijspraktijk.
- De output van evaluatie wordt maximaal aangewend als input waarmee de onderwijspraktijk afgestemd wordt op de leer- en ontwikkelingsnoden van leerlingen (= output=input-principe). Dit resulteert in een onderwijspraktijk waarbij evaluatie het beginpunt vormt van de didactische cyclus.
- Het output=input-principe wordt nog krachtiger indien dit zich ook vertaalt in concrete handvatten voor de leerling zelf. Deze kunnen aangereikt worden via kwaliteitsvolle feedback die hen helpt om de vooropgestelde doelen te bereiken.

## 3. Beslissingnemende functie

- Beslissingen over de schoolloopbaan van leerlingen moeten genomen worden op basis van *faire/billijke* evaluaties die aan de beslissingen voorafgaan.
- Het nemen van beslissingen moet beschouwd worden als een proces in plaats van als een moment.
- Het nemen van beslissingen over de schoolloopbaan is een gedeelde verantwoordelijkheid en daarom verloopt dit proces best in partnerschap met collega-leerkrachten, ouders én de leerlingen zelf.

De aanbevelingen die hierna weergegeven worden, richten zich naar specifieke actoren. De aanbevelingen hebben als doel om het huidige evaluatiebeleid en de evaluatiepraktijken zo bij te sturen zodat ze bijdragen aan de realisatie van de hierboven omschreven algemene aanbeveling. Belangrijk is dat elke actor hierin actief initiatief neemt opdat er op coherente wijze gewerkt kan worden aan een evaluatiebeleid dat de ontwikkeling naar een inclusiever onderwijs bevordert.

### **1. Aanbevelingen ten aanzien van de overheid**

Onderwijsvernieuwingen maken volgens Hopkins (2003) pas kans op slagen indien beleidsinitiatieven aandacht hebben voor zowel een systeem-brede als een systeem-diepe aanpak. De systeem-brede aanpak verwijst naar het voeren van een beleid dat coherent is en ook als coherent gepercipieerd wordt door de betrokken actoren. We adviseren om de plannen in de beleidsnota 2019-2024 om gevalideerde, gestandaardiseerde en genormeerde toetsen te introduceren best te kaderen binnen eerder genomen beleidsinitiatieven die de ontwikkeling naar meer inclusief onderwijs stimuleren. Hierbij is het van belang om de actuele invulling van 'inclusie' voor ogen te houden, namelijk de brede benadering waarbij het onderwijssysteem zo ingericht wordt opdat het voor alle leerlingen mogelijk wordt om hun potentieel ten volle te kunnen ontplooien en ontwikkelen, met oog op het minimaal bereiken van de eindtermen. Daarnaast heeft de overheid ook een rol om haar beleid vertaald te krijgen in de praktijk. Dit kan bevorderd worden door aandacht te besteden aan een systeem-diepe aanpak. Voor zowel de systeem-brede als de systeem-diepe aanpak formuleren we enkele concrete aanbevelingen.



### **Aanbeveling 1.1. Een systeem-brede aanpak realiseren via het bedachtzaam introduceren van gestandaardiseerde toetsen**

Hoewel gestandaardiseerde toetsen mogelijkheden bieden in het streven naar kwaliteitsvol onderwijs, pleiten we voor een bedachtzame introductie ervan. De resultaten van dit onderzoek, en meer specifiek de opvattingen van leerkrachten, lieten zien dat dergelijke toetsen soms als selectie-instrument dreigen aangewend te worden. De implementatie van gevalideerde, gestandaardiseerde en genormeerde toetsen wordt daarom best niet uitsluitend gekaderd binnen de internationale trend richting meer centrale toetsing om de onderwijskwaliteit op te volgen. We adviseren om bij deze implementatie eveneens te waken over het realiseren van de ratificatie van het Verdrag van de Verenigde Naties waartoe Vlaanderen zich sinds 2009 verbindt. Dit bevordert de systeem-brede aanpak – ook in de perceptie van de overige actoren in het onderwijsveld - aangezien dit aansluit bij bijvoorbeeld het M-decreet. Het streefdoel van inclusief onderwijs, dat maximale leer- en ontwikkelingskansen bij álle leerlingen beoogt, dient bijgevolg het uitgangspunt te vormen bij de implementatie van gevalideerde, gestandaardiseerde en genormeerde toetsen. De overheid wordt geadviseerd om hierbij concreet rekening te houden met volgende aandachtspunten:

- De implementatie van gevalideerde, gestandaardiseerde en genormeerde toetsen moet eerder gericht zijn op ontwikkeling van de school, de leerkracht en de leerling dan op verantwoording. We verwijzen hiervoor naar de *good practice* van case D in het basisonderwijs waarbij er op schoolniveau regelmatig gestandaardiseerde taaltoetsen afgenomen worden. De resultaten van deze toetsen worden uitsluitend aangewend voor het opvolgen en verbeteren van het taalonderwijs op school en het nemen van specifieke beleidskeuzes om de taalvaardigheid van leerlingen te bevorderen. Hier gaat het om gestandaardiseerde toetsen die dus vooral een reflexieve functie hebben. Het leidde concreet tot beslissingen op vlak van:
  - o Onderwijskundig beleid: taal werd niet langer als apart vak beschouwd, maar werd geïntegreerd binnen wereldoriëntatie.
  - o Professionaliseringsbeleid: leerkrachten lager onderwijs werden in een traject begeleid in hoe ze taalvaardigheid van leerlingen konden stimuleren en evalueren. Later werden ook leerkrachten van het kleuteronderwijs hierin geprofessionaliseerd.
  - o Beleid inzake kwaliteitsontwikkeling: de gestandaardiseerde toetsen worden regelmatig afgenomen en de evolutie van de resultaten wordt in verband gebracht met genomen acties van de school om na te gaan in welke mate deze de taalvaardigheid bevorderen/belemmeren.
- Verantwoording met betrekking tot de output op deze toetsen is uitermate onwenselijk. Mogelijke verantwoording dient zich te richten op de kwaliteitsontwikkeling van de school als organisatie. Anders gezegd: door scholen verantwoordelijk te houden voor de mate waarin men op basis van de resultaten op deze toetsen geïnformeerde beleidskeuzes maakt, worden scholen gestimuleerd om te ontwikkelen naar kwaliteitsvoller onderwijs ten aanzien van alle leerlingen.
- Een breed instrumentarium aan gevalideerde, gestandaardiseerde en genormeerde toetsen dat recht doet aan de ruime waaier aan na te streven leer- en ontwikkelingsdoelen is nodig.
- Scholen moeten hun beleid inzake kwaliteitsontwikkeling kunnen afstemmen op de contextspecifieke noden. Hiervoor moeten ze over een zekere mate van keuzevrijheid kunnen beschikken inzake de zelfevaluatie-instrumenten die ze hiervoor aanwenden.

**Aanbeveling 1.2. Een systeem-brede aanpak realiseren door de autonomie van scholen te beklemtonen in het realiseren van een evaluatiebeleid dat maximale leer- en ontwikkelingskansen bij álle leerlingen nastreeft.**

Om op schoolniveau een evaluatiebeleid te ontwikkelen dat kan vertrekken vanuit de eigenheid en diverse compositie van de school en dat maximaal de kansen van álle leerlingen bevordert, vormt de autonomie van de school een essentieel aspect. Aangezien leren en evalueren onlosmakelijk met elkaar verbonden zijn, moet een school de ruimte waarover ze beschikt om het evaluatiebeleid af te stemmen op het onderwijskundig beleid ook benutten. Een krachtig evaluatiebeleid sluit aan bij de ambitie om het leren van álle leerlingen maximaal te bevorderen. Daarenboven moeten scholen verantwoordelijk gehouden worden om hun autonomie zinvol aan te wenden om gericht in te spelen op de specifieke noden van de context van de school. Op die manier kan bijvoorbeeld een school met een groot aandeel leerlingen wiens thuistaal niet Nederlands is, opteren voor een frequent gebruik van verschillende evaluatievormen die de taalvaardigheid van leerlingen in kaart brengen om zo de evoluties te monitoren en de taaldidactiek hierop af te stemmen. De resultaten van het onderzoek lieten echter zien dat scholen over het algemeen heel wat kansen laten liggen om op basis van de ruime autonomie waarover men beschikt een schoolspecifiek evaluatiebeleid te ontwikkelen. We pleiten allerminst voor het vergroten van deze autonomie. We adviseren eerder om scholen verantwoordelijk te houden met betrekking tot de invulling van deze autonomie waarbij de nadruk ligt op het ontwikkelen van een evaluatiebeleid dat maximale leer- en ontwikkelingskansen nastreeft. Het raamwerk voor onderwijskwaliteit (Onderwijsinspectie, 2016) dat gehanteerd wordt bij schooldoorlichtingen legt hieromtrent de juiste accenten. Hoewel het casestudieonderzoek liet zien dat een doorlichting een stimulans vormt voor het invullen van deze autonomie, stellen we vast dat dit in de praktijk nog geen gemeengoed is.

**Aanbeveling 1.3. Een systeem-diepe aanpak realiseren door professionalisering te voorzien.**

We adviseren de overheid om sterk in te zetten op een *systeemdiepe* aanpak die ernaar streeft om het handelen van actoren op verschillende niveaus te stroomlijnen volgens dezelfde uitgangspunten. Een heldere communicatie vanuit de overheid waarbij de *systeembrede* aanpak geëxpliciteerd wordt, draagt hiertoe bij.

Verder is het van belang om professionalisering voor scholen mogelijk te maken opdat ze ondersteund worden in het ontwikkelen van een evaluatiebeleid dat maximale leer- en ontwikkelingskansen nastreeft om alle leerlingen vooruit te helpen richting de vooropgestelde doelen. De resultaten van het onderzoek laten namelijk een positieve samenhang zien tussen het deelnemen aan professionalisering en de mate waarin ingezet wordt op een geïntegreerde evaluatiepraktijk. Binnen dergelijke professionalisering is er naast specifieke aandacht voor de ontwikkeling en implementatie van het formele evaluatiebeleid en de evaluatiepraktijken ook aandacht voor de opvattingen omtrent evaluatie van betrokkenen.

De overheid wordt aanbevolen om hierbij professionaliseringsinitiatieven te voorzien die expliciet inzetten op het betrekken van verschillende actoren op school bij de ontwikkeling van dergelijk evaluatiebeleid. Steeds meer onderzoek toont namelijk aan dat gedeeld leiderschap bijdraagt tot organisatieperformantie en een positief effect heeft op leerlingresultaten (Harris & Spillane, 2008). Bovendien versterkt dit de ontwikkeling naar de school als een lerende organisatie die haar werking in

vraag stelt en zo gericht bijstuurt. Hiermee kan tegemoet gekomen worden aan het vastgestelde gebrek aan reflecterend vermogen.

Bovendien moet dergelijke professionalisering ook het beleidsvoerend vermogen van scholen versterken door te streven naar een geïntegreerd beleid. Op deze manier wordt een gefragmenteerd of gesegmenteerd beleid, zoals in het onderzoek werd vastgesteld, voorkomen. De overheid wordt aanbevolen om professionaliseringsinitiatieven te voorzien die scholen ondersteunen in het expliciteren van verbanden tussen het evaluatiebeleid en andere aspecten van het schoolbeleid zoals bijvoorbeeld het taalbeleid, zorgbeleid of het beleid inzake de interne kwaliteitsontwikkeling.

## **2. Aanbevelingen ten aanzien van pedagogische begeleidingsdiensten**

### **Aanbeveling 2.1. Een pro-actief professionaliseringsaanbod voorzien dat scholen ondersteunt in de ontwikkeling van een evaluatiebeleid dat maximale leer- en ontwikkelingskansen vooropstelt.**

De resultaten van zowel het casestudie- als het surveyonderzoek wijzen op een positieve samenhang tussen scholen die inzetten op professionalisering en sterk geïntegreerde evaluatiepraktijken. Deze professionalisering heeft specifieke aandacht voor zowel het formele evaluatiebeleid, de evaluatiepraktijken zelf en de opvattingen omtrent evaluatie van betrokkenen. De professionalisering beoogt hiermee het evaluatiebeleid en de evaluatiepraktijken zo te ontwikkelen dat elk van de drie functies van evaluatie maximale leer- en ontwikkelingskansen nastreeft bij alle leerlingen.

Begeleidingsdiensten en nascholingsinitiatieven wordt aanbevolen om hierbij specifiek in te zetten op gedeeld leiderschap. Het casestudieonderzoek laat zien dat het evaluatiebeleid op een duurzamere wijze vertaald wordt in de praktijk indien meerdere actoren voor langere tijd betrokken worden bij de ontwikkeling hiervan. Het betrekken van verschillende actoren op school bij de ontwikkeling van dergelijk evaluatiebeleid draagt niet alleen bij tot organisatieperformantie maar vertaalt zich ook in een positief effect op leerlingresultaten (Harris & Spillane, 2008).

### **Aanbeveling 2.2. Een pro-actief begeleidingsaanbod voorzien dat scholen specifiek ondersteunt in de ontwikkeling van een geïntegreerd evaluatiebeleid.**

Een evaluatiebeleid ontwikkelen is uitermate complex, precies omdat het op veel verschillende aspecten van het ruimere schoolbeleid betrekking heeft. Hoewel het referentiekader voor onderwijskwaliteit (Onderwijsinspectie, 2016) deze complexiteit weergeeft, zijn scholen zich niet altijd bewust van de afwezige of gebrekkige verbanden in het eigen evaluatiebeleid. De resultaten van dit onderzoek lieten zien hoe een krachtig evaluatiebeleid vaak afwezig is of sterk gefragmenteerd is. Pedagogisch begeleiders kunnen scholen pro-actief begeleiden om inzicht te verwerven in het eigen evaluatiebeleid om zo de leemtes te ontdekken en deze procesmatig aan te pakken. Zo kan er concreet gefocust worden op het verbinden van een evaluatiebeleid dat streeft naar het maximaliseren van leer- en ontwikkelingskansen en het zorgbeleid van de school, of het verband met het beleid inzake ouderbetrokkenheid, het professionaliseringsbeleid, het taalbeleid en of het beleid omtrent kwaliteitszorg.

### **Aanbeveling 2.3. Scholen pro-actief begeleiden in de ontwikkeling van een reflexieve cultuur onder andere door middel van gestandaardiseerde vormen van evaluatie.**

Een specifiek aspect van evaluatie omvat het aanwenden van evaluatie om te reflecteren over het beleid en de praktijken van de school, een cultuur die nog verre van gangbaar is in het Vlaamse onderwijs. De resultaten van het onderzoek laten zien hoe verplichte gestandaardiseerde vormen van

evaluatie nauwelijks voor reflexieve doeleinden worden gebruikt. Met het vooruitzicht op het introduceren van netoverstijgende, gestandaardiseerde toetsen, wordt het des te belangrijker dat scholen de juiste ondersteuning aangereikt krijgen om dergelijke evaluatieresultaten onder de loep te nemen. Scholen zullen begeleid moeten worden in het ontwikkelen van expertise die hen in staat stelt om gestandaardiseerde vormen van evaluatie aan te wenden voor het monitoren van:

- de ontwikkeling van alle leerlingen met het oog op het minimaal bereiken van de eindtermen,
- de professionele ontwikkeling van het team,
- de ontwikkeling van de school als organisatie.

Het zwaartepunt van dergelijke ondersteuning ligt niet noodzakelijk op het monitoren op zich, maar op het ontwikkelen van een reflexieve cultuur. Dit stelt een school in staat om geïnformeerde beleidskeuzes te maken waaraan doelgerichte acties gekoppeld worden. Dergelijke cultuur is niet alleen noodzakelijk om het leren van leerlingen naar een hoger niveau te tillen, maar ook om als school te evolueren richting een lerende organisatie die zichzelf voortdurend bijstuurt.

### **3. Aanbevelingen ten aanzien van lerarenopleidingen**

#### **Aanbeveling 3.1. De kaart van diversiteit trekken**

De onderzoeksbevindingen lieten zien dat de evaluatiepraktijk gekenmerkt wordt door een onderbenutting van de formatieve functie die evaluatie heeft. Daarenboven bleek ook dat de summatieve functie sterk aangestuurd wordt door het waterval-denken. Beide vaststellingen samen vormen een hindernis voor het creëren van maximale leer- en ontwikkelingskansen voor alle leerlingen. Meer zelfs: ze staan een evolutie richting meer inclusief onderwijs in de weg. In navolging van de evolutie naar meer inclusief onderwijs, wordt lerarenopleidingen daarom aanbevolen om de competenties inzake evaluatie aan te brengen via het trekken van de kaart van diversiteit.

Bij de competentieontwikkeling van student-leraren inzake evaluatie dient erover gewaakt te worden dat bij elk van de drie functies van evaluatie de maximale leer- en ontwikkelingskansen van alle leerlingen het streefdoel zijn. Dit vertaalt zich idealiter in een accentverschuiving in het curriculum waarbij uitgegaan wordt van de idee dat evaluatie het beginpunt vormt in het didactisch proces en niet het eindpunt. Het aspect 'evaluatie' vroeg introduceren in het curriculum kan hier al toe bijdragen. Aangezien evalueren en leren onlosmakelijk met elkaar verbonden zijn, is het niet meer dan logisch dat evaluatie van bij het begin meegenomen wordt in de didactische leerlijn van leraren-in-opleiding. We pleiten voor een curriculum waarin evaluatie geen op zichzelf staand thema vormt, maar steeds terugkeert doorheen het curriculum waarbij telkens de verschillende functies van evaluatie geëxpliciteerd worden. Een curriculum van de lerarenopleiding waarbij evaluatie voortdurend aan de orde is, sluit bovendien aan bij het referentiekader voor onderwijskwaliteit (Onderwijsinspectie, 2016) waarbij evaluatie eveneens opduikt in meerdere kwaliteitsrubrieken.

Een dergelijke benadering waarbij de rol van evaluatie als dé ingang tot differentiatiepraktijken reeds vroeg beklemtoond wordt, zal student-leraren vaardiger maken in het ontwikkelen van een 'diversiteits-bril'. De onderzoeksresultaten wezen op dit vlak op een positief verband tussen het kunnen opmerken van diversiteit enerzijds en het afstemmen van zowel instructie- als evaluatiepraktijken op deze diversiteit. Via de bekwaamheid om diversiteit onder leerlingen op te

merken, voelen leerkrachten zich competenter om hun onderwijs- en evaluatiepraktijken af te stemmen op de diversiteit onder leerlingen en in te zetten op hun talentontwikkeling.

### **Aanbeveling 3.2. Sterker inzetten op praktijkgerichte oefenkansen inzake evaluatie**

Het vroeg introduceren van 'evaluatie' in het curriculum en het veelvuldig laten terugkeren hiervan, zal niet volstaan om de evaluatiecompetentie van toekomstige leraren ten volle te ontwikkelen. Leerkrachten geven aan dat ze op dit vlak het meest geleerd hebben via de praktijkervaring die ze hieromtrent opdeden. Lerarenopleidingen wordt aanbevolen om deze verhoogde aandacht ook door te trekken naar praktijkgerichte opdrachten en stage-onderdelen, precies om toekomstige leraren al zoveel mogelijk kansen te geven tot praktijkervaring. Dit zal studenten alleen maar vaardiger maken in hun competentie om hun evaluatiepraktijk zo in te richten opdat de leer- en ontwikkelingskansen van leerlingen gemaximaliseerd worden. Hierbij is een sterke focus nodig op het ontwikkelen van de vaardigheid om evaluatiepraktijken sterk in te bedden in de algemene instructieactiviteiten (Cfr. geïntegreerde evaluatiepraktijk). Daarnaast zijn oefenkansen nodig om de vaardigheden inzake de informerende, ondersteunende en de beslissingnemende functie van evaluatie aan te scherpen.

## **4. Aanbevelingen ten aanzien van schoolleiders**

### **Aanbeveling 4.1. Een evaluatiebeleid dat streeft naar het maximaliseren van ieders leer- en ontwikkelingskansen.**

Een evaluatiebeleid waarin diversiteit een centrale plaats heeft, maximaliseert de leer- en ontwikkelingskansen van elke leerling. Het is hierbij cruciaal dat elk van de drie functies van evaluatie vanuit dit idee vorm en invulling krijgen (zie algemene aanbeveling). Zoniet, dreigt er een risico op het limiteren van de ontwikkelingskansen van (groepen van) leerlingen. Dit onderzoek stelde namelijk vast dat de evaluatiepraktijk gekenmerkt wordt door een overwegend gescheiden evaluatiepraktijk, wat veeleer een drempel vormt in het creëren van maximale leer- en ontwikkelingskansen voor alle leerlingen.

Via het centrale doel om bij het evaluatiebeleid maximale leer- en ontwikkelingskansen voor elke leerling steeds voor ogen te houden beogen we twee verschuivingen in de huidige evaluatiepraktijken van scholen. Enerzijds is er nood aan een sterkere klemtoon op de formatieve functie van evaluatie in plaats van de summatieve functie. Uit het onderzoek blijkt dit een noodzakelijke verschuiving aangezien het onderbenutten van de formatieve functie eerder de leer- en ontwikkelingskansen belemmert in plaats van bevordert. Via het streven naar maximale leer- en ontwikkelingskansen wordt eveneens een verschuiving beoogt met betrekking tot de summatieve functie van evaluatie. Dit houdt in dat de oriëntering van leerlingen niet langer uitgaat van tekorten bij leerlingen, maar veeleer aangestuurd wordt vanuit de vraag welke de talenten zijn van leerlingen en waar deze het best tot uiting kunnen komen. Op deze manier behoedt het evaluatiebeleid zich voor een oriënteringspraktijk die gebaseerd is op het watervaldenken en waarbinnen evaluatie een selecterende functie heeft.

### **Aanbeveling 4.2. Een sterk geïntegreerd evaluatiebeleid nastreven vanuit een reflexieve cultuur**

Om tot een sterk evaluatiebeleid te komen is het van belang om vanuit het ruimere schoolbeleid expliciet stil te staan bij elk van de drie functies van evaluatie. Dit draagt bij tot een sterkere integratie van het evaluatiebeleid en voorkomt een gesegmenteerd evaluatiebeleid wat nu doorgaans te vaak het geval is volgens de resultaten van het onderzoek. De verschillende functies krijgen bij voorkeur een

plaats binnen het evaluatiebeleid waarbij de onderlinge samenhang met de overige beleidsdomeinen bewaakt wordt.

Zo haakt een evaluatiebeleid logischerwijs in op het *onderwijsbeleid* van de school aangezien evalueren en leren onlosmakelijk met elkaar verbonden zijn. Zowel het onderwijs- als het evaluatiebeleid van een school zijn gelinkt aan het *zorgbeleid* net omdat via evaluatie de noden en leerbehoeften van leerlingen vastgesteld worden. Ook vanuit het *communicatiebeleid* van een school kan het evaluatiebeleid verder ingevuld worden. Vragen zoals waarom en hoe men leerlingen, ouders en collega's wil informeren inzake de aanpak en de uitkomsten van evaluatie kunnen van hieruit beantwoord worden.

Verder is er ook een samenhang tussen het evaluatiebeleid en het *kwaliteitsbeleid* van de school. Dit kwaliteitsbeleid krijgt idealiter vorm via een reflexieve cultuur waarbij bijvoorbeeld gestandaardiseerde vormen van evaluatie aangewend worden om de ontwikkeling te bevorderen van leerlingen, leerkrachten en de school als organisatie. Zelfevaluatie op organisatieniveau biedt zo informatie die kan aangewend worden om aan kwaliteitsontwikkeling te doen. Tegelijkertijd haakt dit mogelijks in op het *professionaliseringsbeleid* van de school om zo kwaliteitsvol onderwijs aan alle leerlingen te realiseren. Een reflexieve cultuur zal op termijn niet alleen bijdragen tot een meer geïntegreerd evaluatiebeleid, het stimuleert tegelijkertijd de school om te evolueren richting een lerende organisatie. Ten slotte willen we er op wijzen dat het referentiekader voor Onderwijskwaliteit (Onderwijsinspectie, 2016) hierbij een bruikbaar raamwerk vormt. Het feit dat evaluatie via verschillende kwaliteitsverwachtingen in elk van de vier rubrieken van het referentiekader aan bod komt, draagt bij tot de uitwerking van een evaluatiebeleid dat aandacht besteedt aan enerzijds de verschillende functies van evaluatie en anderzijds de expliciete verbanden met verschillende beleidsaspecten van het ruimere schoolbeleid.

#### **Aanbeveling 4.3. Faciliterende randvoorwaarden scheppen om het evaluatiebeleid te vertalen naar een afdoende evaluatiepraktijk**

Enkele randvoorwaarden faciliteren de vertaalslag van een evaluatiebeleid naar een gedragen beleid dat ook in de praktijk vertaald wordt. Bij dit alles is er een cruciale rol weggelegd voor de directeur die hierin een katalysatorfunctie heeft, zo bleek uit het casestudieonderzoek.

Een eerste randvoorwaarde is om de ontwikkeling van het evaluatiebeleid te realiseren via gedeeld leiderschap. Steeds meer onderzoek toont namelijk aan dat gedeeld leiderschap bijdraagt tot organisatieperformantie wat tevens resulteert in betere leerlingresultaten (Harris & Spillane, 2008). Het casestudieonderzoek omvat *good practices* waarbij via een kernteam een gedragen evaluatiebeleid geïnitieerd wordt. Belangrijk hierbij is dat een visie en een beleid omtrent evaluatie vakoverschrijdend ontwikkeld wordt, om pas in een volgende stap de concrete uitwerking te vertalen binnen een specifiek vak.

Naast het ontwikkelen in gedeeld leiderschap, wijst dit onderzoek ook op de rol van externe professionalisering. Deze kan betrekking hebben op het verruimen van kennis inzake kwaliteitsvolle evaluatie in het algemeen, het procesmatig ontwikkelen van een evaluatiebeleid, of het aanscherpen van specifieke evaluatievaardigheden van leerkrachten.

Ten slotte blijkt het van belang te zijn om voldoende mogelijkheden te voorzien waarbij leerkrachten ook aan interne professionalisering kunnen doen. Bijvoorbeeld door randvoorwaarden te creëren die

onderlinge samenwerking tussen leerkrachten stimuleren om zo hun evaluatiepraktijken in co-creatie te ontwikkelen. Door in te zetten op gedeelde en gezamenlijke professionalisering wordt het mogelijk om gaandeweg naar een gedeelde evaluatiepraktijk te evolueren die in wisselwerking staat met het evaluatiebeleid.

## **5. Aanbevelingen ten aanzien van leerkrachten**

### **Aanbeveling 5.1. Een evaluatiepraktijk die maximale leer- en ontwikkelingskansen nastreeft bij alle leerlingen.**

Een evaluatiebeleid die geënt is op diversiteit maximaliseert de leer- en ontwikkelingskansen van elke leerling. Het is hierbij cruciaal dat elk van de drie functies van evaluatie vanuit dit idee vorm en invulling krijgen (zie algemene aanbeveling). Zoniet, dreigt er een risico op het limiteren van de ontwikkelingskansen van (groepen van) leerlingen. Inzake het realiseren van een evaluatiepraktijk die maximale leer- en ontwikkelingskansen nastreeft, schuiven we drie specifieke aandachtspunten naar voren.

#### **(1) Zet maximaal in op een sterke integratie van evaluatie en instructie**

De onderzoeksresultaten laten zien hoe leer- en ontwikkelingskansen beperkt worden door een overwegend gescheiden evaluatiepraktijk. De kracht van evaluatie zit echter in het feit dat evalueren en leren onlosmakelijk met elkaar verbonden zijn. Dit komt het beste tot uiting in een geïntegreerde evaluatiepraktijk waarbij evaluatie voortdurend deel uitmaakt van de algemene onderwijspraktijk. Uit het onderzoek blijkt bovendien dat er een positief verband is tussen de mate waarin leerkrachten in staat zijn om diversiteit tussen leerlingen op te merken en de mate waarin ze een geïntegreerde evaluatiepraktijk aanwenden. Via dergelijke evaluatiepraktijk kunnen de leer- en ontwikkelingsbehoeften van leerlingen beter opgevangen worden. Een sterkere integratie van evaluatie en instructie resulteert enerzijds in het optimaal afstemmen van de instructieactiviteiten op de leerbehoeften van leerlingen. Anderzijds draagt dit bij tot het bevorderen van het leerproces van leerlingen aangezien een geïntegreerde evaluatiepraktijk veel kansen creëert om leerlingen directe feedback te geven. Op deze manier vormt evaluatie de motor van een inclusieve leeromgeving.

#### **(2) Betrek zoveel mogelijk de leerlingen en hun ouders bij het evaluatieproces**

Zowel met betrekking tot de ondersteunende functie van evaluatie als bij de functie van het nemen van beslissingen, loont het om de leerlingen en hun ouders te betrekken bij het evaluatieproces. Leerlingen kunnen betrokken worden voorafgaand bij het evaluatieproces door met hen heldere, transparante afspraken te maken over wanneer en wat er geëvalueerd wordt, en hoe dit gebeurt. Tijdens de evaluatie kunnen leerlingen nauw betrokken worden door hen een rol als (mede-)evaluator te geven via bijvoorbeeld zelf- of peerevaluatie. Het ondersteunende effect van evaluatie kan gemaximaliseerd worden door leerlingen ook na een evaluatie-activiteit volop te betrekken, bijvoorbeeld door ruime aandacht te besteden aan een feedbackmoment dat volgt na evaluatie. Uit het onderzoek blijkt namelijk dat een hogere betrokkenheid van leerlingen positief samenhangt met een geïntegreerde evaluatiepraktijk die overigens meer rekening houdt met de diversiteit onder leerlingen.

Met betrekking tot het evalueren om beslissingen te nemen over de schoolloopbaan is het eveneens van belang dat de leerling en zijn/haar ouders vroeg betrokken worden: dit om de leerling inzicht te laten krijgen in de eigen mogelijkheden en talenten, maar eveneens om de leerling en zijn/haar ouders te informeren over welke mogelijkheden de onderwijsstructuur biedt. Uit het casestudieonderzoek blijkt bijvoorbeeld dat de oriëntering van leerlingen sterk onderhevig is aan de kenmerken van het onderwijssysteem en gestuurd wordt door het watervaldenken. Een procesmatige aanpak waarbij de verschillende partners betrokken worden stimuleert daarentegen een meer leerlinggerichte oriëntering. Ook na het nemen van een beslissing blijft het van belang om samen met leerling en ouders de vinger aan de pols te houden om zo gemaakte keuzes eventueel te herzien.

### **(3) Maak van je evaluatiepraktijk een gedeelde praktijk**

Leerkrachten wordt geadviseerd om hun evaluatiepraktijk in te richten via co-creatie. Door nauw samen te werken met collega-leerkrachten wordt het evaluatiebeleid van de school ook bottom-up ontwikkeld, zo bleek uit de *good practices* van het casestudieonderzoek. Bovendien draagt het samenwerken en het uitwisselen van praktijkervaringen bij tot het intern professionaliseren van leerkrachten.



# Inhoud

<b>Voorwoord</b>	<b>4</b>
<b>Beleidsamenvatting</b>	<b>5</b>
<b>Inhoud</b>	<b>17</b>
<b>I Opzet van dit onderzoek</b>	<b>19</b>
1. Evaluatie: een cruciaal maar gevoelig thema	19
1.1. Binnen het onderwijs- & leerproces	19
1.2. Voor de leerling & diens toekomst	20
1.3. Bemoedlijkt door (inter)nationale ontwikkelingen	23
1.4. Een complex gegeven voor zowel school, leerkracht als leerling	24
2. Onderzoeksvragen	25
<b>II Een gefaseerd onderzoek</b>	<b>26</b>
1. Mixed-method design	26
2. Literatuurstudie	26
3. Casestudieonderzoek	27
4. Surveyonderzoek	29
<b>III De resultaten gebundeld</b>	<b>31</b>
<b>1. Model voor evaluatiebeleid</b>	<b>31</b>
<b>1. Evaluatiepraktijk van leerkrachten</b>	<b>32</b>
1.1. Beperkte inbedding van evaluatie in de algemene onderwijspraktijk	32
1.2. Evaluatie op momenten van transitie	36
1.3. Evalueren met gevalideerde instrumenten	38
1.4. Evaluatiepraktijk & diversiteit	39
<b>2. De opvattingen van leerkrachten</b>	<b>43</b>
2.1. Opvattingen met betrekking tot evaluatie	43
2.2. Houding ten aanzien van centrale, gestandaardiseerde examens	44
2.3. Opvattingen omtrent evalueren van kennis versus evalueren van competenties	45
2.4. Opvattingen over eigen competenties en bekwaamheid	46
<b>3. Evaluatiebeleid van scholen</b>	<b>48</b>
3.1. Evaluatiepraktijk is eerder leerkracht- dan schoolafhankelijk	48
3.2. Visie bij evaluatiebeleid ontbreekt vaak	48
3.3. Factoren die een evaluatiebeleid op school faciliteren	49
<b>4. De resultaten samengevat</b>	<b>54</b>
<b>IV Discussie</b>	<b>61</b>
<b>1. Onderzoek naar evaluatie in een context van diversiteit</b>	<b>61</b>
<b>2. Formatieve functie van evaluatie onderbenut</b>	<b>62</b>
<b>3. Summatieve functie overschaduwde de formatieve functie</b>	<b>63</b>
3.1. 'Waterval-denken' bepalend voor beslissingen over de schoolloopbaan	63
3.2. Selecteren in plaats van oriënteren	64
<b>4. De rol van het beleidsvoerend vermogen</b>	<b>66</b>
4.1. Gebrek aan geïntegreerd beleid	66
4.2. Gebrek aan reflecterend vermogen	67
<b>V Conclusie &amp; beleidsaanbevelingen</b>	<b>70</b>
<b>o. Algemene aanbeveling: Stel diversiteit centraal bij elk initiatief dat aan evaluatie raakt.</b>	<b>70</b>

<b>1. Aanbevelingen ten aanzien van de overheid</b>	<b>72</b>
<b>2. Aanbevelingen ten aanzien van pedagogische begeleidingsdiensten</b>	<b>75</b>
<b>3. Aanbevelingen ten aanzien van lerarenopleidingen</b>	<b>76</b>
<b>4. Aanbevelingen ten aanzien van schoolleiders</b>	<b>77</b>
<b>5. Aanbevelingen ten aanzien van leerkrachten</b>	<b>79</b>
<b>Bibliografie</b>	<b>81</b>

# I Opzet van dit onderzoek

## 1. Evaluatie: een cruciaal maar gevoelig thema

### 1.1. Binnen het onderwijs- & leerproces

Evaluatie omvat veel meer dan de toetsen en examens die leerlingen op het einde van een leercyclus afleggen. Evaluatie dient meerdere doelen (Philips, Seghers, Versteden, & Ysenbaert, 2013) waardoor evaluatie in de praktijk complexer is dan het op het eerste gezicht lijkt. Binnen de schoolse context speelt evaluatie op meerdere momenten én ten aanzien van verschillende actoren een cruciale rol. Zo heeft evaluatie altijd al een *informerende* functie gehad (Scriven, 1967), waarbij informatie over het leerproces verschaft wordt: wat heeft de lerende opgestoken tijdens de leercyclus? Zijn de doelen wel of niet bereikt? In de literatuur duikt deze functie van evaluatie ook op als summatieve evaluatie of ‘assessment of learning’.

Verder heeft evaluatie een *ondersteunende* functie, wat omschreven wordt als de formatieve functie van evaluatie. Initieel werd in de literatuur voornamelijk de ondersteunende rol ten aanzien van de leerkracht beklemtoond (Bloom, Hastings, & Madaus, 1971). De informatie uit evaluatie wordt aangewend door de leerkracht om verder richting te geven aan de onderwijsactiviteiten en zo didactische keuzes te beïnvloeden, wat omschreven wordt als ‘assessment for learning’ (Havnes, Smith, Dysthe, & Ludvigsen, 2012; Wiliam, 2011). Dankzij ontwikkelingen in de leerpsychologie en motivatietheorie werd deze formatieve functie van evaluatie steeds verder uitgebreid (Brookhart, 2007). De meerwaarde van evaluatie beperkt zich niet tot de leerkracht, maar breidt zich verder uit naar de leerling, die deze informatie kan gebruiken om zijn/haar leerproces bij te sturen (Clark, 2012). Bovendien hebben formatieve vormen van evaluatie een rechtstreekse impact op de motivatie van de leerling (Black & Wiliam, 1998). In recentere literatuur beklemtonen concepten als ‘assessment as learning’ (Dann, 2014; Earl, 2003) en ‘learning-oriented assessment’ (Jones, Saville, & Salamoura, 2016) de meerwaarde van evaluatie in de bevordering van het leerproces van de leerling.

Evaluatie speelt tevens een rol in *het nemen van beslissingen*. We maken een onderscheid tussen twee soorten beslissingen. Enerzijds zijn er de beslissingen in het kader van de schoolloopbaan van leerlingen (Nagy, 2000; Tierney, 2013). Bijvoorbeeld de studiekeuze, het instromen op een school of het doorverwijzen naar externe hulpverlening. Dit omvat beslissingen bij elke overgang naar een volgend leerjaar of de ‘grotere’ transitie momenten van kleuter- naar lager onderwijs en van lager naar secundair onderwijs. Deze functie van evaluatie is in meerdere of mindere mate verbonden met de twee eerder genoemde functies van evaluatie. Afhankelijk van het onderwijssysteem en het beleid op centraal, lokaal en schoolniveau zullen de beslissingen in meerder of mindere mate mee bepaald worden door de resultaten van summatieve en formatieve vormen van evaluatie.

Tegelijk kan evaluatie een rol spelen in beslissingen in het kader van kwaliteitsverbetering op het niveau van onderwijsorganisatie en/of onderwijssysteem (Nagy, 2000; Schildkamp, Lai, & Earl, 2012). Afhankelijk van het onderwijssysteem en het daarbij horende beleid, kan dit soort evaluatie schoolintern of -extern aangestuurd worden. Het feit dat deze vorm van beslissingen zich meer situeren op organisatie- of systeemniveau, neemt niet weg dat ze een impact uitoefenen op het

onderwijs- en leerproces. Integendeel, ze zijn doorgaans bedoeld om de kwaliteit van dit proces te verbeteren.

Het meest recente jaarverslag van de Onderwijsinspectie (2020) stelt dat leerlingenevaluatie in scholen een pijnpunt vormt. Het afstemmen van evaluatie op de eindtermen, het hanteren van evenwichtige en betrouwbare evaluatie die gevarieerd is blijkt moeilijk te zijn. In het secundair onderwijs zijn de uitdagingen het grootst: minder dan de helft van de doorgelichte scholen slaagt erin om hun leerlingen te evalueren volgens verwachting (Onderwijsinspectie, 2020).

## **1.2. Voor de leerling & diens toekomst**

Uiteindelijk heeft evaluatie een heel bepalende rol voor de toekomst van leerlingen. Leerkrachten baseren zich op de informatie van evaluaties om beslissingen te nemen over studiekeuze en bepalen zo mee of de leerling al dan niet moet zittenblijven en in welke onderwijsvorm een leerling terecht komt (Tierney, 2013). Afhankelijk van de mate waarin een onderwijssysteem flexibel is en zo de mogelijkheden voor (her)oriëntering open houdt, zal evaluatie een afdoende rol hebben met betrekking tot de kansen binnen hoger onderwijs. Bijgevolg bepaalt evaluatie ook de latere mogelijkheden op de arbeidsmarkt mee en de algemene levenskwaliteit binnen het volwassen leven. Hoewel dit cruciale aspect van evaluatie vooral verbonden lijkt met de functie van 'beslissingen nemen over de schoolloopbaan', vallen de overige functies van evaluatie niet te onderschatten als het gaat over hun invloed die ze (on)rechtstreeks uitoefenen op de toekomst van leerlingen.

### **1.2.1. Informeren**

Evaluatie informeert de leerling, leerkracht en andere actoren over waar de leerling staat in zijn/haar leerproces. De informatie die uit evaluatie voortvloeit, schept een beeld over de leerling. De mate waarin evaluatie op dit vlak in staat is om een totaalbeeld van de leerling weer te geven, hangt onder andere af van wat de leerkracht beoogt te weten via evaluatie. Dit laatste bepaalt de mate waarin er gepeild wordt naar: kennis en vaardigheden binnen specifieke cognitieve domeinen, sociale vaardigheden, attitudes, leerhouding, motivatie, metacognitieve kennis en zelfregulerende vaardigheden. Afhankelijk van wat er aandacht krijgt bij evaluatie en welk evenwicht hierbij gehanteerd wordt, worden mogelijks bepaalde talenten van een leerling niet opgemerkt. Naast de keuze omtrent 'wat' er geëvalueerd wordt, kan een mogelijke 'blinde vlek' of 'verkeerd' beeld van de leerling ook ontstaan door de keuze rond 'hoe' er geëvalueerd wordt. Bepaalde vormen van evaluatie, zoals bijvoorbeeld pen-en-papier-toetsen lenen zich nu eenmaal beter om de reproductie van woordenschatkennis in kaart te brengen dan bijvoorbeeld de spreekvaardigheid. Verder speelt de wijze waarop er geëvalueerd wordt ook een rol in de mate waarin leerlingen de kans krijgen om te laten zien wat ze kunnen en/of weten. Verschillende studies toonden aan dat sommige manieren van evalueren een hindernis kunnen vormen voor bepaalde (groepen van) leerlingen (De Backer, Van Avermaet, & Slembrouck, 2016; Nayir et al., 2019; Shohamy, 2001). Zo kan de taal waarin een toets afgenomen wordt een drempel vormen voor leerlingen die deze taal niet als moedertaal hebben (Shohamy, 2001). De Backer et al. (2016) verwijzen in dit verband naar het probleem van constructvaliditeit indien de talige complexiteit van een test leerlingen verhindert om te laten zien dat ze de leerinhoud van een niet-taalvak wel degelijk beheersen. De drempels kunnen velerlei zijn en beperken zich niet enkel tot de taal waarin geëvalueerd wordt. Nayir et al. (2019) verwijzen daarom naar '*culturally responsive assessments*' die zoveel mogelijk rekening houden met culturele verschillen onder leerlingen en zo ook hun verschillende leerstijlen respecteren. Dergelijke studies binnen het domein van toetsing en evaluatie zorgen ervoor dat naast de klassieke psychometrische

kwaliteitscriteria zoals validiteit en betrouwbaarheid de aandacht gevestigd wordt op bijkomende criteria zoals 'fairness' (Tierney, 2013) en 'billijkheid' in evaluatie (Van Petegem & Vanhoof, 2002). Deze aandachtspunten beogen precies tegemoet te komen aan de ontorechte, vaak onbedoelde, maar mogelijk nefaste impact van evaluatie op de resultaten van de leerling en bijgevolg ook op diens toekomst. Het spreekt voor zich dat, om recht te doen aan (de toekomst van) de leerling, evaluatiepraktijken in het kader van de informerende functie van evaluatie enige zorgvuldigheid en aandacht vergen. Immers, de informerende functie werkt onrechtstreeks door op de toekomst van leerlingen aangezien deze mee de functie van het nemen van beslissingen over de schoolloopbaan beïnvloedt.

### 1.2.2. Ondersteunen

Dezelfde zorgvuldigheid en aandacht zijn van belang voor de ondersteunende functie van evaluatie. Afhankelijk van wanneer en hoe vaak de leerkracht evalueert, zal het onderwijsproces net meer of minder 'gevoed' worden. De leerkracht kan zo aan de hand van bijvoorbeeld pre-toetsing inzicht verwerven over de beginsituatie van de leerlingen (Tomlinson & Moon, 2013b). Bovendien stellen 'on-the-fly-assessments' (McNamara, 2001) de leerkracht in staat om voortdurend keuzes te maken over bijvoorbeeld het tempo waarmee nieuwe leerstof wordt behandeld. Teleurstellende toetsresultaten die initieel verkregen werden via de informerende functie van evaluatie kunnen als ondersteunende functie aangewend worden door in navolging hiervan een bepaald leerstofonderdeel nog eens te herhalen met (bepaalde) leerlingen. Evaluatie heeft zo een signaalfunctie voor de leerkracht om te achterhalen wie nog bijkomende instructie nodig heeft om minimaal de door de overheid vastgelegde eindtermen te bereiken. Op deze manier vormt evaluatie de motor om te differentiëren en maximaal aan te sluiten bij de leerbehoeften van leerlingen (Tomlinson & Moon, 2013a, 2013b). De vaststelling dat leerkrachten zichzelf niet bekwaam voelen om tegemoet te komen aan de noden van leerlingen (De Boer, Pijl, & Minnaert, 2011) doet vermoeden dat het potentieel van evaluatie in het kader van differentiatie nog onvoldoende benut wordt. Daarnaast stelt onderzoek naar de implementatie van het M-decreet vast dat het afstemmen van het evaluatiebeleid en de evaluatiepraktijken op de noden van leerlingen nog moeilijk loopt (Struyf, Verschueren, Petry, & Van Mieghem, 2018).

Het stimuleren van het leerproces bij de leerling is een aspect dat zijn oorsprong vindt in de ondersteunende functie van evaluatie. Een concreet voorbeeld hiervan is het geven van feedback (Hattie & Timperley, 2007), dat zelfs wordt gezien als één van de meest effectieve elementen om leerlingen te helpen leren (Hattie, 2012). Onderzoek over het gebruik van feedback door leerkrachten laat echter zien dat de mogelijkheden hiervan vaak niet ten volle benut worden. Dit wordt bevestigd door het PISA-onderzoek, waar aan de hand van een vragenlijst voor leerlingen een index voor leerkrachtfeedback samengesteld wordt. De meest recente resultaten laten zien dat Vlaanderen, op vier landen na, de laagste index heeft (De Meyer et al., 2019). Het meest recente jaarverslag van de Onderwijsinspectie bevestigt deze bevinding en stelt dat dit binnen het secundair onderwijs een nog grotere uitdaging vormt dan in het basisonderwijs. In minder dan de helft van de secundaire scholen die doorgelicht werden in schooljaar 2018-2019 werd er feedback gegeven volgens de gehanteerde kwaliteitsverwachting. Hoewel het basisonderwijs het hier iets beter doet, zijn er ook daar nog kansen tot verbetering aangezien men in de helft tot drie vierde van de doorgelichte basisscholen feedback geeft volgens verwachting (Onderwijsinspectie, 2020). Kwaliteitsvolle feedback besteedt aandacht aan het proces en het zelfregulerend vermogen (Hattie & Timperley, 2007), maar Carless, Salter, Yang, and Lam (2011) bevestigen dat slechts een minderheid van de leerkrachten belang hecht aan de ontwikkeling van dit zelfregulerend leren. Terwijl net de zwakker presterende leerlingen hier het

meeste nood aan hebben (Zimmerman, 1990). De ondersteunende functie van evaluatie is dan ook uitermate cruciaal om het leerproces van de leerling te bevorderen, en des te crucialer voor risico-leerlingen.

### 1.2.3. Beslissingen nemen

Bij aanvang van deze paragraaf werd al toegelicht hoe de evaluatiepraktijk vanuit de informerende functie cruciaal kan zijn voor de toekomst van leerlingen. Afhankelijk van de aandacht voor wat en hoe er geëvalueerd wordt, zal een leerling net meer of minder ondersteund worden en zo meer of minder kans maken om in een studierichting terecht te komen waarin zijn/haar talenten ten volle ontwikkeld kunnen worden. Bij het nemen van beslissingen over de schoolloopbaan, speelt er echter meer mee dan alleen de resultaten van de leerling. Onderzoek in de Vlaamse context (Thys, 2018) liet zien dat 15% van de leerkrachten vindt dat de sociale achtergrond van de leerling meegenomen moet worden bij oriëntering. Leerlingen uit een arbeidersgezin maken zo beduidend minder kans om naar het aso georiënteerd te worden in vergelijking met leerlingen wiens ouders hoog opgeleid zijn. Voorstanders van centrale examens zien hierin een oplossing om leerlingen net te oriënteren op basis van 'objectieve' resultaten (Bol, Witschge, Van de Werfhorst, & Dronkers, 2014). Tegelijkertijd worden ook vragen gesteld bij de idee van objectiviteit die aan dergelijke resultaten gekoppeld wordt (Shohamy, 2001). Andere auteurs verwijzen naar nadelige effecten van centrale examinering op het onderwijsleerproces (Amrein & Berliner, 2003; Jones, Jones & Hargrove, 2003; Nichols and Berliner, 2007, 2008): het bederft de leerling-leerkrachtrelatie die belangrijk is om het leren te bevorderen en leerlingen ontwikkelen eerder een negatieve houding ten aanzien van leren.

De functie van het nemen van beslissingen kan zich ook situeren op het niveau van onderwijsorganisatie en/of onderwijssysteem, in het kader van kwaliteitsverbetering (Nagy, 2000; Schildkamp et al., 2012). In de Vlaamse context blijkt dat vijftig tot zestig procent van de scholen niet systematisch inzet op het evalueren van hun onderwijskwaliteit. Voor deze scholen is het onduidelijk hoe ze aan kwaliteitsontwikkeling kunnen doen met behulp van de reeds aanwezige data (Onderwijsinspectie, 2020). Hoewel deze functie van evaluatie de bedoeling heeft om de kwaliteit van onderwijs, en dus ook de resultaten van leerlingen in het algemeen, te bevorderen, oefent dit mogelijk ook invloed uit op de individuele leerling. Volgens Sahlberg (2006) bijvoorbeeld, leidt centrale examinering die kwaliteitsontwikkeling beoogt tot curriculum-verenging: leerkrachten leggen via teaching-to-the-test meer nadruk op taal en wiskunde in plaats van op de 21e eeuwse vaardigheden. Dit illustreert hoe centrale initiatieven met betrekking tot de functie van het nemen van beslissingen mogelijk doorwerkt tot op de informerende functie die evaluatie heeft op de klasvloer. De centrale examens bepalen zo mee 'wat' er geëvalueerd wordt en dragen al dan niet bij aan het vergroten van de blinde vlek als het gaat over het in kaart brengen van de competenties van leerlingen. Anderzijds toont onderzoek van Bol et al. (2014) dat in onderwijssystemen met centrale toetsing, de impact van SES bij de oriëntering van leerlingen net afgezwakt wordt omdat de oriëntering van leerlingen meer gebaseerd is op objectieve indicatoren en minder afhankelijk is van vooringenomenheid van leerkrachten.

Het feit dat alle bovengenoemde functies cruciaal blijken te zijn voor de leerling en diens toekomst, maakt dat evaluatie een delicaat gegeven is. Evaluatie lijkt zo wel een complexe machine te zijn die, indien ze goed aangestuurd wordt, heel doelgericht en functioneel kan zijn. Het delicate zit precies vevat in de aansturing: de verschillende machineonderdelen kunnen potentieel bijdragen aan het ontwikkelen van talenten en competenties van leerlingen, maar kunnen tegelijkertijd ook hun

mogelijkheden beknootten en voorlopig beperkt ontwikkelde talenten negeren. Het hoeft dan ook geen betoog dat evaluatie, en de wijze waarop evaluatie aangestuurd en gehanteerd wordt binnen het onderwijs- en leerproces en het ruimere onderwijssysteem, een belangrijke rol speelt bij het maximaliseren van leer- en ontwikkelingskansen van alle leerlingen.

### **1.3. Bemoeilijkt door (inter)nationale ontwikkelingen**

#### **1.3.1. Ontwikkeling richting meer inclusie**

Naast het feit dat evaluatie an sich al een complexe taak is voor zowel leerkracht als school, zijn er een aantal (internationale) ontwikkelingen die het er niet makkelijker op maken. Enerzijds stellen we vast dat het leerlingenpubliek diverser wordt. Deze diversiteit uit zich op allerlei manieren: leerlingen verschillen onderling in interesses, talenten, motivatie, leervermogen, de talen die ze beheersen, in leertempo, in het al dan niet hebben van fysieke beperkingen of leerstoornissen, in reeds gevolgde trajecten indien het oudere leerlingen betreft, enzovoort. Deze feitelijke diversiteit wordt bovendien geïntensiveerd door een gepercipieerde diversiteit die van hogerhand gestimuleerd wordt. Zo wekken bepaalde beleidsinitiatieven bij leerkrachten de idee op dat de diversiteit waarmee ze moeten omgaan in de klas toeneemt. Het Verdrag van de Verenigde Naties (2006), dat in België geratificeerd werd, bevestigt het recht op onderwijs van personen met een handicap en roept op tot het waarborgen van een inclusief onderwijssysteem op alle niveaus om zo maximale leer- en ontwikkelingskansen voor iedereen mogelijk te maken. Dit verdrag lag mee aan de basis van het M-decreet dat onder meer de bedoeling heeft om meer leerlingen met een profiel buitengewoon onderwijs in het gewoon onderwijs te laten school lopen. Bovendien legt dit M-decreet de nadruk op het realiseren van gepaste zorg die aansluit bij de onderwijsbehoeften van alle leerlingen, ook ten aanzien van deze leerlingen die niet meteen een profiel voor buitengewoon onderwijs hebben. Van scholen wordt verwacht dat ze inzetten op de ontwikkeling van een algemene brede zorgvisie (Petry et al., 2020; Struyf et al., 2018).

De perceptie van steeds diverser wordende klassen wordt via deze trend naar inclusie versterkt. Leerkrachten worden uitgedaagd om hun aanpak hierop af te stemmen en zoeken oplossingen in redelijke aanpassingen, (binnenklas)differentiatie, *universal design for learning*, ... Doorgaans richten deze oplossingen zich op de algemene onderwijs- en instructieactiviteiten. Deze oplossingen doortrekken naar de evaluatiepraktijk blijkt geen evidentie te zijn (James & Pedder, 2006; Shepard, 2000). Onderzoek omtrent de implementatie van het M-decreet wijst hierbij op de specifieke uitdagingen voor het evaluatiebeleid van scholen, zeker met betrekking tot gedifferentieerde evaluatie (Struyf et al., 2018).

#### **1.3.2. Ontwikkeling richting meer standaardisering**

Tegelijkertijd is er een (internationale) tendens om meer in te zetten op centraal gestuurde en gestandaardiseerde vormen van evaluatie. Een trend die aangewakkerd wordt door grootschalige internationale toetsprogramma's, zoals bijvoorbeeld PISA en TIMSS, waarbij de resultaten van verschillende deelnemende landen onderling vergeleken kunnen worden. Teleurstellende resultaten lagen mee aan de basis van bijvoorbeeld het Duitse onderwijsbeleid om centrale toetsen te organiseren om zo het onderwijsniveau te doen stijgen. Deze trend is wereldwijd merkbaar en wordt in de literatuur omschreven als '*global education reform movement*' (Sahlberg, 2006). Onrust met betrekking tot een dalende onderwijskwaliteit wordt door veel landen beantwoord met een toenemende centralisering. Dergelijke centrale evaluatie heeft ook impact op curriculumvorming. Deze ontwikkeling is ook voelbaar in de Vlaamse onderwijscontext. De meest recente cijfers van PISA

bevestigen de al eerder ingezette dalende trend voor de resultaten van Vlaanderen (De Meyer et al., 2019). De tendens naar meer gestandaardiseerde vormen van evaluatie is niet alleen aanwezig in het maatschappelijke en onderwijskundige debat, maar vertaalt zich ook in de beleidsinitiatieven in de beleidsnota onderwijs voor de periode 2019-2024, waarbij verwezen wordt naar *‘het ontwikkelen van gevalideerde, gestandaardiseerde en genormeerde toetsen’* om de kwaliteit van het onderwijs te bewaken. De nota blikkt vooruit op de concrete implementatie van deze toetsen: *“Alle scholen nemen de proeven af bij alle leerlingen op twee momenten in het lager onderwijs en aan het einde van de eerste en de derde graad van het secundair onderwijs.”*

#### **1.4. Een complex gegeven voor zowel school, leerkracht als leerling**

In de praktijk is het geen evidentie om alle functies van evaluatie (zoals toegelicht in 1.1) te realiseren. De veelheid aan functies impliceert dat de leerkracht meerdere rollen aanneemt. Dit alleen al zorgt ervoor dat ‘evaluatie’ een complex gegeven is. Met betrekking tot de ondersteunende functie treedt de leerkracht op als begeleider van het leerproces, als coach van de lerende die het leerproces stimuleert en die afoetst op welke momenten er geremedieerd moet worden. Tegelijkertijd heeft diezelfde leerkracht ook de opdracht om mee te beslissen over de schoolloopbaan van de leerling oftewel om te beoordelen. Het samenkomen van deze – soms schijnbaar tegenstrijdige rollen – maakt dat de leerkracht in een moeilijke, kwetsbare positie komt, waarbij het moeilijk balanceren is tussen de rol van begeleider enerzijds en de rol van beoordelaar anderzijds.

De inzichten uit 1.2. stelden de complexiteit omtrent leerlingenevaluatie scherper en meerbepaald met het oog op het maximaliseren van leer- en ontwikkelingskansen voor éénieder. De impact op de toekomst van leerlingen maakt dat evaluatie voor zowel school als leerkracht een heikele klus vormt. Daarenboven stellen de ruimere (inter)nationale ontwikkelingen, die beschreven werden in 1.3., de evaluatiepraktijk van leerkrachten en het evaluatiebeleid van scholen voor een uitdaging. Het is opmerkelijk hoe er parallel aan de beweging naar meer inclusief onderwijs zich ook een evolutie ontwikkelt richting meer gestandaardiseerde vormen van evaluatie. Enerzijds worden leerkrachten uitgedaagd om zoveel mogelijk aan te sluiten bij éénieders leernoden- en behoeften opdat élke leerling kan excelleren. Anderzijds zullen de leerlingen van diezelfde leerkrachten op vastgelegde momenten misschien allemaal dezelfde toets moeten afleggen opdat het Vlaamse onderwijs opnieuw kan excelleren. Deze ontwikkelingen dagen scholen uit om een evaluatiebeleid uit te werken dat aan beide tendensen het hoofd biedt.



## 2. Onderzoeksvragen

In dit SONO-onderzoek nemen we het evaluatiebeleid van scholen en de evaluatiepraktijk van leerkrachten in de Vlaamse context onder de loep. De onderstaande onderzoeksvragen vormden de leidraad. Een deel van deze vragen gaat in op het beschrijven van de huidige stand van zaken: hoe zien het evaluatiebeleid en de evaluatiepraktijken eruit. Andere vragen beogen te achterhalen welke factoren bijdragen tot de ontwikkeling van een evaluatiebeleid en het bevorderen van evaluatiepraktijken.

OV1: Wat is het evaluatiebeleid, en wat zijn de evaluatiepraktijken in Vlaamse scholen in basis- en secundair onderwijs?

- OV1a: Hebben scholen een evaluatiebeleid? Waarom evalueert men? Wat evalueert men en wat doet men ermee? Kunnen hierin typologieën onderscheiden worden?
- OV1b: Op welke wijze wordt er geëvalueerd? Welke instrumenten worden er gebruikt? Is er een verschil tussen evaluatie op transitiemomenten versus meer continue klasevaluatie?
- OV1c: In welke mate (en hoe) houden scholen en leerkrachten rekening met de diversiteit aan leerlingen (sociaal, etnisch, gender, ...) in hun evaluatiebeleid en -praktijk?
- OV1d: In welke mate houden scholen en leerkrachten rekening met de diversiteit aan ontwikkelingsdoelen en eindtermen in hun evaluatiebeleid en -praktijk?
- OV1e: Hoe verhouden het evaluatiebeleid en de praktijk zich ten opzichte van de werking en de rol van de klassenraden (toelatingsklassenraad, begeleidende klassenraad, ...)?
- OV1f: Hoe verhouden het evaluatiebeleid en de praktijk zich ten opzichte van het attesteringsbeleid?
- OV1g: Hoe verhouden het evaluatiebeleid en de praktijk zich ten opzichte van het oriënteringsbeleid naar A/B-stroom (en heroriëntering van B naar A-stroom)?

OV2: Wat zijn bepalende factoren bij evaluatiebeleid en -praktijk in Vlaamse scholen in het basis- en secundair onderwijs?

- OV2a: Hoe komt een evaluatiebeleid tot stand? Wat zijn bepalende factoren op school-, leerkracht- en leerlingniveau? Welke schoolexterne actoren (pedagogische begeleiding, ...) worden betrokken bij het tot stand komen van een evaluatiebeleid?
- OV2b: (Hoe) communiceren scholen en leerkrachten over de (individuele) evaluatiepraktijk aan collega's, leerlingen en ouders? Op welke manier motiveren scholen en leerkrachten de eigen evaluatiepraktijk?
- OV2c: Hoe motiveren scholen en leerkrachten de mate waarin ze rekening houden met ontwikkelingsdoelen en eindtermen? (Hoe) communiceren ze deze keuzes met collega's, leerlingen en ouders?

OV3: Wat zijn de opvattingen van directies, leerkrachten en leerlingen op vlak van evaluatie in het algemeen en meer specifiek met betrekking tot de evaluatiepraktijk en het -beleid van hun school?

- OV3a: Wat zijn de opvattingen van de verschillende actoren over centrale, gestandaardiseerde instrumenten versus meer continue procesevaluatie? Is er sprake van een gedeelde visie (i.e., cultuur) hieromtrent?
- OV3b: Wat zijn de opvattingen van de verschillende actoren over kennisgerichte versus competentiegerichte evaluatie? Is er sprake van een gedeelde visie (i.e., cultuur) hieromtrent?
- OV3c: Wat zijn de opvattingen van de verschillende actoren over evaluatie op momenten van transitie (bv. jaarlijkse overgang van jaar x naar y; overgang KO-LO; overgang LO-SO; overgang graden SO;) versus dagdagelijkse 'classroom based' evaluatie?

OV4: In welke mate voelen leerkrachten zich competent (handelingsbekwaam) om te evalueren? Voelen ze zich even competent om competentiegericht als kennisgericht te evalueren? Waardoor wordt deze competentie – of het gebrek eraan – bepaald?

Deze onderzoeksvragen houden we tegen het licht van de hierboven omschreven uitgangstelling, namelijk dat evaluatie (1) cruciaal is voor het onderwijs- en leerproces en zo ook (2) voor de leerling en diens toekomst en (3) zich bovendien situeert binnen hedendaagse (inter)nationale ontwikkelingen die bijkomende uitdagingen met zich meebrengen.

## II Een gefaseerd onderzoek

### 1. Mixed-method design

Voor de aanpak werd geopteerd voor een mixed-method design waarbij zowel kwalitatieve als kwantitatieve onderzoeksmethoden gebruikt werden. Hierdoor is de uiteindelijke combinatie van data complementair in sterktes en overlapt dit elkaar niet in zwaktes (Johnson & Turner, 2003). Het mixed method-design omvat drie fases: 1) een literatuurstudie, 2) een casestudieonderzoek en 3) een surveyonderzoek. De onderzoeksluiken volgden elkaar op en werden volgens onderstaand tijdsplan uitgevoerd tussen september 2016 en juli 2020.



**Figuur 2.1:** Fasering van het mixed-method design

In wat volgt, lichten we kort de methodologie toe voor de verschillende fasen. Voor de uitgebreide rapportering verwijzen we naar de op zichzelf staande deelrapporten waarmee elke fase afgerond werd:

- Ysenbaert, J., Van Avermaet, P., & Van Houtte, M. (2017). *Literatuurstudie evaluatie en diversiteit*, Steunpunt Onderwijsonderzoek, Gent.
- Ysenbaert, J., Van Avermaet, P., & Van Houtte, M. (2018). *Evaluatie & diversiteit in het basis- en secundair onderwijs. Deelrapport: casestudieonderzoek*, Steunpunt Onderwijsonderzoek, Gent.
- Ysenbaert, J., Van Avermaet, P., & Van Houtte, M. (n.d.). *Evaluatie & diversiteit in het basis- en secundair onderwijs. Deelrapport: survey-onderzoek*, Steunpunt Onderwijsonderzoek, Gent.

### 2. Literatuurstudie



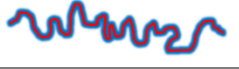
De literatuurstudie biedt een referentiekader aan van waaruit de huidige stand van zaken inzake evaluatiebeleid en –praktijk en hoe diversiteit hierbinnen een plaats krijgt, verklaard kan worden.

Het voorgestelde referentiekader vertrekt vanuit een holistisch perspectief waarbij ‘evaluatie’ niet als een losstaand gegeven bekeken wordt, maar in relatie gezien wordt tot meerdere domeinen dan alleen de school en de klas. Onder ‘evaluatie’ wordt begrepen het evaluatiebeleid van een school en de evaluatiepraktijken van leerkrachten, maar evenzeer de opvattingen en overtuigingen met betrekking tot evaluatie van relevante stakeholders. Dit alles vormt dan ook het voorwerp van dit onderzoek. Het referentiekader geeft weer hoe ‘evaluatie’ ingebed is in zowel een onderwijskundige context alsook in een socio-culturele en (internationale) politiek-economische context. Deze contexten zijn nauw verstrengeld met elkaar en beïnvloeden elkaar voortdurend en oefenen aldus invloed uit op het evaluatiebeleid van een school, de evaluatiepraktijken van leerkrachten en hun opvattingen met

betrekking tot evaluatie. Daarnaast bepalen deze contexten ook mee hoe deze evaluatiepraktijken en het evaluatiebeleid van scholen zich verhouden tot de toenemende diversiteit.

We lichten hier specifiek even de onderwijskundige context uit omdat de literatuurstudie met betrekking tot dit aspect resulteerde in drie modellen. Deze modellen representeren hoe evaluatiepraktijken zich verhouden tot de algemene instructieactiviteiten van een leerkracht. We lichten ze hier concreter toe omdat ze in zowel het kwalitatieve als het kwantitatieve onderzoeksluik als onderzoeksinstrument gebruikt werden.

**Tabel 2.1. Drie modellen voor evaluatiepraktijk: gescheiden, semi-geïntegreerd en geïntegreerd model (Ysenbaert, Van Houtte, & Van Avermaet, 2020)**

	 <b>Gescheiden model</b>	 <b>Semi-geïntegreerd model</b>	 <b>Geïntegreerd model</b>
<b>Leerparadigma</b>	Behaviorisme	Cognitivisme	(Sociaal-)constructivisme
<b>Onderwijzen wordt beschouwd als</b>	het deconstrueren van complexe competenties in verschillende niveaus van kennis en vaardigheden (Shepard, 2000).	het uitbreiden van leerlingen hun mentale modellen (Schunk, 2012).	het faciliteren van kennisconstructie.
<b>Evaluatie wordt beschouwd als</b>	het nagaan of de onderwijsactiviteit geslaagd is (Ertmer & Newby, 2013).	<ul style="list-style-type: none"> <li>- het activeren van voorkennis bij leerlingen (Ertmer &amp; Newby, 2013).</li> <li>- input waarmee didactische keuzes bijgestuurd worden.</li> </ul>	<ul style="list-style-type: none"> <li>- input waarmee didactische keuzes bijgestuurd worden.</li> <li>- input waarmee het leerproces van leerlingen bijgestuurd wordt (metacognitieve en zelfregulerende vaardigheden).</li> </ul>
	<i>Evalueren VAN het leren</i>	<i>Evalueren OM te leren</i>	<i>Evalueren ALS leren</i>
<b>Evaluatie binnen het onderwijsproces situeert zich</b>	op het einde van een onderwijsproces	aan het begin van het onderwijsproces en meer geïntegreerd in het onderwijsproces.	sterk geïntegreerd in het onderwijsproces, waardoor leren mogelijk wordt.
<b>Leerkracht of leerlinggestuurd?</b>	Leerkrachtgestuurd	Leerkrachtgestuurd, maar met een focus op de leerling waardoor toekomstige instructies aangepast worden naar zijn/haar noden.	Leerlingen zijn sterk betrokken bij het evaluatieproces.

### 3. Casestudieonderzoek

Voor het casestudieonderzoek werden zes basisscholen en zes secundaire scholen onderzocht. De scholen werden zo geselecteerd dat er onderling variatie was op vlak van 'diversiteit', dit op basis van de leerlingencompositie op school. Dit selectiecriteria werd geoperationaliseerd aan de hand van de

indicator 'opleidingsniveau moeder'. Zo wordt de helft van de onderzochte scholen gekenmerkt door een eerder homogene leerlingencompositie, waarbij een laag percentage (minder dan 10%) aan leerlingen een laagopgeleide moeder heeft. De andere helft van de scholen wordt gekenmerkt door een meer heterogene leerlingencompositie, waarbij 30 tot 60% van de leerlingen een laagopgeleide moeder heeft. In het secundair onderwijs werd erover gewaakt dat er voldoende variatie is tussen de scholen op vlak van studieaanbod: zo zijn alle onderwijsvormen vertegenwoordigd. Daarnaast is er ook variatie in het type structuur van de school: zowel een middenschool, een bovenbouwschool en een middelbare school met een drie graden-structuur zijn opgenomen. Ten slotte liet de cyclische steekproeftrekking toe dat er voor de scholen van het basisonderwijs één methodeschool meegenomen werd omwille van de specifieke invulling van hun evaluatiebeleid en – praktijk.

Onderstaande tabellen geven de kenmerken weer van de scholen die deelnamen aan het casestudieonderzoek.

**Tabel 2.2. Cases basisonderwijs**

	% leerlingen met laag opleidingsniveau moeder	Provincie	Net
<b>A</b>	4,55	W-VL	Vrij
<b>B</b>	4,46	W-VL	Vrij
<b>C</b>	56,50	VL-Br	OVSG
<b>D</b>	62,91	L	Vrij
<b>E</b>	8,56	A	Vrij
<b>F</b>	35,02	O-VL	OVSG

**Tabel 2.3. Cases secundair onderwijs**

	% leerlingen met laag opleidingsniveau moeder	ASO	BSO	KSO	TSO	1e graad	2e graad	3e graad	Provincie	Net
<b>A</b>	0,19					X			W-VL	GO!
<b>B</b>	0,32	X	X		X	X	X	X	W-VL	GO!
<b>C</b>	0,09	X	X		X	X	X	X	A	Vrij
<b>D</b>	0,18			X			X	X	O-VL	OVSG
<b>E</b>	0,08	X					X	X	W-VL	Vrij
<b>F</b>	0,58		X		X	X	X	X	A	OVSG

De onderzoeksdata werd verzameld via verschillende bronnen. Er werden semi-gestructureerde interviews afgenomen met de directie en met verschillende leerkrachten. De interviews met leerkrachten werden stevast voorafgegaan door één of twee lesobservaties. Via deze observatie werd het vertrouwen tussen respondent en onderzoeker versterkt, waardoor de kans op sociaal wenselijke antwoorden van de respondent afnam. De lesobservatie zorgde tevens voor aanknopingspunten om het gesprek over evaluatie rond op te bouwen.

In het basisonderwijs werden er telkens twee lessen geobserveerd in het derde kleuter, eerste leerjaar, derde leerjaar en zesde leerjaar. In één case (case C) werd de leerkracht van het zesde leerjaar niet

bereid gevonden om deel te nemen aan het onderzoek waardoor de observatie doorging in het vijfde leerjaar. In de methodeschool (case F) is er geen sprake van leerjaren maar van graden waardoor er op deze school in elke graad geobserveerd werd. Tijdens de observaties werd telkens een les taal en een les wiskunde gevolgd. Hierbij werd er gefocust op de mate waarin leerkrachten de kansen die zich voordoen tijdens een les om te evalueren in functie van het leerproces ook daadwerkelijk aangrijpen en naar de mate waarin er rekening gehouden werd met diversiteit. In totaal werden er in het basisonderwijs 48 lessen geobserveerd.

In het secundair onderwijs werd er naar gestreefd om in de mate van het mogelijke, volgens het aanbod van de school telkens te observeren in het eerste jaar, derde jaar en zesde jaar (opdat alle graden vertegenwoordigd zouden zijn). Binnen het aso werd er geopteerd om een les te observeren binnen de vakken 'wetenschappen' en 'Nederlands'. In het tso werd er geobserveerd in 'Nederlands' en een 'technisch vak (TV)'. In het bso werd dit 'PAV' (indien Nederlands opgenomen was binnen PAV) en een 'praktijkvak (PV)'. In het kso werd er geobserveerd binnen 'Nederlands' en een 'kunstvak KV'. In de eerste graad werd ernaar gestreefd om telkens zowel de a-stroom en de b-stroom mee te nemen.

Naast de interviews en observaties werd ook de stem van de leerlingen meegenomen. In het basisonderwijs gebeurde dit aan de hand van individuele interviews met leerlingen van het zesde leerjaar. In het secundair onderwijs werden er focusgroepen georganiseerd per graad waarin verschillende onderwijsvormen vertegenwoordigd werden.

Onderstaand overzicht geeft weer hoeveel data verzameld werd per onderwijsniveau:

Basisonderwijs:

- Semi-gestructureerd interview met directie (6)
- Semi-gestructureerd interview met leerkrachten (24)
- Semi-gestructureerd interview met leerlingen (24)
- Lesobservaties (48)

Secundair onderwijs:

- Semi-gestructureerd interview met directie (6)
- Semi-gestructureerd interview met leerkrachten (36)
- Focusgroepen met leerlingen (13)
- Lesobservaties (38)

Naast bovengenoemde onderzoeksdata werden per case ook relevante documenten opgevraagd: zoals beleidsdocumenten van de school in verband met evaluatiebeleid, voorbeelden van rapporten, het schoolreglement, werkgroepverslagen ... Bij elke observatie werden ook de documenten opgevraagd die gebruikt werden bij de geobserveerde les.

## 4. Surveyonderzoek

Via een online survey werden zowel leerkrachten als directies bevroegd met betrekking tot het evaluatiebeleid van de school en hun evaluatiepraktijken. De survey peilde naar hoe hun evaluatiepraktijk zich verhoudt tot hun algemene lespraktijk, in welke mate men bij evaluatie aandacht heeft voor aspecten zoals eerlijkheid, transparantie, verantwoording, impact en betrokkenheid van

leerlingen. De focus van dit onderzoek ligt niet alleen op leerlingenevaluatie in het algemeen, maar ook op de mate waarin hierbij rekening gehouden wordt met diversiteit onder leerlingen. Hiertoe werd naast de opvattingen omtrent evaluatie ook gepeild naar de opvattingen ten aanzien van de rol van onderwijs inzake het omgaan met diversiteit onder leerlingen en naar de mate waarin leerkrachten zichzelf bekwaam voelen om diversiteit onder leerlingen op te merken.

Voor dit onderzoek werd gebruik gemaakt van twee vragenlijsten: één voor directies en één voor leerkrachten. Zowel scholen met veel diversiteit in het leerlingenpubliek als scholen met een eerder homogeen leerlingenpubliek werden meegenomen in het onderzoek om te kijken of er met betrekking tot leerlingencompositie schoolspecifieke tendensen vastgesteld kunnen worden op vlak van evaluatiebeleid- en praktijk. In totaal werden 47 directeurs en 1120 leerkrachten bereikt uit 56 scholen die deelnamen aan het onderzoek.

In totaal namen er 1120 leerkrachten deel aan de online bevraging, waarvan 419 leerkrachten in het basisonderwijs en 701 leerkrachten in het secundair onderwijs. De vragenlijst voor directies werd in totaal door 47 directies ingevuld waarbij 25 in het basisonderwijs en 22 in het secundair. Dit betekent dat er van de deelnemende scholen van het basisonderwijs twee scholen zijn waarvan er geen respons is van de directeur. In het secundair onderwijs zijn er acht scholen waarvan de vragenlijst voor directie niet beantwoord werd.

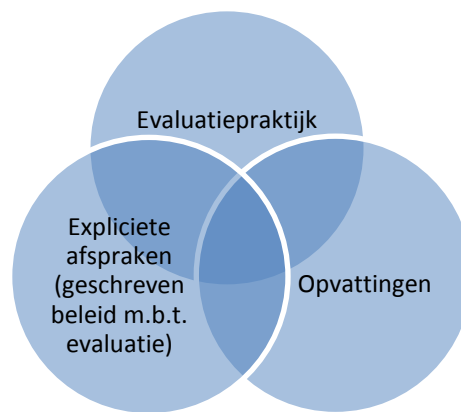
**Tabel 2.4. Overzicht respons voor basis en secundair onderwijs bij surveyonderzoek**

	<b>Gerealiseerde steekproef</b>	<b>Effectieve respons van individuen</b>
<b>Basisonderwijs</b>	27	417 leerkrachten 25 directies
<b>Secundair onderwijs</b>	29	686 leerkrachten 22 directies
<b>Totaal</b>	56	1120 leerkrachten 47 directies

### III De resultaten gebundeld

#### 1. Model voor evaluatiebeleid

Vooraleer in te gaan op de resultaten die voortvloeien uit de drie fases van het onderzoek, willen we het kader toelichten waarbinnen we de resultaten ordenen. Het voorwerp van onderzoek vormt het evaluatiebeleid van scholen. Dit evaluatiebeleid dienen we ruim te interpreteren: als het samengaan van drie componenten: evaluatiepraktijken van leerkrachten, opvattingen omtrent evaluatie en expliciete afspraken (geschreven beleid) op schoolniveau omtrent evaluatie.



Het evaluatiebeleid kan best onderzocht worden door aandacht te hebben voor het samengaan van deze drie componenten. Leerkrachten en directies houden er immers opvattingen op na omtrent hoe evaluatie zou moeten zijn. Deze opvattingen oefenen ook een invloed uit op de evaluatiepraktijk zelf. De evaluatiepraktijk zelf verwijst dan weer naar wat leerkrachten precies doen om hun leerlingen te evalueren. Het geëxpliciteerde evaluatiebeleid of de afspraken hieromtrent op schoolniveau kunnen deze opvattingen ondersteunen of net ingaan tegen de ideeën die leerkrachten erop nahouden. Het evaluatiebeleid van een school kan dus niet verengd worden tot louter het geëxpliciteerde beleid omtrent evaluatie, maar moet gezien worden als het samengaan van deze drie componenten die elkaar onderling beïnvloeden.

Om tegemoet te komen aan deze complexiteit wordt geopteerd om dit model voor 'evaluatiebeleid' te hanteren bij het weergeven van de resultaten van de triangulatie. In wat volgt, zal opeenvolgend ingegaan worden op evaluatiepraktijk van leerkrachten, opvattingen en ten slotte evaluatiebeleid van scholen. Waar dit aan de orde is worden er kruisverwijzingen gemaakt tussen de verschillende onderdelen.

Hoewel we hiermee de structuur van de onderzoeksvragen verlaten, stelt het ons beter in staat om de complexiteit van evaluatiepraktijken van leerkrachten en het evaluatiebeleid van scholen weer te geven. De gebundelde resultaten vatten we in sectie 5 samen door een antwoord te formuleren op de vier centrale onderzoeksvragen.

# 1. Evaluatiepraktijk van leerkrachten

## 1.1. Beperkte inbedding van evaluatie in de algemene onderwijspraktijk

In het onderzoek werd gepeild naar de mate waarin evaluatie-activiteiten van leerkrachten ingebed worden in hun dagdagelijkse onderwijspraktijk. Op basis van de literatuurstudie (Ysenbaert et al., 2017) werd een theoretisch onderscheid gemaakt tussen (1) het gescheiden model, waarbij de instructieactiviteiten sterk gescheiden zijn van de evaluatieactiviteiten, (2) het semi-geïntegreerde model waarin er een matige verwevenheid is van beide activiteiten en (3) het geïntegreerde model waarbij evaluatie sterk geïntegreerd is in de algemene onderwijspraktijk.

Het surveyonderzoek laat zien dat leerkrachten hun eigen evaluatiepraktijk het meest herkennen in het gescheiden model met een gemiddelde score van 5,11 op een schaal van nul tot acht voor leerkrachten van het basisonderwijs en een score van 5,51 bij leerkrachten secundair onderwijs. Over het algemeen sluiten leerkrachten minder aan bij het semi-geïntegreerde model: gemiddeld 3,81 voor basisonderwijs en 3,67 voor secundair onderwijs. Het geïntegreerde model is het minst kenmerkend voor de evaluatiepraktijk van leerkrachten: gemiddeld 3,04 voor basis en 2,70 voor het secundair onderwijs. Deze bevindingen bevestigen de trend die ook vastgesteld werd in het casestudieonderzoek. Hetzelfde geldt bovendien voor de vastgestelde verschillen tussen basis- en secundair onderwijs waarbij zowel het kwantitatief als het kwalitatief onderzoek laten zien dat leerkrachten in het basisonderwijs er meer dan hun collega's in het secundair onderwijs in slagen om evaluatiepraktijken in te bedden in de onderwijspraktijk.

Het mixed-method-design van dit onderzoek levert verder een dieper inzicht op met betrekking tot de evaluatiepraktijken die gerepresenteerd worden door deze drie modellen. We lichten verder toe hoe de modellen samenhangen met de mate waarin leerlingen betrokken worden, de aard van de gebruikte evaluatievormen en de mate waarin er aandacht is voor verantwoording en impact.

### 1.1.1. Leerlingen worden meer betrokken bij (semi-)geïntegreerde evaluatiepraktijk

Volgens de meervoudige lineaire regressies (Ysenbaert et al., n.d.) is er voor elk van de drie evaluatiemodellen een verband met de mate waarin leerlingen betrokken worden bij evaluatie. Concreet blijkt dat hoe meer een leerkracht de evaluatieactiviteiten inbedt in de algemene onderwijspraktijk, hoe hoger de kans is dat de leerlingen zelf hier ook bij betrokken worden.

#### *Betrokkenheid leerlingen bij planning*

In de casestudies zien we dat de betrokkenheid van leerlingen eerder beperkt is als het gaat over de planning van evaluaties. Over het algemeen wordt de planning aangestuurd door de leerkracht. De gesprekken met leerkrachten tonen dat de wijze waarop het onderwijs georganiseerd wordt op school de planning sterk aanstuurt. Ze zien leerlingen soms maar één lesuur per week en/of er is de afspraak op schoolniveau om tegen elke rapportperiode over een minimum aantal evaluaties te beschikken. Dit maakt dat de werkkalender van de school en bijgevolg van de leerkracht soms primeert boven de inspraak van leerlingen bij de planning van evaluaties. De methodeschool in het basisonderwijs vormt hierbij een uitzondering aangezien de leerlingen zelf aangeven wanneer ze zich klaar voelen om deel te nemen aan een evaluatie. In deze school werd bovendien een sterk gedeelde en geïntegreerde evaluatiepraktijk vastgesteld.



### ***Betrokkenheid leerlingen bij bepalen van criteria***

Bij de meerderheid van de bevraagde leerkrachten in de casestudies is het zo dat de evaluatiecriteria bepaald worden door de leerkracht zelf. In het basisonderwijs stellen we vast dat dit in de cases die gekenmerkt worden door meer diversiteit vaker gebeurt en vooral als het gaat om het meebepalen van criteria met betrekking tot het geven van een presentatie. Hetzelfde zien we gebeuren bij leerkrachten Nederlands van de derde graad secundair onderwijs. De methodeschool in het basisonderwijs vormt opnieuw de uitzondering en zet het sterkste in op het betrekken van de leerlingen bij het meebepalen van de evaluatiecriteria, wat de bevinding uit het surveyonderzoek bevestigt dat een geïntegreerd model voor evaluatiepraktijk samen gaat met een hoge mate van leerlingenbetrokkenheid.

### ***Betrokkenheid leerlingen via de vorm van evaluatie***

(Semi-)geïntegreerde evaluatiepraktijken worden volgens het casestudieonderzoek vaker gekenmerkt door peer- en zelfevaluatie en genereren hierdoor een hogere betrokkenheid van leerlingen. Dergelijke evaluatievormen blijken meer in het basis- dan in het secundair voor te komen. In het kleuteronderwijs worden kinderen spontaan gestimuleerd om aan zelf- en peerevaluatie te doen. In het lager onderwijs komt zelf- en peerevaluatie iets meer gefragmenteerd aan bod. De gesprekken met leerkrachten tonen aan dat leerlingen zichzelf en hun medeleerlingen vooral evalueren met betrekking tot attitudes en/of de leergebieden wereldoriëntatie en muzische opvoeding. De wijze waarop zelfevaluatie ingevuld wordt, varieert: mondeling terugblikken op een activiteit of op papier aanduiden via het inkleuren van bolletjes met groen, oranje, rood of het aanduiden van smileys. Inhoudelijk blijkt dat de 'zelfevaluatie' eerder oppervlakkig benaderd wordt. Zo wordt er doorgaans gepeild naar de mate waarin de leerlingen een opdracht of een thema binnen wereldoriëntatie of muzische opvoeding leuk vonden en niet zozeer de mate waarin ze hun eigen prestaties of leerproces inschatten. De methodeschool geeft hier een andere invulling aan: zelfevaluatie wordt meer gekoppeld aan concrete leerdoelen. Peerevaluatie komt in het basisonderwijs frequenter voor dan zelfevaluatie, al is dit in sommige cases ook nog heel beperkt. Vaak gaat het hierbij om het beoordelen van elkaars eindproduct, of indien het om een groepswerk gaat, worden de houding en bijdragen van de groepsleden geëvalueerd.

In het secundair onderwijs zijn zelf- en peerevaluatie over het algemeen minder in gebruik. Indien leerlingen zichzelf en/of elkaar evalueren in het secundair dan is dit vooral met het oog op de vaardigheden en attitudes. In tegenstelling tot de cases het basisonderwijs maken de leerkrachten in het secundair onderwijs dan weer vaker gebruik van zelfevaluatie. Peerevaluatie wordt als minder geschikt beschouwd omwille van tijdsgebrek of omdat men de leerlingen hiervoor nog te jong vindt. Een redenering die ietwat haaks staat op de vaststelling dat leerlingen in het lager onderwijs, en zelfs in het kleuter al leren om elkaar te evalueren.

### **1.1.2. Rol van observatie & kwaliteitsvolle feedback bij een meer ingebedde evaluatiepraktijk**

De modellen voor evaluatiepraktijk (Ysenbaert et al., 2017) geven een continuüm weer voor de mate waarin evaluatie ingebed is in de ruimere onderwijspraktijk: het gescheiden model geeft enerzijds een strikte scheiding weer en het geïntegreerde model representeert een sterke integratie van evaluatieactiviteiten binnen de alledaagse klaspraktijk. Op basis van triangulatie van de literatuurstudie (Ysenbaert et al., 2017), het casestudieonderzoek (Ysenbaert et al., 2018) en het surveyonderzoek (Ysenbaert et al., n.d.) vinden we bevestiging voor de hypothese dat observaties en

kwaliteitsvolle feedback bijdragen tot het integreren van evaluatieactiviteiten in de algemene onderwijspraktijk.

Het surveyonderzoek laat zien dat het (semi-)geïntegreerde model vaker deel uitmaakt van de evaluatiepraktijken in het basisonderwijs dan in het secundair onderwijs. Tegelijkertijd laat het casestudieonderzoek zien dat informele evaluatie eerder onderdeel is van de alledaagse klaspraktijk in het basisonderwijs, en misschien wel nog het sterkste in het kleuteronderwijs. De kleuterleerkrachten wijzen erop dat ze continu informeel evalueren, doorheen al hun dagelijkse activiteiten met de kinderen. Op basis van hun observaties bepalen ze welke activiteiten er gedaan worden, geven ze onmiddellijke feedback aan de kinderen om kort op de bal te kunnen spelen. Hun observaties houden ze bij in een schriftje, notitieboekje of zorgschrift. In het lager onderwijs houdt de meerderheid van de leerkrachten eveneens een (zorg)schriftje of map bij waarin ze aantekeningen maken op basis van hun observaties. Daarnaast geven de leerkrachten uit het lager aan dat er ook informeel geëvalueerd wordt doorheen de lessen via interactie met leerlingen bijvoorbeeld door het stellen van gerichte vragen. In het secundair onderwijs bevestigt de meerderheid van de leerkrachten dat observeren ook deel uitmaakt van hun klaspraktijk en dat ze dit zien als een controle om te zien of leerlingen mee zijn met de les. De observaties in het secundair gebeuren doorgaans op een informele wijze, zonder dat hieromtrent specifieke notities worden bijgehouden. Het bso vormt hierop een uitzondering, waarbij er meer doelgerichte observaties vastgesteld werden, voornamelijk in de praktijkvakken waarbij geëvalueerd wordt met behulp van vooropgestelde evaluatiecriteria die weergegeven worden in observatiewijzers. Dit laatste weerspiegelt ook het negatieve verband dat gevonden werden bij de analyses van de surveydata tussen de onderwijsvorm bso en een gescheiden evaluatiepraktijk van leerkrachten.

Observatie biedt leerkrachten de mogelijkheid om vast te stellen hoe leerlingen het er van afbrengen in hun leerproces. Het levert informatie op die aangewend kan worden om de volgende instructie af te stemmen op daar waar de leerlingen staan. Dit keert terug in de multivariate analyses waarbij er een positieve samenhang werd vastgesteld tussen een semi-geïntegreerde evaluatiepraktijk en het belang dat men hecht aan de impact die evaluatie heeft op het uiteindelijke leergedrag van leerlingen. Volgens Hattie (2012) speelt feedback hierin een cruciale rol, wat tevens blijkt uit de respons van leerlingen die bevraagd werden in het kader van het casestudieonderzoek. Zo blijkt hun voorkeur te gaan naar 'observatie door de leerkracht' net omdat dit leidt tot onmiddellijke bijsturing. De mate waarin er feedback volgt op toetsen en examens is volgens hen dan weer sterk wisselend van leerkracht tot leerkracht. Sommige leerkrachten komen er niet toe om de feedback verder te laten gaan dan het meegeven van de correcte oplossingen, volgens hen is dit te wijten aan tijdsgebrek.

De behoefte aan feedback keert ook terug in de vraag van leerlingen in het secundair onderwijs om meer in te zetten op peerevaluatie dan zelfevaluatie omdat ze veel meer in return krijgen: de woordelijke feedback, de tips om het in het vervolg anders en beter aan te pakken, zichzelf leren kennen met betrekking tot hoe ze overkomen bij medeleerlingen tijdens groepswork vinden ze bijvoorbeeld heel waardevol. Het feit dat de leerlingen in mindere mate te vinden zijn voor zelfevaluatie heeft dan weer te maken met het gebrek aan feedback. De oorzaak hiervoor ligt vooral in het feit dat 'zelfevaluatie' als éénrichtingsverkeer ingevuld wordt waardoor feedbackkansen onderbenut blijven.

De casestudies lieten al zien dat vormen van peer- en zelf-evaluatie frequenter gehanteerd worden in het basisonderwijs dan in het secundair onderwijs (zie ook 1.1.1), wat opnieuw de samenhang met meer geïntegreerde vormen van evaluatie bevestigt.

### **1.1.3. Gescheiden evaluatiepraktijk: verantwoording en toetsen & examens**

#### ***Eénzijdige invulling van evaluatie door middel van toetsen en examens***

De klassieke pen-en-papier toets blijft een heel courante manier van evalueren in de onderzochte cases van zowel basis als secundair onderwijs. Vaak is het een afspraak binnen een lagere school om zoveel mogelijk de toetsen uit de methode te gebruiken omdat de vraagstelling vergelijkbaar is met deze die in de handboeken gebruikt wordt. Meestal passen de leerkrachten de bestaande toetsen aan om de afstemming met het leerplan te bewaken of om tegemoet te komen aan de noden van de leerlingen, zoals bijvoorbeeld de duurtijd inperken bij lange herhalings-toetsen. In het secundair onderwijs wordt er eveneens frequent gebruik gemaakt van de klassieke toets. Al lijkt het erop dat veel leerkrachten deze zelf opstellen omdat er vaak geen bestaande toetsen voor handen zijn, of men laat zich inspireren door bestaande toetsen. Het gebruik van examens is vooral een praktijk die aan de orde is in het secundair onderwijs en dit twee tot drie keer per schooljaar.

Dit courante gebruik van klassieke pen-en-papier-toetsen is kenmerkend voor een gescheiden evaluatiepraktijk. Dit is in tegenstelling tot enerzijds de atypische school binnen de onderzochte cases van het basisonderwijs waar helemaal niet met toetsen geëvalueerd wordt en waar geen enkele leerkracht aanleunt bij het gescheiden model. Anderzijds blijkt de gescheiden evaluatiepraktijk volgens de multilevelanalyses minder aan de orde te zijn in het bso. Leerkrachten in scholen met een bso-aanbod zijn minder geneigd om hun evaluatieactiviteiten te scheiden van instructieactiviteiten. Dit kunnen we verklaren door de aard van de praktijkgerichte vakken waarin vaardigheden centraal staan en die een meer continue evaluatie vergen om ze te ontwikkelen. Verder blijkt dat er in het bso en kso ook meer gebruik wordt gemaakt van permanente evaluatie in plaats van (uitsluitend) examens. Het belang van examens wordt in de cases overigens ook sterker benadrukt in het aso dan in de onderwijsvormen tso en bso.

De bevindingen in het kwalitatieve en kwantitatieve onderzoek in zowel het basis- als het secundair onderwijs bevestigen de hypothese dat een gescheiden model voor evaluatiepraktijk gekenmerkt wordt door een sterkere nadruk op gebruik van toetsen en examens.

#### ***Aandacht voor verantwoording evaluatie***

Er is een positief verband tussen het gescheiden model en de frequentie waarmee er leerlingencontacten georganiseerd worden op school: leerkrachten op scholen die dit meer dan twee keer per schooljaar organiseren lijken vaker 'gescheiden' te evalueren. Op het eerste gezicht lijkt dit een vreemde vaststelling: leerlingcontacten zorgen er namelijk voor dat leerlingen nauwer betrokken worden (Mattheeuws, z.d.) waardoor we verwachten dat dit eerder zou samenhangen met een semi-geïntegreerd of geïntegreerd model voor evaluatiepraktijk. Een verklaring voor deze bevinding ligt mogelijks in de wijze waarop dit leerlingencontact ingevuld wordt. Het casestudieonderzoek liet bijvoorbeeld zien dat leerlingencontacten slechts beperkt onderdeel zijn van de evaluatiepraktijk: slechts drie van de twaalf onderzochte scholen maken hier werk van en bovendien blijkt de wijze waarop dit ingevuld wordt heel leerkrachtafhankelijk. Aangezien het leerlingencontact sterk samenhangt met een gescheiden evaluatiepraktijk vermoeden we dat dit vaak oppervlakkig ingevuld wordt: eerder als wijze om te communiceren over het leerproces in plaats van het uitvoeren van

impact op dit leerproces. Deze verklaring wordt bovendien ondersteund door het significante verband dat gevonden werd met de opvattingen van leerkrachten: leerkrachten die sterk overtuigd zijn van de informerende waarde van evaluatie zijn meer geneigd om hun evaluatiepraktijk in te richten volgens het gescheiden model. In die zin komt het leerlingcontact dicht in de buurt van een oudercontact dat heel vaak georganiseerd wordt na een toetsen- of examenperiode om de resultaten van de leerling mee te delen. De nadruk lijkt meer te liggen op het 'summatief' meedelen van resultaten dan op het 'formatief' invullen van een leerlingcontact. De vaststelling dat een gescheiden evaluatiepraktijk samenhangt met een verhoogde aandacht voor het verantwoorden van de uitkomsten van evaluatie lijkt deze verklaring te ondersteunen.

## 1.2. Evaluatie op momenten van transitie

### 1.2.1. Transitie tussen leerjaren in basisonderwijs

De overgang van het ene leerjaar naar het andere leerjaar wordt in het basisonderwijs besproken op een klassenraad. We stellen veel verschillen vast in de wijze waarop dit aangepakt wordt in de onderzochte cases van het basisonderwijs (Ysenbaert et al., 2018). Een eerste verschil is wie betrokken wordt bij deze beslissing: een minderheid streeft ernaar om dit met het hele team van leerkrachten te doen. In deze scholen ervaart men daarenboven vaak praktische en organisatorische hinder om dit te realiseren waardoor het in de praktijk alsnog met een beperkt aantal leerkrachten doorgaat. Op vlak van gehanteerde criteria besluiten we dat er hieromtrent geen heldere communicatie is. In de meerderheid van de cases wordt er weinig tot niets hierover vermeld in de beleidsdocumenten, met uitzondering van de methodeschool die elke leerling wil laten overgaan naar het volgende leerjaar. Uit de gesprekken lijken de bevroegde leerkrachten het erover eens te zijn dat men bij momenten van doorstroom de leerling *breed* beoordeelt: niet alleen cognitieve elementen worden meegenomen, maar ook de motivatie, het sociaal-emotioneel functioneren en het welbevinden. Op dit vlak lijken er op het eerste gezicht weinig verschillen te zijn tussen de cases, hoewel bepaalde uitspraken van leerkrachten laten vermoeden dat er in realiteit net wel een zwaarder gewicht toegekend wordt aan de cognitieve prestaties van leerlingen. Gelijkaardige verschillen zijn merkbaar op vlak van houding ten aanzien van zittenblijven: waar men in de methodeschool elke leerling wil laten overgaan, omschrijft men zittenblijven in een andere school als '*soms een absolute noodzaak*'.

### 1.2.2. Getuigschrift basisonderwijs & oriëntering naar a/b-stroom

Op basis van het casestudieonderzoek stellen we vast dat ook het beleid en de praktijk inzake het toekennen van het getuigschrift basisonderwijs sterk uiteenlopend is. Op vlak van het formele beleid dat weergegeven wordt in de beleidsdocumenten en schoolreglementen zijn de verschillen eerder beperkt. Eén school (case E) is bijzonder op het vlak van beleid aangezien elke leerling die het zesde leerjaar op die school doorloopt sowieso het getuigschrift ontvangt. Voor de overige cases kan worden samengevat dat het formele beleid een weergave is van de vastgelegde regelgeving. In de praktijk blijken de overige basisscholen onderling wel nog te verschillen in de manier waarop beslist wordt wie wel/geen getuigschrift krijgt. Twee vormen van aanpak kunnen onderscheiden worden. De eerste, *enge* manier wordt gekenmerkt door een primaire rol van een cijfercriterium in combinatie met een delibererend overleg in beperkte kring waarbij enkel de directeur en/of zorgleerkracht samen met de leerkracht van het zesde leerjaar beslist over deze leerlingen die het cijfercriterium niet halen. Het cijfercriterium wordt daarenboven verschillend ingevuld: zo geldt in de ene school (case D) een drempel van gemiddeld 50% bij elk van de vakken Nederlands, Frans en wiskunde, waar in een andere school (case B) de drempel 60% is voor dezelfde vakken en bijkomend nog het vak wereldoriëntatie.

De tweede manier wordt gekenmerkt door een *brede* aanpak waarbij het hele team van leerkrachten betrokken wordt én waarbij tegelijkertijd het totale plaatje van de leerling meegenomen wordt. Een mogelijke, voorzichtige, verklaring voor beide vormen van aanpak zou kunnen liggen in de samenstelling van het leerlingenpubliek. De cases die werken volgens de tweede, *brede* manier worden gekenmerkt door veel diversiteit en/of het beleid van deze scholen omarmt de aanwezige diversiteit. De diversiteit vormt het uitgangspunt, waardoor het een logische verklaring kan zijn waarom deze scholen bij het uitreiken van het getuigschrift veel *breder* kijken naar hun leerlingen. Twee van de drie overige cases, die werken volgens de *engere* vorm om het getuigschrift toe te kennen, worden daarentegen gekenmerkt door een homogene leerlingenpopulatie, waarbij de meerderheid van de leerlingen uit een blank, Nederlandstalig, middenklasse gezin komt.

De manier waarop leerlingen georiënteerd worden naar de a- of de b-stroom wordt voor een groot deel meebepaald door de besluitvorming omtrent het getuigschrift. Wie geen getuigschrift krijgt, gaat doorgaans door naar 1b. Op het moment van de dataverzameling was 1a ook nog een optie, wat inmiddels niet meer het geval is. We stellen vast dat de oriëntering naar de b-stroom op deze manier een negatieve keuze is op het moment van de dataverzameling: wanneer een leerling niet alle doelen bereikt heeft of niet het zesde leerjaar doorlopen heeft, zal hij/zij georiënteerd worden naar de b-stroom. Met andere woorden, het niet-succesvol afronden van het lager onderwijs resulteert in een secundaire schoolloopbaan in de b-stroom en later in het bso (Goosen et al., 2017). Of anders gezegd: de keuze voor b-stroom is niet zozeer een keuze maar een gegeven bij onderwijsachterstand. Hier zien we dat de mechanismen van het watervalstelsel al opduiken nog vóór de leerlingen starten in het secundair onderwijs.

Opvallend is dat de meeste scholen in het kader van de oriëntering van leerlingen al op het einde van het tweede trimester een advies formuleren – veelal door de leerkracht van het zesde leerjaar en een CLB-medewerker – en dit wordt gecommuniceerd naar ouders en leerling. In twee van de zes onderzochte basisscholen is er vooral bij de directies ontevredenheid over hoe de oriëntering verloopt en wenst men hier in de toekomst meer tijd in te investeren. Zo wil men dit grondiger aanpakken en ook de verschillende partners (ouders, leerlingen, secundaire scholen) bij betrekken.

### **1.2.3. Attestering in secundair onderwijs & (her)oriëntering**

Het casestudieonderzoek toont aanzienlijke verschillen in de deliberatiepraktijk van de zes secundaire scholen. De verschillen tussen scholen in het secundair lijken groter te zijn dan deze vastgesteld in het lager onderwijs. Deze verschillen doen zich voor op vlak van gegevensverzameling, bijvoorbeeld de mate waarin er gekeken wordt naar dagelijks werk enerzijds en de examenresultaten anderzijds. Daarnaast verschillen de scholen uit het casestudieonderzoek ook in de criteria die gehanteerd worden en de mate waarin deze doorslaggevend zijn voor de deliberatiebeslissing, zoals bijvoorbeeld de aandacht die geschonken wordt aan het gedrag van een leerling. Deze bevindingen bevestigen wat eerder onderzoek, hoewel uitgevoerd in de Franstalige gemeenschap in België, aantoonde, namelijk dat scholen onderling verschillen in de wijze waarop deliberatiebeslissingen genomen worden (Anthoons, Deprez, Thyssen, & Verbruggen, 2004). Deze variatie aan deliberatiecriteria kan verklaard worden door de pedagogische vrijheid die scholen hebben – naast de algemene richtlijnen – om te evalueren in welke mate de eindtermen, ontwikkelingsdoelen en leerplandoelen bereikt zijn (Teijssen, De Fraine, & Wouters, 2017).

Bovendien wordt het deliberatiebeleid niet steeds gedragen door het team, zoals in case F waar de leerkrachten het oneens zijn met het beleid van de directie dat de nadruk legt op het uitreiken van A-attesten, ook al zijn er veel tekorten bij de resultaten van een leerling. Hoewel de casestudies het niet toelaten om veralgemenende uitspraken te doen, zijn er toch wel aanwijzingen binnen deze cases die pleiten voor een beleid dat sterk inzet op communicatie en gedeeld leiderschap om tot een gedragen deliberatiebeleid te komen waarin éénieder zich kan vinden.

De (her)oriëntering van leerlingen in het secundair onderwijs krijgt beleidsmatig beperkte aandacht in de meerderheid van de onderzochte scholen. De oriënterende functie van de klassenraad blijkt in de praktijk doorgaans gestuurd te worden door het watervalstelsel. Oriëntering lijkt pas aan de orde te komen wanneer men vaststelt dat de slaagkansen voor volgend schooljaar in dezelfde studierichting klein zijn. De oriëntering is in die zin vooral een *negatieve* oriëntering. Deze trend die eigenlijk al ingezet wordt op het einde van het lager onderwijs (zie 1.2.2), wordt dus doorgetrokken in het secundair onderwijs. Dit sluit aan bij eerdere bevindingen van Nicaise, Spruyt, Van Houtte, and Kavadias (2014) die aantonen dat leerkrachten eerder op basis van prestaties en studiehouding adviseren dan op basis van belangstelling of talenten van leerlingen. In één case (aso-school) gaat het zelfs nog een stap verder: de oriëntering fungeert via de adviezen als een selectiemechanisme. Zo wordt bij een a-attest alsnog een advies opgenomen waarin de leerling gestimuleerd wordt om het aso te verlaten. Op deze manier houdt deze school enkel de sterkste leerlingen over.

### 1.3. Evalueren met gevalideerde instrumenten

Volgens het surveyonderzoek maken alle scholen gebruik van gevalideerde evaluatie-instrumenten. De casestudies laten echter zien dat het gebruik van deze instrumenten heel miniem én oppervlakkig ingevuld wordt. In de cases van het basisonderwijs worden gevalideerde toetsen onder stimulans van de directie afgenomen. Daarenboven blijkt de meerderheid van de directies hiervoor te opteren omwille van druk van buitenaf (bijvoorbeeld de verplichte taalscreening), of omwille van het feit dat het een stand van zaken oplevert (zonder dat hier verder iets mee gedaan wordt). De bevraagde leerkrachten in het basisonderwijs lijken hieromtrent een onverschillige houding aan te nemen. Er is slechts één case in het basisonderwijs waarbij het gebruik van gevalideerde instrumenten bedachtzaam gebeurt. In case C geven deze instrumenten en vooral de resultaten ervan richting aan het beleid van de school. De directie zet deze instrumenten in in functie van interne kwaliteitszorg: via het herhaaldelijk meten wordt het eigen onderwijs gemonitord én bijgestuurd. De directie laat zich hierdoor leiden om prioriteiten te stellen in het beleid en om met het team in te tekenen op professionaliseringsinitiatieven die specifiek inzetten op het wegwerken van vastgestelde tekorten.

In het secundair onderwijs stellen we een gelijkaardige trend vast in het casestudieonderzoek: enkel de verplichte taalscreening bij aanvang van het secundair onderwijs wordt afgenomen en doorgaans is dit een taak van de leerkracht Nederlands. De opvolging en remediëring die hierop zou moeten volgen, wordt heel beperkt ingevuld.

In de meerderheid van de cases worden er genormeerde instrumenten gebruikt, maar zelden vanuit de meerwaarde die dit kan hebben voor de kwaliteitsontwikkeling. De verdere opvolging ervan is bijgevolg onbestaande of slechts heel miniem. De *good practice* in het casestudieonderzoek onderstreept de cruciale rol van de directie bij het functioneel inzetten van genormeerde evaluatie-instrumenten.

## 1.4. Evaluatiepraktijk & diversiteit

### 1.4.1. (Semi-)geïntegreerde evaluatiepraktijk & diversiteit

De literatuurstudie liet zien dat het geïntegreerde model voor evaluatiepraktijk het sterkst de diversiteit onder leerlingen erkent. Dit is het gevolg van het feit dat het geïntegreerde evaluatiemodel voortvloeit uit het constructivisme dat, in tegenstelling tot het behaviorisme en cognitivisme, tegemoet komt aan diversiteit onder leerlingen. Aangezien het constructivisme aandacht heeft voor de individuele leerling en de wijze waarop hij/zij kennis en vaardigheden ontwikkelt, spreekt het voor zich dat het leerproces als een sterk geïndividualiseerde aangelegenheid beschouwd wordt. Binnen het constructivisme ligt de klemtoon van evaluatie precies op de stimulerende rol die het heeft voor het leerproces van de leerling. Evaluatie is hierin cruciaal precies omdat evaluatie het mogelijk maakt om voortdurend na te gaan waar de leerling staat.

Dit komt ook tot uiting in de multilevelanalyses: er is samenhang tussen de mate waarin leerkrachten differentiëren in hun algemene klaspraktijk en de mate waarin ze evaluatiepraktijken integreren in hun algemene onderwijspraktijk. Leerkrachten die veel aandacht hebben voor differentiatie, en dus gericht inspelen op de specifieke leerbehoeften van leerlingen, zullen eerder een (semi-)geïntegreerde evaluatiepraktijk hanteren. Het gescheiden model voor evaluatie laat daarentegen een negatief verband zien: hoe minder aandacht voor differentiatie tijdens de les, hoe meer evaluatieactiviteiten gescheiden worden van onderwijsactiviteiten. Het feit dat de evaluatiepraktijk van leerkrachten over het algemeen meer gekenmerkt wordt door een gescheiden evaluatiepraktijk in plaats van een semi-geïntegreerde of geïntegreerde evaluatiepraktijk, betekent dat het aanwenden van evaluatie als input voor het onderwijsproces eerder beperkt is. Deze situatie zorgt ervoor dat groeikansen mogelijks onderbenut blijven voor leerlingen die de vooropgestelde leerdoelen nog niet bereikt hebben. Eveneens belemmert dit mogelijks leerlingen die het potentieel hebben om hun competenties te ontwikkelen tot een niveau dat verder reikt dan de vooropgestelde leerdoelen.

Verder blijkt dat wie zich bekwaam voelt om diversiteit onder leerlingen op te merken eerder zal inzetten op geïntegreerde evaluatiepraktijken.

Naast de leerkrachtkenmerken, zoals de mate waarin ze differentiëren en ze diversiteit onder leerlingen waarnemen, is er ook een specifiek schoolkenmerk dat relevant blijkt inzake diversiteit: de leerlingencompositie. Hoe meer het leerlingenpubliek van een school gekenmerkt wordt door een hoge SES, hoe minder leerkrachten op deze school een (semi-)geïntegreerde evaluatiepraktijk zullen aanwenden. Dit betekent dat hoe meer lage SES-leerlingen het leerlingenpubliek vertegenwoordigen, hoe meer leerkrachten de uitkomsten van evaluatie zullen aanwenden als input voor hun eigen onderwijspraktijk. Dit zorgt ervoor dat de instructie afgestemd wordt op waar de lerende staat in zijn/haar leerproces en/of dat er aan de hand van feedback concrete handvatten aangereikt worden waarmee de lerende verder kan. Ten aanzien van leerlingen met een lage SES is dit een positieve vaststelling. Tegelijkertijd kunnen we hier ook uit afleiden dat een leerlingenpubliek dat vooral gekenmerkt wordt door een hoge SES, veel minder kan rekenen op instructieactiviteiten die afgestemd worden op de uitkomsten van de leerlingenevaluaties. Deze bevinding kan op twee manieren geïnterpreteerd worden. Enerzijds kunnen leerkrachten het niet nodig vinden om de instructie nog af te stemmen op basis van de evaluaties voor deze groep van leerlingen omdat men (mede door de hoge verwachtingen die men heeft ten aanzien van deze hoge SES-leerlingen) ervan uitgaat dat deze leerlingen zelf een extra inspanning kunnen leveren om de beoogde doelen te bereiken. Deze



redenering sluit aan bij het verantwoordingsgerichte aspect dat leerkrachten, die eerder aansluiten bij het gescheiden evaluatiemodel, belangrijk vinden. Hoge SES-leerlingen worden in dit geval sterker aangesproken op hun verantwoordelijkheid. Anderzijds kan dit verwijzen naar het feit dat leerkrachten menen dat de beoogde doelen al behaald zijn. Dit laatste zou betekenen dat differentiatie en uitdaging die verder gaat dan de beoogde doelen, specifiek voor hoge SES-leerlingen, weinig aandacht krijgen van leerkrachten.

#### **1.4.2. Minder aandacht voor diversiteit in hogere leerjaren.**

Het rekening houden met diversiteit bij evaluatie krijgt meer aandacht in het basisonderwijs dan in het secundair onderwijs. Dit blijkt uit zowel het survey- als het casestudieonderzoek dat aantoont dat (semi-)geïntegreerde evaluatiepraktijken vaker voorkomen in het basisonderwijs.

In vrijwel alle cases van het basisonderwijs is het duidelijk dat leerlingen die nood hebben aan ondersteuning op momenten van evaluatie hier op kunnen rekenen. De manier waarop er beslist wordt wie hier wel of geen aanspraak op kan maken is binnen dit casestudieonderzoek niet geheel scherp gesteld. Het lijkt er sterk op dat dit bepaald wordt op basis van aanvoelen van de klasleerkracht. Het is lang niet zo dat een attest van een leerling een voorwaarde vormt om ondersteuning te bieden. Bovendien worden er in het basisonderwijs iets vaker klassikale maatregelen genomen om tegemoet te komen aan de noden van leerlingen dan in het secundair onderwijs, bijvoorbeeld door *alle* leerlingen voldoende tijd te geven en door de toets of het examen mondeling te overlopen met de leerlingen, zodat er eventueel nog vragen ter verduidelijking gesteld kunnen worden. In het secundair onderwijs lijkt het rekening houden met diversiteit op momenten van evaluatie minder evident. De praktijken worden er meer afgestemd op het 'label' dat een leerling heeft. Dit gebeurt voornamelijk door de leerlingen die een attest hebben meer in het oog te houden en hen extra te ondersteunen via de sticordi- of redicodimaatregelen. Het 'rekening houden met diversiteit' wordt sterker benaderd vanuit een 'deficit-denken'. De leerlingen met bijvoorbeeld een leerstoornis of met een andere thuistaal moeten bijkomend ondersteund worden of geremedieerd worden. Het gaat hierbij nooit over leerlingen die vanuit hun sterktes of talenten een andere aanpak nodig zouden hebben of een andere invulling van hun opdrachten zouden moeten krijgen. Hierdoor wordt duidelijk dat 'diversiteit' in hoofden van leerkrachten als 'negatief' gepercipieerd wordt omdat het vooral ingevuld wordt als 'leerlingen die *niet* meekunnen, die *geen* aandacht kunnen houden, het Nederlands *niet* kunnen', ... .

Een beperkter aantal leerkrachten uit de cases van het secundair geeft aan dat ze niet of weinig rekening houden met diversiteit omdat ze niet weten hoe ze dit moeten doen, omdat het te veel tijd vraagt of omdat men ernaar streeft om het leerplan af te werken met (de meerderheid van) de klasgroep.

De methodeschool uit het basisonderwijs (case F) vormt opnieuw de uitzondering. De wijze waarop deze school omgaat met diversiteit bij evaluatie vertrekt vanuit een heldere en gedragen onderwijsvisie. In deze onderwijsvisie wordt er uitgegaan vanuit de idee dat diversiteit een gegeven is en aangezien het beleid en de praktijk met betrekking tot evaluatie hierop gestoeld is, wordt diversiteit op dit vlak beschouwd als de norm en niet als een deficit. Het leidt tot een evaluatiepraktijk die nauw verweven is met de onderwijspraktijk en waarin leerlingen heel sterk betrokken worden. Het is bovendien opvallend dat de evaluatiepraktijk veel gelijkenissen vertoont met de gerapporteerde evaluatiepraktijken in het kleuteronderwijs, waarin nog sterker uitgegaan wordt vanuit het feit dat diversiteit onder leerlingen een gegeven is.



Gedifferentieerd evalueren aan de hand van differentiatie in doelen is volgens het casestudieonderzoek eerder uitzonderlijk. In het basisonderwijs gebeurt het iets vaker, wat we verklaren door het feit dat de eindtermen geformuleerd zijn voor het einde van het zesde leerjaar. Dit biedt meer vrijheid om gedurende de zes voorafgaande leerjaren een leerling voor één of meerdere leergebieden te laten meevolgen in een hoger of lager leerjaar. Tegelijkertijd stellen we een ambiguïteit vast. Enerzijds lijkt het alsof men differentiatie in doelen bij evaluatie wil vermijden omdat het praktisch gezien eenvoudiger is – ook gegeven het jaarklassensysteem – om alle leerlingen in eenzelfde traject op dezelfde manier te evalueren. Anderzijds gaan er stemmen op om net sneller over te schakelen naar evaluatie waarbij er gedifferentieerd wordt in getoetste doelen opdat de evaluatie meer zou aansluiten bij het niveau van het kind. Men ervaart hierbij een spanning tussen enerzijds de leerling aan een getuigschrift helpen en anderzijds werken op maat van de leerling zodat de leerling nog succeservaringen kan opdoen die hem/haar motiveren in het verdere leerproces. Uit de gesprekken met de leerkrachten en de directies is het duidelijk dat die spanning nog het meeste tot uiting komt in de contacten met ouders wanneer leerlingen omwille van curriculumdifferentiatie, niet meer in aanmerking zouden komen voor het getuigschrift. In de perceptie van sommige leerkrachten vormen de ouders zelfs bijna een hindernis om over te schakelen naar gedifferentieerde evaluatie. Case C in het basisonderwijs zet hier via een sterk beleid inzake ouderbetrokkenheid op in door tijdig én frequent met ouders hierover te communiceren.

In het secundair onderwijs leeft de overtuiging bij leerkrachten die deelnamen aan het casestudieonderzoek dat differentiëren in doelen in de praktijk moeilijk haalbaar is en veel werklast met zich meebrengt. Men vindt dit niet 'fair' en vindt en dat alle leerlingen (van eenzelfde klas) dezelfde doelen moeten halen, of men is ervan overtuigd dat de ouders en leerlingen hier problemen mee zouden hebben. Deze vaststelling zou verklaard kunnen worden door het feit dat men in het secundair onderwijs sterker vertrekt vanuit vooropgestelde doelen net omdat deze doelen ook vaak per graad geformuleerd zijn en niet voor een periode van zes jaar zoals dat in het basisonderwijs het geval is. Tegelijkertijd is de bevinding omtrent leerkrachten die differentiatie in doelen als *onfair* beschouwen paradoxaal: respondenten in het secundair geven aan dat ze op een toets aan bepaalde leerlingen bijvoorbeeld tips geven zodat deze leerlingen het beter zouden doen. Of achteraf quoteert men 'ruimer' voor bepaalde leerlingen om hen gemotiveerd te houden. Hoewel er niet expliciet naar gepeild werd, gaven leerlingen zelf spontaan aan dat ze het aspect van 'faire' evaluatie net heel belangrijk vinden. Opvallend is dat ze accommodaties en ondersteuning voor bepaalde leerlingen in de klas pas fair vinden als iedereen die er nood aan heeft er ook gebruik van kan maken. Een verschillende manier in quoteren ervaart men dan wel weer als oneerlijk.

### **1.4.3. Geïntegreerde evaluatiepraktijk en leerlingencompositie**

Zowel de casestudies als het surveyonderzoek laten zien dat leerkrachten in scholen met meer lage SES-leerlingen zich competentier voelen in het opmerken van diversiteit en meer zullen inzetten op zowel differentiatie als een geïntegreerde evaluatiepraktijk. Bovendien stellen we vast dat de scholen of de contexten (zoals bijvoorbeeld het bso) die door meer diversiteit gekenmerkt worden op de grenzen van het onderwijssysteem stuiten. Anders gezegd: het tegemoet komen aan diversiteit onder leerlingen op momenten van evaluatie lijkt te worden belemmerd door de wijze waarop onderwijs georganiseerd is in onderwijsvormen en jaarklassen.

### *Beperkingen onderwijsorganisatie: onderwijsvormen*

De casestudies wijzen op een tendens die samenhangt met de opdeling in onderwijsvormen. Leerkrachten in een categoriale aso-school geven aan dat ze beperkte ervaring hebben met diversiteit omdat ze vooral te maken krijgen met leerlingen uit de blanke, Nederlandstalige middenklasse. Tegelijkertijd hebben ze moeite met de idee van 'gedifferentieerd' evalueren. Differentiëren tussen studierichtingen is voor hen aannemelijker dan differentiëren binnen éénzelfde klas.

Bekeken vanuit de structuur van onderwijsvormen voelen aso-leerkrachten zich minder 'genoodzaakt' om te differentiëren of extra te ondersteunen bij evaluatie precies omdat de onderwijsstructuur nog andere 'opties' voorziet via de 'lagere' onderwijsvormen tso, kso en bso. Leerlingen die moeilijkheden ervaren worden binnen de huidige structuur van het secundair onderwijs sneller gepercipieerd als 'ze zitten niet op hun plaats' waardoor ze geheroriënteerd worden naar een 'lagere' onderwijsvorm. Leerkrachten voelen zich minder uitgedaagd om te differentiëren zolang die leerlingen nog kunnen doorstromen naar andere onderwijsvormen. Leerkrachten in het bso daarentegen, worden veel meer uitgedaagd door diversiteit. Dit verklaart mogelijks waarom deze leerkrachten meer en makkelijker overgaan naar een praktijk die rekening houdt met diversiteit. De multilevelanalyses bevestigen eveneens dat het geïntegreerde model voor evaluatiepraktijk eerder gehanteerd wordt door leerkrachten die les geven in het bso. Dit leunt bovendien aan bij de bevinding van een directeur uit het casestudieonderzoek die ervaart dat aso-leerkrachten zich minder flexibel opstellen ten aanzien van diversiteit dan hun collega's in tso en bso.

De literatuur bevestigt dat leerkrachten in het bso veel sterker geconfronteerd worden met diversiteit in de klassen. In dit licht is de samenhang in de multilevelanalyses tussen de onderwijsvorm bso en een geïntegreerde evaluatiepraktijk heel logisch. Volgens Van Praag et al. (2015) blijkt dat leerlingen met een lagere SES vaker kiezen voor meer praktijkgerichte studierichtingen. De groep leerlingen in het bso wordt gekenmerkt door een groot aandeel van leerlingen met een lage SES. Eerder onderzoek van Van Praag et al. (2013) verklaart dit door het waternetstelsel in het onderwijs. Een ongelijke spreiding van leerlingen met een verschillende etnische en socio-economische achtergrond over de verschillende onderwijsvormen leidt tot een zeer specifieke etnische en socio-economische compositie per onderwijsvorm met als gevolg dat het bso veel leerlingen heeft met een lagere SES. Bovendien beperkt de diversiteit in het bso zich niet tot de etnische en socio-economische compositie: er is ook een grotere diversiteit in gevolgde schoolloopbanen en dus ook in de voorkennis van leerlingen in vergelijking met meer academisch georiënteerde richtingen (Van Praag et al., 2015). Verschillende studies toonden aan dat leerlingen uit het bso zich minder gemotiveerd voelen omwille van een sterker gevoel van zinloosheid (Van Houtte, 2016; Van Houtte & Stevens, 2010, 2015). Het samenkomen van al deze uitdagingen in het bso verklaart mogelijks mee waarom de leerkrachten uit case F (met aanbod tso-bso) zich machteloos voelen in het tegemoet komen aan de leernoden van hun leerlingen.

Dat aso-leerkrachten minder geneigd zijn om rekening te houden met diversiteit, kunnen we dus mogelijks verklaren door de structuur van het onderwijssysteem. Ze voelen zich minder 'genoodzaakt' om over te gaan naar differentiatie omdat het onderwijssysteem nog andere 'opties' voorziet voor deze leerlingen via de 'lagere' onderwijsvormen tso, kso en bso. Leerlingen die het moeilijker hebben in het aso worden binnen de huidige structuur van het secundair onderwijs sneller geheroriënteerd.

### ***Beperkingen onderwijsorganisatie: het jaarklassensysteem***

Volgens het casestudieonderzoek in het basisonderwijs lopen scholen die meer gekenmerkt worden door diversiteit aan tegen de grenzen van het jaarklassensysteem. Ze geven aan dat ze wensen af te stappen van het jaarklassensysteem en al kleine voorzichtige initiatieven om hierin verandering te brengen. Eén school werkt met graadsklassen waarin flexibel omgegaan wordt met leerplandoelen. Twee scholen groepen de kleuters niet langer volgens leeftijd. Een andere school wil in de toekomst naar een sporenmodel evolueren waarbij leerlingen voor verschillende leergebieden kunnen kiezen tussen vier verschillende 'sporen'.

## **2. De opvattingen van leerkrachten**

### **2.1. Opvattingen met betrekking tot evaluatie**

Het is opvallend dat leerkrachten uit het surveyonderzoek die overtuigd zijn van de informerende waarde van evaluatie hun evaluatiepraktijk vaker inrichten volgens het gescheiden model, wat haaks lijkt te staan op de idee dat de informatie verkregen uit evaluatie aangewend kan worden om het onderwijs- en leerproces te bevorderen. Met andere woorden: de overtuiging dat evaluatie een informerende waarde heeft vertaalt zich niet in praktijken die deze informatie ook daadwerkelijk vertalen naar concrete, overeenkomstige acties. Dit doet vermoeden dat de ondersteunende of formatieve functie die evaluatie heeft voor het onderwijs- en leerproces slechts beperkt tot uiting komt in de evaluatiepraktijk. Het lijkt erop dat de stap om de informatie ook daadwerkelijk aan te wenden, ontbreekt. De informerende functie zou in dit geval beperkt blijven tot het informeren, zonder meer. Dit wordt bevestigd door de afwezigheid van een significant verband tussen evaluatiepraktijk van leerkrachten enerzijds en hun opvatting omtrent de impact die evaluatie heeft op hun onderwijspraktijk anderzijds. Dit sluit aan bij de vaststelling van Bonner (2016) namelijk dat de opvattingen van leerkrachten en hun praktijken niet steeds éénduidig zijn. De verklaring voor de mismatch in de opvattingen van leerkrachten die overtuigd zijn van de impact die evaluatie heeft op het onderwijsproces enerzijds en hun evaluatiepraktijken die anderzijds eerder sterk gescheiden verlopen van onderwijsactiviteiten moet gezocht worden in de ruimere context waarbinnen evaluatiepraktijken vorm krijgen (Bonner, 2016; Ercikan & Solano-Flores, 2016). Hiervoor kijken we naar het ruimere schoolbeleid dat met betrekking tot evaluatie over het algemeen eerder afwezig is. Dit blijkt uit zowel het casestudie- als het surveyonderzoek. Het feit dat er volgens de surveydata op schoolniveau ook geen sprake is van een cultuur omtrent leerlinggerichtheid en/of leerstofgerichtheid duidt op de algemene afwezigheid van een gedeelde visie op onderwijs en leren. Dit vormt echter een noodzakelijke voorwaarde om te komen tot een weloverwogen afstemming tussen onderwijsactiviteiten en evaluatieactiviteiten die zo bijdragen aan het leerproces van leerlingen (James & Pedder, 2006).

Leerkrachten die eerder een kritische houding hebben ten aanzien van evaluatie zouden volgens het surveyonderzoek vaker het geïntegreerde model toepassen. De kritische houding van respondenten in het surveyonderzoek ten aanzien van evaluatie verwijst naar de opvatting dat evaluatie meetfouten kan bevatten en oneerlijk kan zijn voor sommige leerlingen. Dit verband lijkt logisch te zijn als we ervan uitgaan dat 'evaluatie' gepercipieerd wordt als *one-size-fits-all*. Deze leerkrachten zullen vaker gebruik maken van evaluatie die sterk ingebed is in de lespraktijk en zullen leerlingen er meer bij betrekken precies om tegemoet te komen aan mogelijke meetfouten of oneerlijke gevolgen. Deze leerkrachten zullen dus vaker rekening houden met de noden van leerlingen.

Deze spanning tussen enerzijds een evaluatiepraktijk die gericht is op *one-size-fits-all* en anderzijds de overtuiging dat dit voor bepaalde leerlingen een hindernis kan vormen, duikt in het casestudieonderzoek eveneens op als het gaat over gedifferentieerd evalueren. Differentiatie in doelen die geëvalueerd worden, gebeurt in het basisonderwijs iets vaker dan in het secundair onderwijs. In het secundair onderwijs blijken leerkrachten hiermee moeite te hebben omwille van een onderliggende opvatting, namelijk dat men dit niet 'fair' vindt en dat alle leerlingen (van eenzelfde klas) aan dezelfde doelen horen tegemoet te komen. Of men verwijst hiervoor naar dergelijke opvattingen van ouders en leerlingen. Tegelijkertijd is de bevinding omtrent leerkrachten die differentiatie in doelen als *onfair* beschouwen paradoxaal aangezien sommige leerkrachten in het secundair aangeven dat ze op een toets aan bepaalde leerlingen bijvoorbeeld tips zullen geven zodat deze leerlingen het beter zouden doen. Of bij de quotering achteraf wordt er 'ruimer' gequoteerd voor bepaalde leerlingen. De leerkrachten lijken dit te doen om de leerlingen die het moeilijker hebben toch nog gemotiveerd te houden. Opvallend hierbij is dat enkele leerlingen er op wezen dat ze accommodaties en ondersteuning voor bepaalde leerlingen in de klas geen probleem vinden, maar ze vinden het pas fair als iedereen die er nood aan heeft er ook gebruik van kan maken.

## 2.2. Houding ten aanzien van centrale, gestandaardiseerde examens

De opvattingen omtrent centrale gestandaardiseerde examens werden uitsluitend bevestigd in het casestudieonderzoek. Het is opmerkelijk dat respondenten vaak geen helder onderscheid maken tussen enerzijds centrale examens die opgelegd worden voor alle leerlingen van een bepaalde leeftijd en gestandaardiseerde examens anderzijds. Deze laatste worden niet noodzakelijk verplicht afgenomen en kunnen bijdragen in de gegevensverzameling van een school om zo geïnformeerde beleidskeuzes te maken die de onderwijskwaliteit bevorderen. Directies en leerkrachten in het basisonderwijs zijn overwegend negatief als het gaat over centrale en/of gestandaardiseerde examens, in het secundair onderwijs zijn de houdingen, vooral onder de leerkrachten meer gevarieerd.

In de onderzochte cases van het basisonderwijs hebben zowel de leerkrachten en de directies een kritische houding ten aanzien van centraal georganiseerde examens. De bedenkingen tonen veel gelijkenissen met wat reeds in de literatuur verscheen omtrent negatieve gevolgen die centrale toetsing met zich mee kunnen brengen. Men verwijst naar het uniformiseren van het curriculum en de overtuiging dat het ons onderwijs nóg prestatiegerichter zou maken. De respondenten in basisonderwijs zijn het er allen over eens dat het resultaat van dergelijke examens niet zou mogen leiden tot een bindend advies inzake de oriëntering van leerlingen in hun studiekeuze omdat dit gebaseerd is op een momentopname. Bovendien verwijst men ook naar het feit dat leerlingen op de leeftijd van 12 jaar nog veel veranderen doorheen de jaren die volgen na het basisonderwijs. Hoewel alle respondenten negatieve elementen aanhalen omtrent centrale gestandaardiseerde examens, wordt er ook verwezen naar positieve elementen. Volgens sommige leerkrachten is het nuttig om te kunnen vergelijken met anderen, om een duidelijk beeld te brengen naar de ouders toe of om leerlingen die nog twijfelen over hun studiekeuze een houvast te bieden. Opvallend is dat deze positieve elementen enkel aangehaald worden door die leerkrachten die tewerkgesteld zijn in een school die gekenmerkt wordt door een leerlingenpubliek met hoge socio-economische status. Bij leerkrachten wiens werkcontext gekenmerkt wordt door meer diversiteit en die dus meer leerlingen hebben met een lage socio-economische status, stellen we een éézijdig negatieve houding vast. Deze leerkrachten verwijzen hiervoor naar de diversiteit in hun leerlingenpubliek om te illustreren hoe een uniform examen volgens hen haaks staat op een klaspraktijk die inzet op differentiatie. In één

specifieke school verwijst men naar hoe dit ingaat tegen de visie die men heeft op onderwijs en leren omdat die vertrekt van het individuele leerproces van de leerling.

In de onderzochte cases van het secundair onderwijs zijn de houdingen van de directies gelijklopend aan deze in het basisonderwijs. Onder de leerkrachten zijn de visies meer verdeeld: waar er in het basisonderwijs een groep leerkrachten is met een ambivalente houding omdat ze verwijzen naar zowel positieve als negatieve elementen, is er in het secundair onderwijs een minderheid van leerkrachten die duidelijk voorstander is van centrale, gestandaardiseerde toetsen. Er kunnen hierbij geen verbanden vastgesteld worden tussen hun houding ten aanzien van dergelijke toetsen en kenmerken van deze leerkrachten zoals bijvoorbeeld het type vak dat ze geven, de onderwijsvorm waarin ze les geven of het aantal jaar ervaring dat ze hebben. Opvallend is de vaststelling dat de meerderheid van die leerkrachten die een invoering van centrale examens wenselijk zou vinden uit één specifieke case komen: een tso-bso-school die gekenmerkt wordt door veel diversiteit. De argumenten pro centrale examens van deze leerkrachten vinden hun oorsprong voornamelijk in het attesteringsbeleid van deze school. De leerkrachten voelen namelijk druk vanuit het schoolbeleid om zoveel mogelijk A-attesten uit te reiken, wat volgens hen leidt tot 'onterechte A-attesten' en uiteindelijk tot minderwaardige diploma's. Deze leerkrachten zien in centrale examens een tegenwicht voor het evaluatiebeleid van hun school omdat dit in hun ogen een oplossing zou kunnen bieden voor de uitwassen van het huidige evaluatiebeleid van de school, namelijk een te enge focus op het uitreiken van A-attesten zonder dat hieraan ook een beleid van zorg en remediëring gekoppeld wordt.

### **2.3. Opvattingen omtrent evalueren van kennis versus evalueren van competenties**

De kwalitatieve data wijzen op veel onzekerheid omtrent competentiegericht evalueren bij leerkrachten in het basisonderwijs. De evaluatiepraktijk lijkt in het basisonderwijs nog sterk aan te leunen bij kennisgerichte evaluatie, hoewel er op basis van de rapportering van leerkrachten met een langere loopbaan in het onderwijs, aanwijzingen zijn voor een evolutie richting meer competentiegerichte evaluatie. De vergelijking tussen de cases onderling laat zien dat de leerkrachten uit de methodeschool zich minder onzeker voelen met betrekking tot competentiegerichte evaluatie. Dit blijkt ook uit hun dagdagelijkse evaluatiepraktijk waarin men frequent inzet op evaluatievormen die zich lenen tot het evalueren van competenties zoals observaties en zelf- en peerevaluatie.

In het secundair onderwijs hangt de mate waarin er aan competentiegerichte evaluatie gedaan wordt voornamelijk samen met het soort vak. Volgens de rapportering door leerkrachten, maakt competentiegerichte evaluatie eerder deel uit van de evaluatiepraktijk van leerkrachten die een praktijkvak, kunstvak of PAV geven. Dit sluit aan bij de leerplannen van deze vakken die meer inzetten op competenties dan op louter kennis.

Met betrekking tot de taalvakken in het secundair onderwijs, waarin de leerplannen eveneens de competenties benadrukken, stellen we een wisselend beeld vast. De leerkrachten uit het casestudieonderzoek verwijzen naar een percentage om het evenwicht tussen de evaluatie van enerzijds kennis en anderzijds vaardigheden weer te geven. Dit varieert sterk van leerkracht tot leerkracht: er zijn er bij wie het voornamelijk om competenties draait en er zijn er die een onderscheid maken tussen 50% kennis en 50% vaardigheden. Enkele leerkrachten worstelen met het onderscheid tussen kennis en vaardigheden omdat dit in realiteit sterk verweven is met elkaar wat het heel lastig maakt om transparant te kunnen rapporteren naar leerlingen en ouders.

Er kan besloten worden dat de kennisgerichte evaluatie doorheen de jaren evolueerde naar meer competentiegerichte evaluatie, al blijven leerkrachten zich hier nog onzeker bij voelen. In het secundair onderwijs liggen de opvattingen én de ervaringen van leerkrachten verder uit elkaar: afhankelijk van het soort vak en de bijhorende leerplannen zal er sprake zijn van een meer competentiegerichte evaluatiepraktijk. Praktijkgerichte vakken in kso en bso lenen zich meer tot competentiegerichte evaluatie dan kennisgerichte evaluatie. Met betrekking tot taalvakken zijn de opvattingen (en de praktijken) van leerkrachten in het secundair meer verdeeld.

## **2.4. Opvattingen over eigen competenties en bekwaamheid**

### **2.4.1. Doelmatigheidsbeleving**

De gesprekken met leerkrachten in de casestudies geven aan dat leerkrachten verschillen in de mate waarin ze zich competent voelen om leerlingen te evalueren. De verschillen zijn voornamelijk terug te voeren naar verschil in ervaring. Startende leerkrachten voelen zich nog veel meer onzeker. Ze zijn zoekende naar hoe je doelgericht moet evalueren en hoe je evaluatie en instructie vlot op elkaar kan laten inspelen. Ze beschikken vanuit de opleiding over de theoretische kennis, maar ze missen volgens hen oefening en ervaring in de praktijk om dit vlot te kunnen doen. De leerkrachten met meer ervaring bevestigen dat ze hierin gegroeid zijn doorheen de jaren. Een directie ondervindt dat beginnende leerkrachten vaker de evaluatiepraktijk zullen inrichten op basis van wat de gebruikte methodes hierrond voorschrijven en dus minder kritisch kijken naar bestaand evaluatiemateriaal.

Het surveyonderzoek stelt echter dat het aantal jaar ervaring niet zozeer de bepalende factor is voor de mate waarin men zich competent voelt om evaluatie in te bedden in de onderwijspraktijk. Het gaat veeleer om de doelmatigheidsbeleving (*self-efficacy*). De mate waarin men gelooft in de eigen competenties inzake instructiestrategieën, klasmanagement en leerlingenbetrokkenheid blijkt positief samen te hangen met de mate waarin leerkrachten hun evaluatieactiviteiten integreren in onderwijsactiviteiten. Leerkrachten met een sterkere doelmatigheidsbeleving, en dus niet per se over meer ervaring beschikken, zullen vaker evalueren volgens het geïntegreerde model.

### **2.4.2. Bekwaamheid om diversiteit onder leerlingen op te merken**

De mate waarin men zichzelf bekwaam acht om de diversiteit onder leerlingen op te merken is volgens de multilevelanalyses positief gerelateerd aan een evaluatiepraktijk die (semi-)geïntegreerd is. Een logisch verband want het in staat zijn om verschillen tussen leerlingen op te merken vormt een voorwaarde om er ook op te kunnen inspelen tijdens momenten van instructie, zodat er gericht gedifferentieerd kan worden.

De gesprekken met leerkrachten uit de casestudies wijzen op een grotere onzekerheid bij leerkrachten, als het gaat over evaluatie waarbij men rekening houdt met diversiteit. Leerkrachten geven aan dat het nog steeds zoeken is naar manieren om dit nog anders/ beter te doen. Op dit vlak lijken alle leerkrachten aan te geven dat er nog groeimarge is. Deze bevinding sluit aan bij het feit dat leerkrachten over het algemeen hoofdzakelijk inzetten op een eerder gescheiden evaluatiepraktijk dan een (semi-)geïntegreerde praktijk die meer beantwoordt aan de diversiteit onder leerlingen. De bekwaamheid om diversiteit onder leerlingen op te merken lijkt hier wel een bepalende rol in te spelen.

### **2.4.3. Professionalisering leerkrachten**

Dat de bekwaamheid van leerkrachten een rol speelt, leiden we ook af uit de multilevelanalyses die een positieve samenhang laten zien tussen de geïntegreerde evaluatiepraktijk en de mate waarin leerkrachten inzetten op professionalisering omtrent evaluatie. Mogelijks kan dit effect van professionalisering verklaard worden door het feit dat leerkrachten die aan de slag willen gaan met geïntegreerde evaluatiepraktijk net vaker op zoek gaan naar professionalisering om te leren hoe ze dit het beste kunnen doen. Anderzijds kan het ook omgekeerd waarbij professionalisering inzake evaluatie net leidt tot meer geïntegreerde evaluatiepraktijken. Het surveyonderzoek toont aan dat de mate waarin leerkrachten zich professionaliseren omtrent evaluatie een schoolkenmerk is. Dit sluit aan bij de casestudies waar de mate van professionalisering sterk beïnvloed wordt door aanmoediging vanuit de school zelf. Uit de casestudies blijkt bovendien dat leerkrachten nauwelijks op eigen initiatief voor nascholing omtrent evaluatie zullen kiezen, zelfs al vinden ze dat de lerarenopleiding hen beperkt voorbereidde. Professionalisering op schoolniveau kwam in de cases tot uiting via een schoolbeleid dat voortdurend aandacht schenkt aan evaluatie én het blijvend professionaliseren van leerkrachten stimuleert, onder andere via een langdurig begeleidingstraject met een externe partner.



### 3. Evaluatiebeleid van scholen

#### 3.1. Evaluatiepraktijk is eerder leerkracht- dan schoolafhankelijk

De intraclasscorrelaties van het gescheiden model en de (semi-)geïntegreerde modellen (telkens lager dan 0,60) laten zien dat de evaluatiepraktijken van leerkrachten onderling veel verschillen en dus niet beschouwd kunnen worden als een schoolkenmerk. Dit ligt in de lijn van de bevindingen die gedaan werden bij de twaalf onderzochte cases in basis- en secundair onderwijs. Binnen eenzelfde school werd er doorgaans een variatie aan evaluatiepraktijken vastgesteld, wat aantoont dat evaluatie eerder bepaald wordt door de individuele leerkracht dan door de school. Slechts één case (F) wijkt hiervan af, aangezien alle bevraagde leerkrachten in kleuter- en lager onderwijs eenzelfde model voor evaluatiepraktijk hanteren. Bovendien gaat dit om het geïntegreerde model wat, volgens het surveyonderzoek, het minst frequent gebruikt lijkt te worden in het Vlaamse basis- en secundair onderwijs.

Dat evaluatie een aangelegenheid is van de individuele leerkracht, blijkt tevens uit de opvattingen. De intraclasscorrelaties voor elk van de subschalen (ICC informeren = 0,37; ICC kritisch = 0,41 en ICC geen impact op instructie = 0,37) laten zien dat er voor de onderzochte scholen geen sprake is van een schoolcultuur omtrent evaluatie. Niet alleen op vlak van praktijk, maar ook op vlak van opvattingen is een gedeelde schoolcultuur omtrent evaluatie afwezig.

#### 3.2. Visie bij evaluatiebeleid ontbreekt vaak

Het feit dat de praktijken en de opvattingen inzake leerlingenevaluatie eerder leerkrachtafhankelijk dan schoolafhankelijk zijn, kan worden verklaard door de wijze waarop dit beleid ingevuld wordt. De onderzochte cases in het basisonderwijs wijzen erop dat de meerderheid van de scholen wel over een evaluatiebeleid beschikt, maar dat de invulling hiervan slechts heel beperkt is. In vijf van de zes cases ontbreekt het aan een scherp gestelde en gedragen visie omtrent waarom er geëvalueerd wordt. Het beleid zelf beperkt zich in deze gevallen vooral tot praktische afspraken omtrent de rapportering, het gebruikte evaluatiemateriaal (vaak vertrekkend vanuit methodetoetsen) en de frequentie waarmee er geëvalueerd wordt. Met uitzondering van één case (F), waarbij het evaluatiebeleid vertrekt vanuit de 'waarom-vraag' en tevens inhaakt op de ruimere visie op onderwijs en leren die de school erop na houdt. Deze visie vertaalt zich bovendien in de evaluatiepraktijken van leerkrachten.

Dezelfde tendens duikt op in de cases van het secundair onderwijs, waar evaluatiebeleid zich vaak beperkt tot het beantwoorden van de hoe-, wanneer- en wat-vragen. Bovendien stellen we vast dat hier een specifieke rol toebedeeld wordt aan de vakgroep. Afhankelijk van de kwaliteit van de vakgroepwerking worden per vak afspraken gemaakt over wat geëvalueerd wordt (het evenwicht tussen kennis en vaardigheden, product versus proces), wanneer en hoe vaak er geëvalueerd wordt en hoe dit gerapporteerd wordt aan leerlingen en ouders. Binnen de vakgroep wordt er bovendien ook gewaakt over de horizontale afstemming, opdat de evaluatiepraktijk van parallelleerkrachten eenvormig zou zijn, en de verticale afstemming, opdat de leerlijn van het vak ook gereflecteerd wordt in de evaluaties over de leerjaren heen. Slechts in beperkte mate wordt er in de onderzochte cases van het secundair onderwijs aandacht besteedt aan waarom er geëvalueerd wordt. In twee cases zijn er signalen die aangeven dat men begaan is met deze vraag, al verschilt de wijze waarop dit zich uit heel sterk: gaande van een visie op papier die niet gekend is bij het team tot een leerkrachtenteam dat actief op zoek gaat naar manieren om *een papieren visie* ook te vertalen in de eigen vak- en klaspraktijk.



### 3.3. Factoren die een evaluatiebeleid op school faciliteren

#### 3.3.1. Inbedding van evaluatiebeleid in het ruimere schoolbeleid

Het surveyonderzoek geeft aan dat leerkrachten van scholen die over een sterker beleidsvoerend vermogen beschikken met hun evaluatiepraktijk vaker aanleunen bij het geïntegreerde model voor evaluatiepraktijk. De casestudies bieden meer inzicht in hoe dit beleidsvoerend vermogen er concreet kan uitzien. Zo worden er enkele good-practices gevonden waarbij het evaluatiebeleid van scholen verbonden wordt aan enkele andere aspecten van het ruimere schoolbeleid: onderwijsbeleid, zorgbeleid en het kwaliteitsbeleid.

##### *Onderwijsbeleid*

Het evaluatiebeleid van een school wordt door heel wat bevraagde leerkrachten en directies in de casestudies gepercipieerd als het geheel van afspraken en impliciete, ongeschreven regels. Een evaluatiebeleid dat vertrekt vanuit een visie op onderwijs ontbreekt hierbij. Een visie op evaluatie die voortvloeit uit de vraag ‘waarom doen we aan evaluatie’ of ‘waartoe willen we dat evaluatie dient’, blijkt nochtans cruciaal te zijn voor het evaluatiebeleid en de daarop afgestemde evaluatiepraktijken. Onderwijs blijkt namelijk pas effectief en succesvol te zijn als de evaluatie ook afgestemd is op de onderwijspraktijk (James, 2006).

Het surveyonderzoek laat specifiek met betrekking tot die visie op onderwijs een hiaat zien. De onderwijsopvattingen van leerkrachten worden bevraagd aan de hand van twee dimensies of subschalen: leerlinggerichtheid en leerstofgerichtheid. Deze laatste dimensie verwijst naar het belang dat men besteedt aan beroepsvoorbereiding, discipline en conformiteit van de leerlingen (Denessen, 2000). Deze dimensie sluit aan bij een productgerichte onderwijsbenadering met nadruk op het leveren van prestaties en het behalen van goede cijfers. De leerlinggerichte dimensie stelt de vorming van de lerende meer centraal, waarbij de lerende ook mee inspraak krijgt en waarbij de nadruk ligt op het proces van leren en het leren van elkaar. Over het algemeen tonen de gemiddelde scores dat de leerlinggerichte opvattingen sterker aanwezig zijn dan de leerstofgerichte opvattingen, dit voor zowel het basisonderwijs als het secundair onderwijs. Opvallend is echter dat de intraclasscorrelaties te laag zijn om te kunnen spreken over gedeelde overtuigingen of een cultuur op schoolniveau met betrekking tot deze onderwijsopvattingen. Dit wijst mogelijk op een onvoldoende uitgewerkte en gedragen visie op onderwijs op schoolniveau, waardoor deze ook geen richting kan geven aan een visie op evaluatie van waaruit een evaluatiebeleid uitgewerkt kan worden.

Het casestudieonderzoek laat op dit vlak een uitzondering zien waar in case F van het basisonderwijs een nauwe verbinding is tussen het onderwijsbeleid en het evaluatiebeleid. Vanuit de visie op onderwijs die sterk leerlinggericht is, wil men de ontwikkeling van leerlingen holistisch benaderen: *“... wat betekent dat we het ‘vakjesdenken’ overstijgen. We zien elk kind als een persoon in ontwikkeling, met zijn specifieke talenten en tekorten die beide een groeimarge hebben. We hebben respect voor het begrip ‘meervoudige intelligentie’: geen enkele van de intelligenties mag belangrijker worden geacht dan een andere.”* In hetzelfde beleidsdocument van deze school lezen we dat dit doorgetrokken wordt in de wijze van evalueren: *“We evalueren continu om elk kind ten dienste te zijn, succes te laten kennen in zijn werk, verder te helpen in de ontplooiing. (...) Er wordt steeds geëvalueerd in een sfeer van respect voor het kind.(...) We respecteren het natuurlijk leerritme van elk kind.”* Deze visie wordt weerspiegeld in een evaluatiepraktijk van leerkrachten die stevast aansluit bij het geïntegreerde model. Leerlingen worden er opvallend sterk betrokken bij het evaluatieproces: ze kunnen mee bepalen wanneer ze een

evaluatie gaan afleggen, evalueren hun klasgenoten en geven feedback aan elkaar, evalueren zichzelf frequent om na te gaan waarop ze nog moeten inzetten.

Een gelijkaardig patroon van afstemming tussen visie op leren en visie op evalueren dook op in case C van het secundair onderwijs. Hoewel deze school in vergelijking met case F van het basisonderwijs, veel recenter werk maakte van een evaluatiebeleid op schoolniveau, is deze afstemming al zichtbaar in de wijze waarop klassenraden ingevuld en georganiseerd worden. Zoals de directeur van deze secundaire school zelf aangeeft: *“evalueren moet het leren dienen, niet het leren belemmeren”*. Zo is er in hun werking een duidelijk onderscheid tussen de begeleidende, delibererende en oriënterende klassenraden. Dit blijkt niet alleen uit het formele beleid van de school, maar eveneens uit wat leerkrachten vertellen over hun evaluatiepraktijk. De visie op evaluatie, namelijk dat evaluatie het leren moet dienen, vertaalt zich onder andere in het aantal begeleidende klassenraden die georganiseerd worden. In de eerste graad worden er vijf klassenraden georganiseerd per schooljaar waarop alle leerlingen aan bod komen. De eerste hiervan is in het begin van het schooljaar: een portretterende klassenraad waarbij voor elke leerling de sterktes en talenten in kaart gebracht worden. Deze positieve elementen worden zo ook meegenomen bij het (her)oriënteren van leerlingen. Voor de tweede en derde graad zijn er ongeveer zes à zeven klassenraden per schooljaar. De begeleidende klassenraden hebben als doel om na te gaan hoe leerlingen verder ondersteund kunnen worden in hun leerproces. Ook tijdens delibererende klassenraden wil men *“het leren blijven dienen”*. Opvallend is dat de school bij deliberaties niet terugvalt op neergeschreven afspraken omtrent percentages en/of het aantal tekorten dat een leerling heeft. Er gelden geen vastgelegde criteria, noch formeel, noch informeel. Het beleidsdocument *‘Deliberatiehouding tijdens klassenraden’* geeft richting aan hoe beslist wordt over de toekomstige schoolloopbaan van leerlingen. De kernwoorden hierbij zijn ‘positief’, ‘toekomstgericht’ en ‘collegiaal & consensusgericht’. Deze deliberatiehouding, gecombineerd met het in kaart brengen van de talenten van leerlingen, verhoogt de kans op een positieve oriëntering van leerlingen en stelt nog steeds het leerproces van de leerling centraal.

Beide voorbeelden uit het basis- en secundair onderwijs illustreren hoe een visie op onderwijs en leren vertaald wordt naar een visie op evalueren, die tevens weerspiegeld wordt in de evaluatiepraktijk van leerkrachten op deze scholen.

### **Zorgbeleid**

In de formele beleidsdocumenten over het evaluatiebeleid van scholen die deelnamen aan het casestudieonderzoek zijn er over het algemeen weinig tot geen verwijzingen naar hoe er rekening gehouden wordt met diversiteit tijdens momenten van evaluatie. Het formele beleid van case F uit het basisonderwijs daarentegen verwijst hier wel naar via kleine nuances in de gebruikte formuleringen. Zo is het bijvoorbeeld opvallend dat hier vaak gesproken wordt over *‘elk kind’* in plaats van over *‘het kind’* of *‘de leerling’*. In het secundair onderwijs is er ook één case die in de beleidsdocumenten van de school expliciet verwijst naar *“het afstemmen van ons evaluatiesysteem op de noden van de lerende.”*

Naar analogie met het algemene evaluatiebeleid, beperkt het beleid van een school omtrent evaluatie en diversiteit zich niet tot het formele beleid, maar omvat dit eveneens de praktijken in de school en de afspraken of procedures hierrond op schoolniveau. De interviews met directies en leerkrachten laten zien dat er op momenten van evaluatie wel degelijk maatregelen zijn die tegemoet komen aan de diversiteit, ook in die scholen die gekenmerkt worden door een beperktere diversiteit.

De methodeschool van het basisonderwijs (case F) vormt hier opnieuw een uitzondering omdat deze school vanuit haar visie op onderwijs en leren de principes omtrent rekening houden met diversiteit zodanig geïnternaliseerd heeft dat ze automatisch doorgetrokken worden bij de evaluatie en verweven zitten in de dagdagelijkse praktijk. Omgaan met diversiteit, of het nu tijdens een gewoon lesmoment is, of tijdens een evaluatiemoment, het zit verweven in het DNA van de werking van deze school. Meer nog, in deze school is er vaak geen duidelijk onderscheid tussen een klassikaal lesmoment of een klassikaal evaluatiemoment omdat dit in de praktijk vaak geïntegreerd verloopt. In die zin wijkt deze school het sterkste af van de andere cases. In de meerderheid van de onderzochte scholen stellen we vast dat het rekening houden met diversiteit tijdens evaluatie inhoudt dat er afgeweken wordt van de dagdagelijkse gang van zaken bij evaluatie en er wordt dan bij bepaalde leerlingen anders geëvalueerd dan bij de meerderheid van de leerlingen. Dit blijkt uit het gebruik van termen als 'maatregelen' en 'aanpassingen' als men het heeft over omgaan met diversiteit tijdens evaluaties.

In de meerderheid van deze cases wordt stevast verwezen naar de sticordi- of redicodi-maatregelen. 'Sticordi' staat voor stimuleren, compenseren, remediëren en dispensereren en wordt sinds de invoering van het M-decreet 'redicodi' genoemd. Hierbij lijkt men voornamelijk compenserende maatregelen aan te wenden. Dit omvat, bijvoorbeeld, het gebruik van hulpmiddelen of accommodaties zoals koptelefoons, rekenmachines, woordenboeken, opzoekkaarten voor de maaltafels, leesteksten op voorhand meegeven aan leerlingen met dyslexie, het mondeling toelichten van vragen bij leerlingen met een andere thuistaal, extra tijd krijgen,.... Differentiërende maatregelen waarbij dezelfde leerdoelen geëvalueerd worden, maar dan op een andere manier, komen in mindere mate aan bod. Dispenserende maatregelen bleken eerder uitzonderlijk en worden eerder gebruikt bij leerlingen die een individueel aangepast curriculum hebben en dus vrijgesteld zijn van een aantal doelen. Dispenserende maatregelen worden niet gerapporteerd in de casestudies van het secundair onderwijs. Dat bepaalde leerlingen *anders* geëvalueerd worden, komt nog sterker tot uiting in het secundair onderwijs waar de meerderheid van de scholen een aparte klas (zorgklas of sticordiklas genoemd) inricht tijdens de examenperiode voor die leerlingen die meer tijd of extra ondersteuning nodig hebben.

Voor de bovengenoemde voorbeelden uit de casestudies wordt er vaak verwezen naar het zorgbeleid op school. Voor leerlingen in het lager onderwijs worden afspraken gemaakt in overleg met de zorgleerkracht of zorgcoördinator. In het secundair onderwijs wordt verwezen naar de rol van de leerlingenbegeleiders omdat zij vaker over meer ervaring beschikken als het gaat over het omgaan met diversiteit.

### **Kwaliteitsbeleid**

Evaluatie biedt informatie over de onderwijspraktijk van de leerkracht die op basis van de uitkomsten van evaluatie kan bepalen welke bijsturing nodig is. Deze functie van evaluatie kan aangegrepen worden op schoolniveau om te reflecteren over hoe de school het doet. De resultaten van toetsen houden de school een spiegel voor en kunnen aanleiding geven tot het aanpakken van bepaalde hiaten. De 'toolkit gevalideerde toetsen basisonderwijs' is bijvoorbeeld bedoeld voor het versterken van de kwaliteitszorg in het basisonderwijs. Het surveyonderzoek bevestigt dat alle scholen gebruik maken van gevalideerde toetsen. De deelnemende scholen van het casestudieonderzoek geven echter weinig gevolg aan de resultaten op deze toetsen. Case C van het basisonderwijs vormde hierop een uitzondering.

Gekenmerkt door een groot aandeel leerlingen met een andere thuistaal was school C op zoek naar manieren om in te spelen op de moeilijkheden inzake taalontwikkeling. De school maakte hiervoor gebruik van gestandaardiseerde taaltoetsen om het gevoel dat er al langer was bevestigd te zien: de leerlingen scoorden heel laag met betrekking tot taalvaardigheid Nederlands. Nog belangrijker is dat in deze school de resultaten van de toetsen gebruikt worden om een aantal initiatieven, over verschillende jaren gespreid, te nemen. Zo deed de school beroep op de begeleidingsdienst (professionaliseringsbeleid) om advies te krijgen inzake het bevorderen van taalvaardigheid. Onder deze begeleiding werd besloten om de leergebieden 'taal' en 'wereldoriëntatie' niet langer als aparte domeinen aan te bieden, maar geïntegreerd. Dit had tevens een impact op de onderwijsaanpak: bestaande methodes voor taal werden niet langer gebruikt, maar werden door de leerkrachten zelf uitgewerkt (onderwijsbeleid). Daarenboven werden leerkrachten begeleid in hun zoektocht naar hoe ze spreekvaardigheid het beste konden evalueren (evaluatiebeleid). Ondertussen werden de gestandaardiseerde taaltoetsen jaarlijks hernomen en lieten de resultaten een stijgende trend zien. Na de positieve effecten van de initiatieven in het lager onderwijs, werd er na een paar jaar besloten om ook in het kleuteronderwijs de aandacht te vestigen op de bevordering van taalvaardigheid.

Deze case (C) uit het basisonderwijs vormt een *good practice* als het gaat over het aanwenden van evaluatie als instrument voor zelfevaluatie. Het reflecterend vermogen van deze school stelt de school in staat om op basis van voortdurende monitoring de kwaliteitsontwikkeling te bevorderen door specifieke beleidskeuzes te koppelen aan wat men vaststelt op basis van de toetsresultaten. Deze toetsen worden afgenomen op aansturen van de directie en de resultaten spelen verder geen enkele rol in de beoordeling van individuele leerlingen.

Binnen de casestudies van het secundair onderwijs zou een dergelijke aanpak een mogelijke oplossing kunnen vormen voor case F, waar het schoolbeleid te eng gericht is op het uitreiken van A-attesten zonder de nodige remediëring en ondersteuning te voorzien die ervoor zouden kunnen zorgen dat een A-attest terecht behaald wordt.

### **3.3.2. Van een ingebed evaluatiebeleid naar een gedragen evaluatiebeleid**

De vertaalslag van een formeel evaluatiebeleid naar een gedeeld beleid dat tevens ingang vindt in de evaluatiepraktijken van leerkrachten gebeurt niet zomaar. Zowel de casestudies als het surveyonderzoek reiken een aantal aspecten aan die deze vertaalslag bevorderen: professionalisering, communicatie en samenwerking.

#### ***Professionalisering***

De resultaten van de casestudies geven aan dat het loont om kennis over evaluatie binnen te brengen in het team. Kennisverwerving ondersteunt de visieontwikkeling met betrekking tot evaluatie. Dit kan op verschillende manieren vorm krijgen: via het volgen van externe nascholing, het uitnodigen van een expert of het inschakelen van een trajectbegeleider. In de onderzochte cases zagen we dat de visieontwikkeling volgens een gefaseerd proces verloopt. Vaak verdiept een (vakoverstijgende) werkgroep zich in het thema evaluatie om vervolgens over te gaan naar het ontwikkelen van een schoolvisie die via overleg en afstemming met de collega's ook een gedeelde visie wordt. Een belangrijke voorwaarde hierbij is dat het team dat begeleid zal worden, ook daadwerkelijk de behoefte voelt om zich hier in te verdiepen. Dit kan extern aangestuurd worden, bijvoorbeeld na een doorlichting van de school, maar even goed kan dit vanuit het team zelf komen, zoals bijvoorbeeld in school C van het secundair onderwijs, waar men vond dat er te veel tijd verloren ging aan

leerlingenevaluatie, ten koste van de onderwijstijd. Hiertoe werden alle professionaliseringsinitiatieven specifiek gelinkt aan evaluatiebeleid en -praktijk.

Het surveyonderzoek geeft bovendien aan dat de mate waarin leerkrachten professionalisering volgen inzake evaluatie een voorspeller is voor de mate waarin ze het (semi-)geïntegreerde model aanwenden in hun evaluatiepraktijk. Deze bevinding lijkt aan te sluiten bij de bevindingen in het casestudieonderzoek. Hoewel de initiatieven tot professionalisering over het algemeen nog vrij beperkt zijn, stellen we vast dat het evaluatiebeleid vorm begint te krijgen in die scholen die werk maken van professionalisering op dit vlak. Daarenboven zijn de *good practices* uit de casestudies steeds gericht op het bevorderen van het leren van alle leerlingen, wat aansluit bij de meer geïntegreerde vormen van evaluatiepraktijk die in het surveyonderzoek samenhangen met een hogere mate van professionalisering.

### **Communicatie**

Het spreekt voor zich dat de kennisverwerving gemakkelijker leidt tot kennisdeling indien er een cultuur is op school die gekenmerkt wordt door communicatie. Doeltreffende communicatie vormt dan ook één van de indicatoren voor een sterk beleidsvoerend vermogen van scholen (Vanhoof, Deneire, & Van Petegem, 2011). Uit het casestudieonderzoek blijkt dat een vakoverstijgende werkgroep die zich buigt over evaluatie draagt enerzijds bij tot het loskomen van de *eilandjes* die gevormd worden door de vakgroepen. Zo gaat er meer aandacht naar de visieontwikkeling op schoolniveau omdat er niet louter afspraken gemaakt worden omtrent vragen die focussen op hoe en wanneer er geëvalueerd wordt binnen één bepaald vak. Anderzijds is het ook nodig dat er vanuit deze vakoverstijgende werkgroep gecommuniceerd wordt met collega-leerkrachten. De visie op evaluatie wordt door de werkgroep afgetoetst bij de collega-leerkrachten om eventueel bijgestuurd en vertaald te worden naar concrete acties. Een werkgroep of kerngroep rond evaluatie heeft dus niet alleen een rol in kennisverwerving, maar evenzeer in kennisdeling rond evaluatie. Via terugkoppeling naar en overleg met de collega's en/of vakgroepen wordt het mogelijk om tot een gedeelde en gedragen visie te komen. Op deze manier ontwikkelt een evaluatiebeleid zich via gedeeld leiderschap.

Bij een gebrek aan communicatie dreigt evaluatie een aangelegenheid te worden van de individuele leerkracht. Door in te zetten op zowel formele als informele vormen van communicatie wordt het dissemineren van professionaliseringsinitiatieven net gefaciliteerd. Ook bij de komst van nieuwe leerkrachten in het team is het van belang dat ze ingeleid worden in de visie die de school heeft op evaluatie en de manier waarop deze visie zich vertaalt in de evaluatiepraktijk.

De casestudies tonen aan dat communicatie des te belangrijker wordt als het gaat over gedifferentieerde evaluatie. Gedifferentieerd evalueren betekent dat men evaluatie afstemt op de noden van de leerling, maar omdat elke situatie anders is, ervaren leerkrachten dat ze geen beroep kunnen doen op een standaard scenario. Wel is het zo dat leerkrachten hier steeds vaardiger in worden nét omdat ze zich kunnen beroepen op eerdere ervaringen. Communicatie moet er daarom naar streven om dergelijke ervaringen om te zetten naar gedeelde expertise en zo vermijden dat gedifferentieerde evaluatie vooral als het domein van de zorgleerkracht/-coördinator gezien wordt.

### **Samenwerking**

Tot slot blijkt uit de gesprekken met leerkrachten en directies uit de twaalf deelnemende scholen dat samenwerking onontbeerlijk is om tot een gedragen en gedeelde evaluatiepraktijk te komen. Het belang van samenwerking is heel nauw verweven met het belang van professionalisering en

communicatie. Het is precies via samenwerking dat er kennis verworven én gedeeld kan worden en dat men van gedachten kan wisselen over hoe men de visie op evaluatie kan vertalen naar de evaluatiepraktijk. Bovendien geven leerkrachten in het onderzoek aan dat samenwerken met collega's ervoor zorgt dat ze zich minder onzeker voelen over hun competenties inzake leerlingevaluatie. Samenwerking leunt zo aan bij interne professionalisering van het team: de leerkrachten met beperkte ervaring in het basisonderwijs geven aan dat ze veel leren via collega's die al langer in het vak staan. Hier moet aan toegevoegd worden dat de interne samenwerking bevorderd kan worden door de wijze waarop een school haar werking organiseert. Er werd veel nauwer samengewerkt in basisscholen waar er parallelklassen zijn en/of waar er systematisch klasdoorbrekend gewerkt wordt voor bepaalde leergebieden. In het secundair onderwijs laten de casestudies zien dat er veel samengewerkt wordt tussen leerkrachten via de vakgroepen, zowel binnen een graad, als over de graden heen.

Gezien het feit dat evalueren en leren onlosmakelijk met elkaar verbonden zijn, draagt samenwerking tevens bij tot de algemene onderwijskwaliteit. Bijvoorbeeld door te waken over horizontale en verticale afstemming in de evaluatiepraktijk. Via de horizontale afstemming wordt er onder parallelcollega's gestreefd naar bijvoorbeeld eenzelfde interpretatie van evaluatiecriteria en/of eenzelfde weging van deze criteria om te beoordelen of een leerdoel al dan niet behaald werd. De verticale afstemming verwijst naar de mate waarin er een lijn zit in de wijze van evalueren over de leerjaren heen: slagen de gekozen of ontwikkelde evaluatie-instrumenten erin om de verschillende vaardigheidsniveaus te evalueren? Of bijvoorbeeld met betrekking tot zelf- en peerevaluatie over de jaren heen: worden leerlingen in de eerste graad hierin begeleid zodat ze zich deze wijze van evalueren meer eigen kunnen maken tegen de derde graad?

#### **4. De resultaten samengevat**

De bovenstaande bevindingen vatten we samen aan de hand van de vier onderzoeksvragen die ingaan op (1) het evaluatiebeleid van scholen en de evaluatiepraktijken van leerkrachten, (2) de bepalende factoren hierbij, (3) de opvattingen hieromtrent en ten slotte (4) de competenties inzake evalueren.

### **Onderzoeksvraag 1: Wat is het evaluatiebeleid, en wat zijn de evaluatiepraktijken in Vlaamse scholen in basis- en secundair onderwijs?**

#### **1A) Evaluatiebeleid**

Het evaluatiebeleid in basis- en secundair onderwijs lijkt in het overgrote deel van de scholen beperkt en/of heel gefragmenteerd ingevuld te zijn.

Met *beperkte* invulling wordt bedoeld dat het beleid doorgaans oppervlakkig ingevuld is en vaak uit (ongeschreven) schoolbrede afspraken bestaat omtrent hoe vaak men evalueert, wanneer en hoe dit gecommuniceerd wordt via rapporten en ouder- en/of leerlingcontacten. Daarbij zijn ook steeds afspraken omtrent de frequentie waarmee begeleidende klassenraden georganiseerd worden, maar ook hier blijft de invulling ervan soms wat oppervlakkig. Kortom, het 'beperkte' beleid is gestoeld op de invulling van praktische vragen zoals, 'wat', 'wanneer', 'hoe vaak' en 'wie'. Het gaat om afspraken die met andere woorden eerder samengevat staan in de schoolkalender dan dat ze voortvloeien uit een visietekst. De waarom-vraag krijgt hierbij geen of heel beperkte aandacht waardoor evaluatiepraktijken eerder gebaseerd zijn op traditie en gewoonte in plaats van dat ze voortvloeien uit een bepaalde visie. De vrijheid waarover een school beschikt om een eigen evaluatiebeleid uit te



werken lijkt op deze manier te worden doorgeschoven naar de individuele leerkracht, wat verklaart waarom de evaluatiepraktijk van leerkrachten weinig gedeeld wordt binnen eenzelfde school.

Met *gefragmenteerde* invulling verwijzen we naar de vaststelling dat de ‘waarom-vraag’ in sommige scholen wel aandacht krijgt op schoolniveau, maar dat het antwoord op deze vraag zich beperkt tot één van de verschillende functies die evaluatie omvat. We denken aan het voorbeeld van de school uit het casestudieonderzoek dat een impliciet beleid heeft omtrent het nemen van beslissingen over de schoolloopbaan waarbij bedoeld wordt op het uitreiken van A-attesten, zonder een beleid uit te werken dat inzet op de ondersteunende functie van evaluatie net opdat leerlingen hun A-attest kunnen krijgen. Of de methodeschool in het casestudieonderzoek die sterk de ondersteunende functie van evaluatie benadrukt, maar hierbij uit het oog verliest om als school door middel van gevalideerde toetsen te reflecteren over de eigen onderwijspraktijk. Een focus op slechts één van de verschillende functies van evaluatie, maakt dat de overige functies van evaluatie een blinde vlek worden wat zich vertaalt in een evaluatiepraktijk die ‘onbewust’ kansen laat liggen om leer- en ontwikkelingskansen op niveau van de leerling, de leerkracht en/of de school te bevorderen. Dit maakt dat we alsnog uiteenlopende verschillen vaststellen tussen het evaluatiebeleid van scholen onderling.

### **1B) Evaluatiepraktijk**

Hoewel de evaluatiepraktijk van scholen niet gedeeld wordt op het niveau van de school, stellen we vast dat de algemene evaluatiepraktijk sterk gekenmerkt wordt door een strikte scheiding tussen evaluatie en algemene instructie-activiteiten. Omwille van een gebrek aan beleid op schoolniveau zien we dat de evaluatiepraktijk voor een groot deel samenhangt met de structuur van het onderwijs. In het basisonderwijs verloopt evaluatie vaker (semi-)geïntegreerd dan in het secundair onderwijs. Binnen het secundair onderwijs stellen we vast dat de evaluatiepraktijk en dagdagelijkse lespraktijk sterker geïntegreerd zijn het bso. Een mogelijke verklaring voor deze verschillen zou gevonden kunnen worden in het soort doelen dat vooropgesteld wordt: in de onderwijsniveaus en/of onderwijsvormen waarin kennisgerichte doelen meer op de voorgrond treden, lijkt evaluatie meer losgekoppeld te worden van de dagdagelijkse klaspraktijk. In de praktijkgerichte vakken van het bso en leggen de eindtermen en leerplannen meer nadruk op competenties en vaardigheden in vergelijking met de vakken in het aso en tso die meer kennisgericht zijn. De casestudies brachten aan het licht dat niet uitsluitend het onderwijsniveau of de onderwijsvorm bepalend is voor de praktijk, maar dat ook een specifiek schoolbeleid inzake evaluatie een rol kan spelen. De methodeschool uit het basisonderwijs, waar de evaluatiepraktijk sterk geïntegreerd is in de onderwijspraktijk, vormt hiervan een voorbeeld. De visie in deze school ligt aan de basis van de evaluatiepraktijk en het feit dat deze visie op evaluatie ingebed is in een ruimere visie op onderwijs en leren dat oog heeft voor de individuele leerling.

De evaluatiepraktijk van de onderzochte scholen op transitie momenten (overgang tussen leerjaren, overgang van kleuter naar lager en van lager naar secundair) blijkt onderling verschillen te tonen. Deze transities gaan vaak gepaard met het uitreiken van een getuigschrift of een attest en hiervoor blijken scholen verschillende – niet altijd helder gecommuniceerde – criteria te hanteren. In de zes cases van het basisonderwijs was er op dit vlak een verschil merkbaar tussen scholen die gekenmerkt worden door een leerlingenpubliek met hoge SES en met een lage SES. De scholen met een hoge SES bij hun leerlingenpopulatie hanteren een *enge* aanpak waarbij beroep gedaan wordt op een cijfercriterium om te bepalen of een leerling het getuigschrift ontvangt. Bijkomend volgt er een delibererend overleg in beperkte kring waarbij enkel de directeur en/of zorgleerkracht samen met de leerkracht van het zesde leerjaar beslist over deze leerlingen die het cijfercriterium niet halen. In de scholen met een lage

SES bij de leerlingenpopulatie is er veel meer sprake van een *brede* aanpak: het hele team van leerkrachten wordt betrokken bij het nemen van de beslissing en er wordt verder geen cijfercriterium gehanteerd, integendeel, het totale plaatje van de leerling wordt meegenomen.

Het gebruik van genormeerde en gestandaardiseerde toetsen is over het algemeen heel beperkt, met uitzondering van één case waar ze ingezet worden voor het bewaken en monitoren van de onderwijskwaliteit. In de meerderheid van de cases is het gebruik van gestandaardiseerde toetsen beperkt tot het minimum en gebeurt dit vooral op aansturen van externe actoren zoals inspectie, of omdat het een wettelijke verplichting is.

## **Onderzoeksvraag 2: Wat zijn de bepalende factoren bij evaluatiebeleid en – praktijk in Vlaamse scholen in het basis- en secundair onderwijs?**

Een aantal contextfactoren van scholen lijken een bepalende factor te zijn voor de evaluatiepraktijken van leerkrachten. Zo blijkt uit zowel het casestudie- als het surveyonderzoek dat het onderwijsniveau een rol speelt: in het basisonderwijs verloopt evaluatie meer geïntegreerd in de dagdagelijkse onderwijspraktijk dan in het secundair onderwijs. Binnen dit laatste niveau stellen we een bijkomend verschil vast naargelang de onderwijsvorm: de evaluatiepraktijk en dagdagelijkse lespraktijk zijn minder sterk gescheiden in het bso. Dit kan verklaard worden door de aard van de vakken in het bso: bij praktijkgerichte vakken staan vaardigheden centraal en vindt de evaluatie/het bijsturen van deze vaardigheden meer plaats terwijl de leerlingen deze vaardigheden oefenen.

Verder blijkt de leerlingencompositie van de school een bepalende factor te zijn voor de evaluatiepraktijk van leerkrachten: hoe meer het leerlingenpubliek van een school gekenmerkt wordt door een hoge SES, hoe minder de leerkrachten zullen kiezen voor een (semi-)geïntegreerde evaluatiepraktijk. Leerkrachten zullen hun evaluatiepraktijk meer integreren in de algemene lespraktijk als ze les geven op een school die gekenmerkt wordt door een groter aandeel lage SES-leerlingen.

Op basis van zowel de kwalitatieve als de kwantitatieve onderzoeksdata stellen we, naast de factoren die onveranderlijk zijn, ook factoren vast die er toe doen en die bovendien kaderen binnen de bewegingsruimte van een school. Zo blijkt het beleidsvoerend vermogen van de school een grote rol te spelen. Een sterker beleidsvoerend vermogen hangt samen met een geïntegreerde evaluatiepraktijk van leerkrachten. Bij scholen met een sterker beleidsvoerend vermogen lieten de casestudies zien dat deze scholen een evaluatiebeleid hebben dat inhaakt op overige aspecten van het schoolbeleid. Zo is er een duidelijke link met het algemene onderwijsbeleid, het zorgbeleid en de visie op leren die men als uitgangspunt neemt. Hoe sterker men het leren van elke leerling wil bevorderen, hoe meer de evaluatiepraktijken zullen ingebed zijn de ruimere onderwijsactiviteiten.

De rol van de directeur is cruciaal in de ontwikkeling van een evaluatiebeleid, het uitdragen en verspreiden ervan en het afstemmen van de evaluatiepraktijk op het beleid. De directie heeft een katalysatorfunctie in het faciliteren van samenwerking tussen leerkrachten en het professionaliseren van het team waardoor de betrokkenheid van leerkrachten verhoogd wordt om mee het evaluatiebeleid vorm te geven. Professionalisering van het team specifiek met betrekking tot evaluatie blijkt een factor te zijn die er toe doet om enerzijds een gedragen evaluatiebeleid te ontwikkelen op school. Anderzijds blijkt de mate van professionalisering ook sterk samen te hangen met een evaluatiepraktijk van leerkrachten die gekenmerkt wordt door een sterkere integratie in het



onderwijsproces. De casestudies boden inzicht in een aantal onderliggende factoren die dit beleidsvoerend vermogen inzake evaluatie versterken. Aansluitend bij 'professionalisering' vormt 'samenwerking tussen leerkrachten onderling' een bepalende factor. In zekere zin kan dit beschouwd worden als een vorm van professionalisering: het zet aan tot het reflecteren over de eigen evaluatiepraktijk en de afstemming ervan met de praktijk van collega's en het evaluatiebeleid van de school. De mate van samenwerking is afhankelijk van de cultuur die er hieromtrent is op de school. In het secundair onderwijs stellen we vast dat de vakgroepen hier een centrale rol in spelen, al is de kwaliteit van de samenwerking afhankelijk van vakgroep tot vakgroep. Hierbij aansluitend speelt ook het aspect van 'communicatie' een rol: niet alleen bij leerkrachten onderling over hoe ze hun evaluatiepraktijk aanpakken, maar ook met betrekking tot de voortgang van het leerproces van leerlingen. Het aspect van 'communicatie' reikt bovendien verder dan het communiceren tussen leerkrachten onderling: het betreft eveneens de mate waarin men aandacht heeft voor communicatie over het leerproces met de leerling zelf en de ouders. In scholen die inzetten op leerlingcontacten en deze meer dan twee keer per schooljaar organiseren, blijkt de evaluatiepraktijk van leerkrachten meer gescheiden te zijn van de onderwijsactiviteiten. Verder zijn er verschillen merkbaar in de mate waarin scholen inzetten op communicatie met ouders over het leerproces van hun kinderen. Hoewel alle scholen hiervoor gebruik maken van oudercontacten, zien we dat basisscholen met een meer diverse leerlingenpopulatie hiervoor ook vaker beroep doen op informelere contacten.

Naast de factoren op schoolniveau blijken ook factoren op het niveau van de individuele leerkracht bepalend te zijn voor de mate waarin ze aansluiten bij een gescheiden of (semi-)geïntegreerd model. Zo blijkt uit het surveyonderzoek dat de opvattingen van leerkrachten een rol spelen: leerkrachten die overtuigd zijn van de informerende waarde van evaluatie zullen in hun praktijk aansluiten bij het gescheiden model, leerkrachten die eerder kritisch zijn ten aanzien van evaluatie omdat ze menen dat evaluatie meetfouten kan bevatten en oneerlijk kan zijn voor sommige leerlingen, zullen meer het geïntegreerde model toepassen. Tegelijk blijkt ook dat die leerkrachten die zichzelf bekwaam voelen om diversiteit onder leerlingen op te merken vaker het (semi-)geïntegreerde model zullen aanwenden. Mogelijks gaan deze leerkrachten specifiek op zoek naar professionalisering, precies omdat zij met de diversiteit onder leerlingen ook aan de slag willen gaan. Leerkrachten die ervan overtuigd zijn dat wat ze doen in de praktijk ook een positief effect heeft op het leren en ontwikkelen van leerlingen, zullen ook vaker aansluiten bij het geïntegreerde model. Ten slotte bleek ook dat het geslacht van de leerkracht een positief verband heeft met het geïntegreerde model: mannen zouden eerder dan vrouwen gebruik maken van evaluatie die sterk ingebed is in het ruimere onderwijsleerproces. Dit verband houdt bovendien stand na controle voor onderwijsvorm. Aangezien dit betekent dat de leerlingen zelf heel sterk betrokken worden, sluit dit positieve effect van 'geslacht' aan bij eerder onderzoek dat aantoont dat mannen meer dan vrouwen inzetten op het ondersteunen van autonomie (Lietaert & Consuegra, 2015).

### **Onderzoeksvraag 3: Wat zijn de opvattingen van directies, leerkrachten en leerlingen op vlak van evaluatie in het algemeen en meer specifiek met betrekking tot de evaluatiepraktijk en het - beleid van hun school?**

Het onderzoek laat zien dat de opvattingen van leerkrachten met betrekking tot evaluatie wisselend zijn. Enerzijds is men vooral overtuigd van de informerende waarde van evaluatie. Hoewel het informerend karakter van evaluatie kan verwijzen naar informerend ten aanzien van de leerkracht, de school, als de leerling zelf zien we dat de nadruk voornamelijk ligt op de leerling en de leerkracht. De

bevroegde leerkrachten in de cases van het kleuter en lager onderwijs leggen hierbij vooral de nadruk op het belang van de formatieve functie van evaluatie: men meent dat evaluatie in de eerste plaats het leerproces van de leerling moet ondersteunen. Evaluatie vindt volgens hen het best voortdurend plaats zodat er actie ondernomen kan worden van zodra men ergens moeilijkheden vaststelt. Bij de opvattingen bij leerkrachten uit het secundair onderwijs verwijst het informerende karakter in mindere mate naar het feit dat de informatie kan aangewend worden om het leer-en onderwijsproces bij te sturen. In het secundair onderwijs leggen de bevroegde leerkrachten meer de nadruk op het informeren, zonder dat hier bijsturingen uit voortvloeien. Het informeren wordt enger ingevuld waarbij evaluatie dient om de leerling en de ouders op de hoogte te houden over de stand van zaken in het leerproces. De overtuiging dat evaluatie impact heeft op de eigen onderwijsactiviteiten is minder sterk aanwezig in de cases van het secundair onderwijs. In het secundair vinden we meer leerkrachten die naar de oriënterende functie verwijzen, namelijk dat de evaluatie dient om beslissingen te nemen over de verdere schoolloopbaan van de leerlingen. De verschillen in opvattingen bij leerkrachten uit basis- en secundair onderwijs zijn analoog met de trend we vaststellen in de evaluatiepraktijk: de mate van geïntegreerdheid die gemiddeld genomen afneemt naargelang men zich op een hoger onderwijsniveau bevindt.

Anderzijds stellen we, weliswaar in mindere mate, ook een negatievere houding vast ten aanzien van evaluatie bij leerkrachten. Zo is men kritisch ten aanzien van evaluatie omdat men ervan overtuigd is dat evaluatie meetfouten kan bevatten en oneerlijk kan zijn voor sommige leerlingen. Deze opvattingen zijn typerend voor leerkrachten die les geven op een school met een leerlingenpubliek dat gekenmerkt wordt door een lage SES. Dit verband lijkt logisch te zijn als we ervan uitgaan dat 'evaluatie' gepercipieerd wordt als *one-size-fits-all*. Deze leerkrachten zullen vaker gebruik maken van evaluatie die sterk ingebed is in de lespraktijk, ze zullen leerlingen meer betrekken bij het evalueren en hun algemene differentiatiepraktijk zal men ook meer doortrekken naar de evaluatiepraktijk. Dit precies om tegemoet te komen aan mogelijke meetfouten of oneerlijke gevolgen voor de leerling. Deze leerkrachten zijn meer geneigd om rekening te houden met de noden van leerlingen, ook bij evaluatie.

Met betrekking tot gestandaardiseerde en centrale examens stellen we vast dat leerkrachten hier niet steeds een onderscheid in maken. Leerkrachten en directies houden er over het algemeen een kritische houding op na. In het basisonderwijs hebben alle respondenten bedenkingen bij centrale examens. Hierbij wordt vooral verwezen naar de nadelen zoals uniformisering van het curriculum en een trend naar een nóg prestatiegerichter onderwijssysteem. Bovendien is men het er unaniem over eens dat dergelijke uitkomsten van één momentopname in geen geval invloed mogen hebben op de studiekeuze van 12 jarigen die nog sterk evolueren. Enkel in basisscholen waarbij het leerlingenpubliek gekenmerkt wordt door een hoge SES, zijn er leerkrachten die positieve elementen aanhalen. Volgens hen is het nuttig om te kunnen vergelijken met anderen, om een duidelijk beeld te brengen naar ouders of leerlingen die nog twijfelen over hun studiekeuze. Bij respondenten wiens werkcontext gekenmerkt wordt door meer diversiteit werden geen positieve opvattingen vermeld. Deze leerkrachten verwijzen dan ook naar de diversiteit in hun leerlingenpubliek: namelijk dat een uniform examen haaks staat op een klaspraktijk die inzet op differentiatie.

In het secundair onderwijs zijn de houdingen vergelijkbaar met deze in het basisonderwijs. Waar er in het basisonderwijs een groep leerkrachten is met een ambivalente houding die zowel naar negatieve als positieve elementen verwijzen, stellen we in de onderzochte scholen van het secundair onderwijs vast dat er een minderheid is aan leerkrachten die duidelijk voorstander is van centrale examens.

Opvallend is dat de meerderheid van deze groep uit één specifieke case komt en centrale examens ziet als een tegenwicht voor het evaluatiebeleid van hun school dat volgens hen inzet op het 'onterecht' uitreiken van A-attesten. Deze uitgesproken houding pro centrale examens vindt zijn oorsprong vooral in het attesteringsbeleid van deze school: de leerkrachten voelen namelijk druk vanuit het schoolbeleid om zoveel mogelijk A-attesten uit te reiken, wat volgens hen leidt tot 'onterechte A-attesten' en uiteindelijk tot minderwaardige diploma's. De opvattingen van deze leerkrachten worden hier specifiek bepaald door de context waarin ze werken: in hun ogen vormt centrale examinering een oplossing voor de uitwassen van het evaluatiebeleid van de school waarbij de nadruk ligt op het uitreiken van A-attesten zonder dat ingezet wordt op een beleid omtrent zorg en remediëring.

Ten slotte stelden we een verband vast tussen de opvattingen van leerkrachten ten aanzien van evaluatie en het beleid van de school omtrent evaluatie. Het surveyonderzoek laat zien dat opvattingen omtrent evaluatie nauwelijks gedeeld worden onder leerkrachten. Dit bevestigt de hypothese uit het casestudieonderzoek dat het bij de meerderheid van de scholen ontbreekt aan een gedeeld en gedragen evaluatiebeleid. In de minderheid van de scholen uit het casestudieonderzoek waar er wel een expliciet of impliciet evaluatiebeleid aanwezig is, stellen we vast dat de gedeelde opvattingen onder leerkrachten zich specifiek verhouden tot het evaluatiebeleid van de school. Deze verhouding kan positief zijn bijvoorbeeld in het geval waar het expliciete evaluatiebeleid van de school gestoeld is op een visie die gedragen wordt door leerkrachten waarbij men ervan overtuigd is dat evaluatie dient om het leren van leerlingen te optimaliseren. Anderzijds kan de verhouding ook net negatief zijn zoals in het voorbeeld van de school waarbij de directeur impliciet aanstuurt op het uitreiken van A-attesten waardoor dit leidt tot een sterkere overtuiging in centrale examens bij het team van leerkrachten.

#### **Onderzoeksvraag 4: In welke mate voelen leerkrachten zich competent om te evalueren? Voelen ze zich even competent om competentiegericht als kennisgericht te evalueren? Waardoor wordt deze competentie – of het gebrek eraan- bepaald?**

Het surveyonderzoek laat zien dat leerkrachten met een hogere doelmatigheidsbeleving meer geïntegreerd zullen evalueren. Deze leerkrachten geloven sterk in hun eigen competenties met betrekking tot instructiestrategieën, klasmanagement en leerlingenbetrokkenheid. In het casestudieonderzoek stellen we vast dat leerkrachten met weinig ervaring hier nog onvoldoende sterk in zijn: zij voelen zich vaker onzeker als het gaat over hun evaluatiepraktijk en verwijzen naar moeilijkheden zoals doelgericht evalueren en het integreren van evaluatieactiviteiten tijdens momenten van instructie.

De gesprekken met leerkrachten uit de casestudies wijzen op een grote onzekerheid bij leerkrachten, als het gaat over evaluatie waarbij men rekening houdt met diversiteit. Leerkrachten – ook de meer ervaren leerkrachten – geven aan dat het nog steeds zoeken is naar manieren om dit nog anders/nog beter te doen. Op dit vlak lijken alle leerkrachten aan te geven dat er nog groeimarge is. Deze bevinding sluit aan bij het feit dat leerkrachten over het algemeen hoofdzakelijk inzetten op een eerder gescheiden evaluatiepraktijk dan een (semi-)geïntegreerde praktijk die meer beantwoordt aan de diversiteit onder leerlingen. De bekwaamheid om diversiteit onder leerlingen op te merken lijkt hier een bepalende rol in te spelen.

Er kan besloten worden dat de kennisgerichte evaluatie doorheen de jaren plaats gemaakt heeft voor een evolutie naar meer competentiegerichte evaluatie, al voelen leerkrachten zich nog onzeker om competentiegericht te evalueren. In het secundair onderwijs liggen de opvattingen én de ervaringen van leerkrachten verder uit elkaar: afhankelijk van het soort vak en de bijhorende leerplannen zal er sprake zijn van een meer competentiegerichte evaluatiepraktijk. Praktijkgerichte vakken in kso en bso lenen zich meer tot competentiegerichte evaluatie dan kennisgerichte evaluatie.

## IV Discussie

### 1. Onderzoek naar evaluatie in een context van diversiteit

De evolutie naar een steeds diverser wordende leerlingenpopulatie doet zich niet enkel voor in de Vlaamse context, maar is een wereldwijd fenomeen (Banks & al., 2007). Dit uit zich in doelstellingen van internationale samenwerkingsverbanden, zo stelt het Verdrag van de Verenigde Naties (2006) inclusief onderwijs als doel voorop om het recht op kwaliteitsvol en toegankelijk onderwijs wereldwijd te garanderen. Het VN-verdrag werd door België geratificeerd in 2009, wat betekent dat België sindsdien moet kunnen aantonen op welke wijze ingezet wordt op het realiseren van zo kwaliteitsvol mogelijk onderwijs voor alle leerlingen. De definiëring van inclusief onderwijs is daarenboven door de jaren heen getransformeerd (Opertti et al., 2013). Een doelgroepenbenadering die evolueerde van een focus op leerlingen met een handicap en later minderheidsgroepen (kansarmen, migratieachtergrond, ...) ruimde plaats voor een systeembenadering. In de huidige benadering van het VN-verdrag ligt de klemtoon op het zodanig inrichten van het onderwijssysteem opdat álle leerlingen bereikt worden, ongeacht hun specifieke achtergrond of eigenschap(pen) waarmee ze aan een (doel)groep gelinkt worden. De idee van een doelgroep die nood heeft aan specifieke maatregelen wordt hier opgeheven doordat het onderwijssysteem maximale kansen creëert voor álle leerlingen. Onderwijs wordt dan op zodanige wijze ingericht dat alle lerenden samen kunnen leren, ook al verschillen ze onderling op vlak van gender, sociale klasse, etnische herkomst, cognitieve capaciteiten, ... Op deze manier vormt inclusiviteit een waarde die bijdraagt aan een gunstig leerklimaat. Dit impliceert immers respect ten aanzien van alle leerlingen en tegelijkertijd omvat dit het engagement om álle leerlingen kansen te bieden om minimaal de eindtermen te bereiken en ondertussen hun talenten maximaal te ontwikkelen (Van Avermaet & Sierens, 2010).

Omgaan met diversiteit binnen een inclusieve omgeving, vormt voor scholen en leerkrachten een uitdaging. Een recente bevraging bij Vlaamse leerkrachten bracht aan het licht dat men nog heel wat vragen heeft over hoe men hiermee moet omgaan in de dagdagelijkse klaspraktijk (Vantieghem, Van Avermaet, Groenez, & Lambert, 2018). Ook specifiek met betrekking tot evaluatie ervaren scholen en leerkrachten de trend naar inclusie als een uitdaging (Struyf et al., 2018). Leerlingenevaluatie maakt niet alleen deel uit van deze dagdagelijkse praktijk, maar is tegelijkertijd uitermate cruciaal voor de realisatie van maximale leer- en ontwikkelingskansen.

Leerkrachten baseren zich bijvoorbeeld op de informatie van evaluaties om beslissingen te nemen over studiekeuze en bepalen zo mee in welke onderwijsvorm de leerling terecht komt en/of de leerling al dan niet moet zittenblijven (Tierney, 2013). Bijgevolg beïnvloedt evaluatie ook de mogelijkheden binnen hoger onderwijs en de latere kansen op de arbeidsmarkt. Zo bepaalt leerlingenevaluatie ook mee de algemene levenskwaliteit binnen het volwassen leven. Hoewel dit cruciale aspect van evaluatie vooral verbonden lijkt met de functie van 'beslissingen nemen over de schoolloopbaan', zijn de overige functies van evaluatie minstens even cruciaal als het gaat over hun invloed die ze (on)rechtstreeks uitoefenen leerlingen hun toekomst.

De informerende functie schept een beeld over de leerling en diens leerproces. Verschillende studies toonden aan dat de evaluatiepraktijk een hindernis kan vormen voor bepaalde (groepen van) leerlingen om te laten zien wat ze kunnen (De Backer et al., 2016; Nayir et al., 2019; Shohamy, 2001). Afhankelijk van welk gevolg men geeft aan evaluatie, leidt de ondersteunende functie van evaluatie tot kwaliteitsvol onderwijs voor de leerling. Op deze manier kan evaluatie aangewend worden als motor om te differentiëren en maximaal aan te sluiten bij de leerbehoeften van leerlingen (Tomlinson & Moon, 2013a, 2013b). De vaststelling dat leerkrachten zichzelf niet bekwaam voelen om tegemoet te komen aan de noden van leerlingen (De Boer et al., 2011) doet vermoeden dat het potentieel van evaluatie in het kader van differentiatie nog onvoldoende benut wordt.

Elke functie van evaluatie blijkt tegelijk cruciaal én gevoelig te zijn in het realiseren van maximale leer- en ontwikkelingskansen van alle leerlingen. Via de centrale onderzoeksvragen binnen dit onderzoek gingen we na hoe deze functies aangewend worden in het Vlaamse basis- en secundair onderwijs. Specifiek gingen we na wat het evaluatiebeleid en de evaluatiepraktijken van scholen betekenen voor de realisatie van maximale leer- en ontwikkelingskansen. In wat volgt, gaan we dieper in op de bevindingen om de mechanismen die hieraan gerelateerd zijn bloot te leggen.

## **2. Formatieve functie van evaluatie onderbenut**

De evaluatiepraktijk van leerkrachten in zowel het basis- als het secundair onderwijs wordt over het algemeen gekenmerkt door een gescheiden evaluatiepraktijk waarbij evaluatie en instructie los van elkaar staan. De evaluatiepraktijk staat op deze manier sterk in het teken van ‘evalueren van het leren’ en vormt aldus het eindpunt. Een meer geïntegreerde evaluatiepraktijk daarentegen komt veel meer tegemoet aan de leer- en ontwikkelingsbehoeften van leerlingen. De leer- en ontwikkelingsbehoeften worden in de eerste plaats veel meer opgemerkt door de leerkracht, waardoor hier op ingespeeld kan worden om leerlingen te ondersteunen in het bereiken van de eindtermen en ontwikkelingsdoelen. Bij geïntegreerde evaluatiepraktijken worden de uitkomsten van evaluatie door de leerkracht gebruikt om enerzijds het onderwijsproces bij te sturen (Havnes et al., 2012; Wiliam, 2011). Anderzijds dient de uitkomst van evaluatie om het leren van de leerling rechtstreeks te bevorderen, bijvoorbeeld via het geven van feedback (Clark, 2012; Hattie, 2012). Via beide manieren wordt evaluatie ‘formatief’ aangewend: de uitkomst van evaluatie vormt het beginpunt van een nieuwe, bijgestuurde onderwijs- en leersequens. Dit ‘output = input – principe’ werd door verschillende auteurs naar voren geschoven als cruciaal voor een gedifferentieerde onderwijspraktijk die streeft naar een inclusievere leeromgeving (Struyven, Gheysens, Coubergs, De Doncker, & De Neve, 2019; C. A. Tomlinson, 2001). Dit blijkt ook uit de resultaten van het surveyonderzoek (Ysenbaert et al., n.d.) waarbij er een positief verband is vastgesteld tussen de algemene differentiatiepraktijk van leerkrachten en een evaluatiepraktijk die meer geïntegreerd is in de onderwijspraktijk. Een gescheiden evaluatiepraktijk liet daarentegen een negatief verband zien met de aandacht die men besteedt aan differentiatie. De onderwijsactiviteiten binnen een gescheiden evaluatiemodel worden niet afgestemd op de uitkomsten van evaluatie. Deze onderwijsactiviteiten zijn eerder afgestemd op de ‘gemiddelde leerling’ in plaats van dat er rekening gehouden wordt met specifieke leernoden.

Wanneer we de resultaten met betrekking tot de ondersteunende functie van evaluatie bekijken vanuit de opzet van dit onderzoek, namelijk de mate waarin evaluatie bijdraagt tot maximale leer- en ontwikkelingskansen bij alle leerlingen, merken we twee zaken op. Ten eerste stellen we vast dat de ondersteunende functie van evaluatie, hoewel belangrijk geacht door de leerkrachten, over het

algemeen slechts beperkt tot uiting komt in hun praktijk. Meer zelfs, de evaluatieactiviteiten van leerkrachten worden eerder gekenmerkt door het feit dat ze los staan van de algemene onderwijsactiviteiten. De ‘output’ van evaluatie wordt dus beperkt als ‘input’ gebruikt, waardoor er kansen onbenut blijven om het leren van álle leerlingen te bevorderen. Dat er op dit vlak kansen onderbenut blijven, weerklonk ook bij de bevraagde leerlingen in het casestudieonderzoek (Ysenbaert et al., 2018). Zij verkiezen een frequenter gebruik van observatie door hun leerkracht en peerevaluatie met hun medeleerlingen omdat deze evaluatievormen volgens hen het meest aangewezen zijn om onmiddellijke feedback te verkrijgen. Het feit dat evaluatie ter ondersteuning van het leerproces vaak ontbreekt, sluit aan bij het recente PISA-onderzoek waar vastgesteld werd dat Vlaanderen, op vier landen na, achterop hinkt als het gaat over de mate waarin leerlingen zich geholpen voelen door hun leerkracht (De Meyer et al., 2019). Deze bevinding ligt bovendien in lijn met de vaststelling van het meest recente jaarverslag van de Onderwijsinspectie (2020) waaruit blijkt dat het geven van feedback een knelpunt vormt in de klaspraktijk van leerkrachten.

Naast de vaststelling dat evaluatiepraktijken kansen laten liggen om het onderwijs af te stemmen op de leernoden van leerlingen omdat ze overwegend ingericht worden volgens het gescheiden model, is het opmerkelijk dat geïntegreerde evaluatiepraktijken eerder aangewend worden in scholen met leerlingen met vooral een lage SES. In scholen waar de leerlingencompositie gekenmerkt wordt door een hoge SES worden de leerlingen overwegend geëvalueerd volgens het gescheiden model, waardoor ze in die scholen in mindere mate inspelen op individuele leernoden.

### **3. Summatieve functie overschaduwde de formatieve functie**

#### **3.1. ‘Waterval-denken’ bepalend voor beslissingen over de schoolloopbaan**

De summatieve functie van evaluatie verwijst naar het nemen van beslissingen over de schoolloopbaan van leerlingen. Dit gebeurt op meerdere momenten: bij de overgang van kleuter naar lager en van lager naar secundair, maar ook bij de transities van leerjaar X naar leerjaar Y. Hierbij zijn ook de keuzes inzake het doorstromen naar de A- of B-stroom en naar de onderwijsvormen aso, bso, kso en tso aan de orde. De manier waarop deze beslissingen genomen worden, verschilt van school tot school. Zowel in het basis- als in het secundair onderwijs zijn er onderlinge verschillen waaruit blijkt dat de transitie van leerlingen geen exacte wetenschap is (Ysenbaert et al., 2018).

In het basisonderwijs stellen we vast dat de beslissingen omtrent het uitreiken van een getuigschrift vaak op een ‘enge’ manier worden ingevuld. Hierbij speelt het cijfercriterium een primaire rol. Er volgt pas een (delibererend) overleg indien leerlingen het cijfercriterium niet halen. Bovendien gaat dit overleg in beperkte kring door waarbij beslist wordt door de directeur en/of zorgleerkracht samen met de leerkracht van het zesde leerjaar. Dit sluit aan bij de vaststellingen van het TRANSBASO-onderzoek dat stelt dat de studiekeuzes van basis- naar secundair onderwijs in Vlaanderen vooral bepaald worden door de cognitieve prestaties –specifiek voor Nederlands en wiskunde – van leerlingen (Goosen et al., 2017). Het cijfercriterium is overigens wisselend: waar in de ene school een drempel van 50% geldt voor de vakken Nederlands, Frans en wiskunde, hanteert een andere school het criterium van 60% voor deze vakken en wereldoriëntatie. We merkten op dat in scholen waar het leerlingenpubliek diverser is, er vaker een *brede* aanpak gehanteerd wordt, waarbij men zich meer richt op het totale plaatje van de leerling. Hierbij gaat ook meer aandacht naar de interesses, talenten en

toekomstbeelden van leerlingen. Bovendien streeft men er in deze brede aanpak naar om de beslissing te nemen met het hele team van leerkrachten dat ooit les gaf aan de leerling.

Vergelijkbaar met de resultaten van het onderzoek van Anthoons et al. (2004), stelt het casestudieonderzoek (Ysenbaert et al., 2018) vast dat de wijze waarop secundaire scholen delibereren sterk verschilt van school tot school. Deze verschillen doen zich voor op vlak van gegevensverzameling, bijvoorbeeld de mate waarin enerzijds dagelijks werk en anderzijds examenresultaten meegenomen worden. Er zijn ook verschillen de criteria die men hanteert en de mate waarin deze doorslaggevend zijn voor de deliberatiebeslissing, zoals bijvoorbeeld de attitude van een leerling. Deze variatie aan criteria kan verklaard worden door de pedagogische vrijheid die scholen hebben – naast de algemene richtlijnen – om te evalueren in welke mate de eindtermen, ontwikkelingsdoelen en leerplandoelen bereikt zijn door de leerlingen (Teijssen et al., 2017).

Een constante bij de meeste onderzochte scholen is dat het nemen van beslissingen over de schoolloopbaan van leerlingen sterk gestuurd wordt vanuit het ‘waterval-denken’. De oriëntering naar de b-stroom is een negatieve keuze: wanneer een leerling niet alle doelen bereikt heeft of niet het zesde leerjaar doorlopen heeft, wordt hij/zij georiënteerd naar de b-stroom. De mechanismen van het watervalstelsel duiken al op nog vóór de leerlingen starten in het secundair onderwijs. Dit ‘deficit-denken’ wordt in de meerderheid van de secundaire scholen doorgetrokken in de praktijk van oriënterende klassenraden. Hierbij lijkt ‘oriëntering’ pas aan de orde te komen indien vastgesteld wordt dat de slaagkansen voor volgend schooljaar in dezelfde studierichting klein zijn. De oriëntering is in die zin vooral een *negatieve* oriëntering. Dit ligt in lijn met eerder onderzoek in Vlaanderen dat aantoont dat adviezen omtrent studiekeuze veeleer gebaseerd zijn op basis van prestaties en studiehouding dan op basis van de belangstelling of talenten van leerlingen (Goosen et al., 2017; Nicaise et al., 2014).

Het feit dat de oriënteringspraktijk in de onderzochte scholen onderling verschilt, valt te verklaren door de pedagogische vrijheid waarover scholen beschikken (Teijssen et al., 2017). Deze vrijheid wordt echter niet aangegrepen om via oriëntering de leer- en ontwikkelingskansen van leerlingen te maximaliseren. Integendeel, de oriënteringspraktijk in de meerderheid van de onderzochte scholen wordt passief ingevuld en men laat zich hierbij leiden door de mechanismen van het watervalstelsel. Het gebrek aan een actieve oriënteringspraktijk waarbij frequent gewerkt wordt aan het in kaart brengen van talenten en interesses van leerlingen, maakt dat leerlingen minder kans maken om terecht te komen in een studierichting die aansluit bij hun talenten. Op deze manier vormt de oriënteringspraktijk van scholen een belemmering in het maximaliseren van ontwikkelingskansen van leerlingen.

### **3.2. Selecteren in plaats van oriënteren**

De oriënteringspraktijk die gestuurd wordt door het ‘waterval-denken’ resulteert in een praktijk die eerder gericht is op selecteren dan op oriënteren. Dit wordt geïllustreerd in het casestudieonderzoek (Ysenbaert et al., 2018), waar een categoriale aso-school via het advies bij een A-attest alsnog leerlingen stimuleert om het aso te verlaten. Door op deze manier reeds vroeg te selecteren, houdt de school in de hogere leerjaren van het aso enkel nog de sterkste leerlingen over.

Eerder onderzoek omtrent oriëntering in het Vlaamse onderwijs stelt bovendien dat de adviezen van leerkrachten beïnvloed worden door stereotypering (Thys, 2018; Vantieghem et al., 2018). De bril waardoor leerkrachten naar hun leerlingen kijken, blijkt bepalend voor het uiteindelijke



oriënteringsadvies. De socio-economische achtergrond van leerlingen blijkt door te wegen in het advies dat ze krijgen. Leerlingen met een hoge SES maken zo beduidend meer kans op een A-attest dan leerlingen die dezelfde resultaten behalen maar uit een gezin met lage SES komen (Vantieghem et al., 2018). De sociale bias in oriënteringsadviezen maakt dat de sociale selectie in onderwijs versterkt wordt. Dit verklaart niet alleen mee waarom het aandeel lage SES-leerlingen disproportioneel hoger is in het bso, maar illustreert tevens hoe de evaluatiepraktijk aan sociale bias onderhevig is.

Verder stellen we vast dat het mechanisme van evaluatie als selectiemechanisme gepaard gaat met een verminderde aandacht voor de ondersteunende functie van evaluatie. Dit betekent met andere woorden dat de alledaagse evaluatiepraktijk vaker gekenmerkt wordt door een gescheiden evaluatiepraktijk in die contexten waar een transitie in de schoolloopbaan nabij is. Twee vaststellingen wijzen hier op: (1) Het geïntegreerde model voor evaluatiepraktijk wordt gemiddeld genomen vaker aangewend in het lager onderwijs dan in het secundair onderwijs. Binnen dit onderwijsniveau is 'selecteren' veel minder aan de orde aangezien het studiekeuzeproses pas op het einde van het basisonderwijs relevant wordt. (2) Binnen de onderwijsvorm bso wordt er beduidend vaker geëvalueerd volgens het geïntegreerde model. Ook in het bso is 'selecteren' minder aan de orde aangezien het bso zich onderaan de hiërarchie van onderwijsvormen bevindt waardoor selecteren niet meer als een optie wordt gezien.

De spanning tussen de functie van enerzijds het nemen van beslissingen over de schoolloopbaan en anderzijds de functie van het ondersteunen van het leerproces, lijkt dus sterk onderhevig te zijn aan de eigenschappen van het onderwijssysteem, dat gekenmerkt wordt door vroege selectie die daarenboven beïnvloed wordt door sociale bias. Onderzoek van Bol et al. (2014) toont aan hoe centrale examinering meer gelijke kansen creëert in onderwijssystemen waarin leerlingen op jonge leeftijd een studiekeuze moeten maken. De studiekeuze wordt meer gebaseerd op objectieve indicatoren en is minder beïnvloed door sociale bias.

Opmerkelijk is dat in het Vlaamse onderwijs, dat eveneens gekenmerkt wordt door vroege studiekeuze, de ondersteunende functie van evaluatie pas aan de orde blijkt te zijn van zodra het nemen van beslissingen minder dwingend wordt. Dit sluit ook aan bij wat Bol et al. (2014) concluderen in hun eerder genoemde studie, namelijk dat het langer inzetten op comprehensief onderwijs of met andere woorden het uitstellen van de studiekeuze nog het meeste bijdraagt tot het creëren van gelijke onderwijskansen. Uit het casestudieonderzoek blijkt dat begeleidende klassenraden vaak inboeten aan 'formatieve' waarde omdat ze zich beperken tot het communiceren over bij welke leerlingen het moeilijk loopt. Mogelijke ondersteunende acties blijven zo te vaak uit, waardoor de begeleidende klassenraden indicatief zijn voor de delibererende klassenraden die een summatieve functie hebben. Door het uitblijven van concrete ondersteuning ten aanzien van leerlingen met moeilijkheden na een begeleidende klassenraad, vervaagt de grens tussen het formatieve en summatieve karakter van evaluatie. Dit fenomeen lijkt vooral op te duiken waar transities nabij zijn waardoor de oriënterende/selecterende functie van evaluatie de bovenhand neemt over de ondersteunende functie van evaluatie. In dit geval wordt formatieve evaluatie eerder summatief aangewend, wat het bevorderen van leer- en ontwikkelingskansen belemmert.

In het kader van het nastreven van maximale leerkansen voor alle leerlingen is de ondersteunende of formatieve functie van evaluatie echter cruciaal (zie 1.) Evaluatiepraktijken waarbij de formatieve functie overschaduw wordt door de summatieve functie – die daarenboven aangestuurd wordt door

onbewuste selectiemechanismen – zijn allerminst bevorderlijk voor maximale onderwijs- en ontwikkelingskansen. Deze selectiemechanismen belemmeren leerlingen om in een onderwijsvorm of studierichting terecht te komen die het beste aansluit bij hun talenten en interesses.

## 4. De rol van het beleidsvoerend vermogen

### 4.1. Gebrek aan geïntegreerd beleid

In de voorgaande sectie zagen we hoe een overwegend gescheiden evaluatiepraktijk van leerkrachten verklaard kan worden door specifieke eigenschappen van het onderwijssysteem, wat aansluit bij de bevindingen van Sach (2015). De evaluatiepraktijken van leerkrachten zijn echter niet uitsluitend onderhevig aan kenmerken van het onderwijssysteem zelf, maar zijn ook het resultaat van hoe verschillende niveaus van dit systeem op elkaar inwerken (Ercikan & Solano-Flores, 2016). Dit brengt ons bij een bijkomende verklaring voor de overwegend gescheiden evaluatiepraktijken van leerkrachten: het samenspel van het school- of mesoniveau en het beleid op macroniveau. Zowel het surveyonderzoek (Ysenbaert et al., n.d.) als het casestudieonderzoek (Ysenbaert et al., 2018) laten zien dat er op schoolniveau doorgaans geen sprake is van een gedeelde cultuur als het over evaluatie gaat, noch op vlak van praktijken noch op vlak van opvattingen. Hoewel Vlaamse scholen over een grote autonomie beschikken om vorm te geven aan een eigen evaluatiebeleid (Penninckx, Vanhoof, & Van Petegem, 2011), slaagt de overheid er te weinig in om scholen te stimuleren in het benutten van deze autonomie om een krachtig evaluatiebeleid te ontwikkelen. Het is precies de afwezigheid van een evaluatiebeleid op schoolniveau die ervoor zorgt dat een overwegend gescheiden evaluatiepraktijk het resultaat is van het systeem dat gekenmerkt wordt door selectie. Hierbij vallen leerkrachten terug op een concept van evaluatie waarbij dit veeleer beschouwd wordt als het eindpunt van een onderwijsproces dan dat dit het beginpunt vormt of ingebed is in het onderwijsproces. Het survey- en het casestudieonderzoek laten namelijk zien dat het beleidsvoerend vermogen van de school er toe doet in de realisatie van evaluatiepraktijken die meer ingebed zijn in het algemene onderwijsproces.

Een krachtig evaluatiebeleid wordt gekenmerkt door een visie op evaluatie die voortvloeit uit een visie op onderwijs en leren, waarbij het accent ligt op het bevorderen van het leerproces van alle leerlingen (Ysenbaert et al., 2018). Uit het jaarverslag van de Onderwijsinspectie (2020) blijkt ook dat het onderwijskundig beleid van Vlaamse scholen nog veel ruimte laat voor verbetering. Een onderwijskundig beleid dat uitgaat van een inclusieve visie op onderwijs vertaalt zich in evaluatiepraktijken die leiden tot gedifferentieerde onderwijspraktijken (Struyven et al., 2019; C. A. Tomlinson, 2001). Struyf et al. (2018) besluiten in hun onderzoek naar de implementatie van het M-decreet dat een *'whole-school approach'* nodig is om te evolueren richting inclusief onderwijs. Alle aspecten binnen de school dragen zo bij tot het bevorderen van éénieders leeruitkomsten (Struyf et al., 2018). Zo krijgt de ondersteunende functie van evaluatie bij voorkeur een meer centrale rol in de visie die de basis vormt van het evaluatiebeleid. Zoals Vandenberghe (2005) echter stelt, volstaat een sterk formeel beleid of visie niet indien dit los staat van de pedagogisch-didactische praktijken op school. De *good practices* uit de casestudies verschillen van de overige cases net omdat er een sterk verband is tussen de visie enerzijds en de didactische praktijk anderzijds. Een visie waarbij evalueren het leren bevordert krijgt zo vorm via frequente observatie, peerevaluatie en klassenraden die expliciet aandacht besteden aan de sterktes van leerlingen zodat men bij de oriëntering kan focussen op het maximaal ontplooiën van de talenten van leerlingen.

Een didactische vertaling van een visie die evaluatie als essentieel ziet voor het bevorderen van het leerproces, volstaat niet om tot een krachtig evaluatiebeleid te komen. Devos, Vandenberghe, and Verhoeven (1989) verwijzen naar het belang van een zekere mate van geïntegreerd beleid, waarmee verwezen wordt naar de koppeling en coherentie tussen verschillende beleidsdomeinen van de school. Een algemene, overkoepelende visie die aandacht heeft voor het hele schoolgebeuren voorkomt een gefragmenteerd beleid en bevordert de ontwikkeling van een geïntegreerd beleid (Devos et al., 1989). Het nastreven van maximale leer- en ontwikkelingskansen dient dus niet alleen het leidend principe te zijn binnen het evaluatiebeleid, maar dient ook doorgetrokken te worden naar andere beleidsdomeinen van de school waarop het evaluatiebeleid inhaakt. De casestudies lieten zien hoe de visie op evaluatie soms doorgetrokken wordt in het communicatiebeleid waarin men de ontwikkeling van leerlingen in kaart brengt in een competentierapport. Of er blijkt een link met het beleid inzake ouderbetrokkenheid waarbij ouders heel vroeg betrokken worden indien men vermoedt dat een leerling het getuigschrift niet kan halen. Verder kan gedifferentieerde evaluatie vorm krijgen vanuit het zorgbeleid op school. Uit het casestudieonderzoek blijkt dat het evaluatiebeleid van de school op een te beperkt aantal beleidsdomeinen inhaakt. Ze maken veeleer deel uit van een gefragmenteerd of, zoals Devos et al. (1989) het zouden noemen, een 'gesegmenteerd' beleid. Een krachtig evaluatiebeleid van een school haakt idealiter in op meerdere deeldomeinen van het ruimere schoolbeleid waarbij het maximaliseren van leer- en ontwikkelingskansen bij elke leerling het uitgangspunt vormt.

## 4.2. Gebrek aan reflecterend vermogen

De afwezigheid van een krachtig en geïntegreerd evaluatiebeleid houdt risico's in voor het garanderen van maximale leer- en ontwikkelingskansen voor alle leerlingen. Hiermee lijkt het Vlaamse onderwijs zich in het tweede kwadrant van onderstaand model van de Nederlandse Onderwijsraad (2001) te situeren. Hoewel scholen over een grote formele beleidsruimte beschikken (Penninckx et al., 2011), is het beleidsvoerend vermogen onvoldoende om deze ook te benutten.

	<b>formele ruimte klein ( - )</b>	<b>formele ruimte groot ( + )</b>
<b>beleidsvoerend vermogen klein ( - )</b>	<p><i>zowel een geringe formele ruimte ( - ) als een gering beleidsvoerend vermogen ( - ) :</i></p> <p>Situatie levert als zodanig geen spanning op, maar is niet het globale streefbeeld. Regulering biedt garantie voor kwaliteit en toegankelijkheid van het onderwijs.</p>	<p><i>formele ruimte groter( + ) dan het beleidsvoerend vermogen ( - ) :</i></p> <p>Situatie levert spanning op. Scholen kunnen de ruimte niet aan, er zijn risico's voor de kwaliteit en toegankelijkheid van het onderwijs.</p>
<b>beleidsvoerend vermogen groot ( + )</b>	<p><i>formele ruimte kleiner ( - ) dan het beleidsvoerend vermogen ( + ) :</i></p> <p>Situatie levert spanning op. Scholen kunnen meer aan dan dat er ruimte is (frustratie voor de scholen) of doen meer dan dat er formeel kan (risico's voor de kwaliteit en toegankelijkheid van het onderwijs).</p>	<p><i>zowel veel formele ruimte ( + ) als een ontwikkeld beleidsvoerend vermogen ( + ) :</i></p> <p>Situatie levert als zodanig geen spanning op en is het globale streefbeeld. Beleidsvoerend vermogen (binnen kaderstellende regulering) biedt garantie voor de kwaliteit en toegankelijkheid van het onderwijs.</p>

Fig. 4.1. Relatie tussen de formele beleidsruimte (autonomie) en het beleidsvoerend vermogen van scholen (Onderwijsraad, 2001: p.23)

Eén van de indicatoren voor het beleidsvoerend vermogen van scholen is volgens Vanhoof et al. (2011) het reflecterend vermogen van de school, ofwel de mate waarin een school in staat is om haar eigen werking in vraag te stellen en bij te sturen. Dit kunnen we ook doortrekken naar het evaluatiebeleid en de evaluatiepraktijken van een school. De informerende functie van evaluatie beperkt zich niet tot informatie over het leerproces van de leerling, maar levert ook informatie op over de onderwijskwaliteit. De resultaten van leerlingen kunnen aanzetten tot reflectie bij de leerkracht om didactische keuzes te herzien. Op deze manier vormt evaluatie een instrument voor zelfevaluatie met betrekking tot de eigen onderwijspraktijk. Deze zelfevaluatie kan ook op het niveau van de school, bijvoorbeeld aan de hand van gestandaardiseerde toetsen. De resultaten vertellen iets over de mate waarin de school erin slaagt om bepaalde doelen te bereiken met de leerlingen, leerwinst in kaart te brengen of om na te gaan in welke mate de leerlingen op school het doen in vergelijking met een vergelijkbare referentiegroep. Dergelijke gestandaardiseerde toetsen zijn niet alleen een instrument voor zelfevaluatie, maar ze vormen eveneens een tool voor de interne kwaliteitsontwikkeling van de school. Afhankelijk van de mate waarin dergelijke informatie meegenomen wordt in de beleidskeuzes van de school hebben ze potentieel om de onderwijskwaliteit te bevorderen. Op basis hiervan kan men prioriteiten stellen en nagaan hoe de onderwijskwaliteit verbeterd kan worden. Zo kan men besluiten om bepaalde methodes niet langer te gebruiken, het onderwijs anders in te richten, met het team een professionaliseringstraject te volgen, etc. In de meerderheid van de scholen in het casestudieonderzoek (Ysenbaert et al., 2018) blijkt er hieromtrent echter geen cultuur te zijn. Dit sluit aan bij de bevindingen van de Onderwijsinspectie bij de doorlichtingen van vorig schooljaar: de kwaliteitsontwikkeling op schoolniveau is ontoereikend. Vooral in het secundair onderwijs slaagt minder dan de helft van de scholen erin om de beschikbare data te analyseren en aan te wenden in functie van kwaliteitsontwikkeling (Onderwijsinspectie, 2020). Opvallend is dat veel scholen uit het casestudieonderzoek wel gestandaardiseerde toetsen afnemen maar ze worden nauwelijks gebruikt voor de interne kwaliteitszorg. Integendeel, de meerderheid van de scholen lijkt zich te beperken tot het afnemen van wat hieromtrent door de overheid verplicht wordt. De gevalideerde toetsen op bijvoorbeeld het einde van het zesde leerjaar worden afgenomen zonder hier verder gevolg aan te geven. Het SONO-onderzoek van collega's Vanbuel, Vandommele, and Van den Branden (2020) dat het taalbeleid van scholen onder de loep neemt, stelt hetzelfde vast: hoewel scholen de verplichte taalscreening afnemen, worden de resultaten nauwelijks geanalyseerd en/of meegenomen in het bijsturen van het taalbeleid op school. De afwezigheid van een cultuur waarbij beschikbare data aangewend wordt in het kader van kwaliteitsontwikkeling wijst opnieuw naar een algemeen beperkt beleidsvoerend vermogen van scholen.

Tegelijkertijd wijzen de bevindingen van het casestudieonderzoek over gestandaardiseerd toetsen op een bijkomend risico. Meer specifiek werd vastgesteld dat leerkrachten er een overwegend negatieve houding op nahouden als het gaat over gestandaardiseerde toetsen. Een houding die nog sterker tot uiting kwam bij leerkrachten in scholen met een grotere diversiteit onder leerlingen. De overtuiging leeft dat dergelijke toetsen haaks staan op hun alledaagse onderwijs- en evaluatiepraktijk die gekenmerkt wordt door differentiatie. Deze vaststelling illustreert hoe het werkveld de beleidsplannen die het gebruik van gestandaardiseerde toetsen wil stimuleren als incoherent dreigt te beschouwen met eerder gemaakte beleidskeuzes die inclusief onderwijs nastreven. Anderzijds leverden de gesprekken met leerkrachten tijdens het casestudieonderzoek ook positieve houdingen op, enkel bij die leerkrachten die tewerkgesteld zijn in een school die gekenmerkt wordt door een leerlingenpubliek met hoge socio-economische status. Zij zien in deze toetsen een soort van 'stok-achter-de-deur-

functie' bij twijfel over studiekeuze. Op deze manier dreigen dergelijke toetsen veeleer aangewend te worden met het oog op de selecterende functie in plaats van de informerende waarde voor de kwaliteitsontwikkeling.

Deze bevindingen wijzen op een spanning die betrekking heeft op wat Hopkins (2003) een *systembrede* en *systemdiepe* aanpak noemt. Volgens Hopkins (2003) zijn beide aspecten van cruciaal belang voor het welslagen van beleidsinitiatieven. Het *systembrede* aspect onderstreept het belang van sterke coherentie tussen verschillende beleidsinitiatieven. Het *systemdiepe* aspect verwijst naar de vereiste dat de actoren op verschillende niveaus (beleidsmakers, schoolleiders, leerkrachten, ...) handelen vanuit dezelfde uitgangspunten. Het feit dat leerkrachten gestandaardiseerde toetsen percipiëren als een praktijk die lijnrecht staat tegenover hun algemene differentiatiepraktijk wijst op een hiaat inzake het *systembrede* aspect. Het beperkt beleidsvoerend vermogen van scholen lijkt daarenboven een hindernis te vormen bij de realisatie van een *systemdiepe* aanpak.

Samengevat stelt dit onderzoek de nood scherp aan centrale beleidsvoering die specifieke aandacht heeft voor twee zaken. Enerzijds aandacht om toekomstige beleidskeuzes expliciet te kaderen binnen het maximaliseren van leer- en ontwikkelingskansen van alle leerlingen. Anderzijds aandacht voor het stimuleren van scholen hun ontwikkeling naar een sterk beleidsvoerend vermogen dat hen in staat stelt om een krachtig evaluatiebeleid uit te werken dat eveneens maximale leer- en ontwikkelingskansen bij alle leerlingen nastreeft.

## V Conclusie & beleidsaanbevelingen

Het evaluatiebeleid van scholen en de evaluatiepraktijken van leerkrachten spelen een cruciale rol in het realiseren van kwaliteitsvol onderwijs voor alle leerlingen. In het licht van de (inter)nationale ontwikkelingen richting inclusief onderwijs en de toenemende diversiteit is specifieke aandacht voor evaluatiebeleid en -praktijken cruciaal. De onderzoeksbevindingen laten zien hoe het evaluatiebeleid en de evaluatiepraktijken in zowel basis- als secundair onderwijs in Vlaanderen kansen onbenut en/of onderbenut laat met betrekking tot het maximaliseren van leer- en ontwikkelingskansen van alle leerlingen.

Tegelijkertijd blijkt dat deze onderbenutte of gemiste kansen enerzijds het gevolg zijn van een niet-bewustzijn van en/of een gebrekkige aandacht voor het feit dat evaluatie uitermate essentieel is in het realiseren van maximale leer- en ontwikkelingskansen voor alle leerlingen. Anderzijds stelt het onderzoek vast dat evaluatie al te sterk een zaak blijkt te zijn van de individuele leerkracht. Dergelijke cultuur staat een gedragen en geïntegreerd evaluatiebeleid dat maximale leer- en ontwikkelingskansen voor alle leerlingen nastreeft in de weg.

In wat volgt, worden beleidsaanbevelingen geformuleerd die bijdragen tot de realisatie van een evaluatiebeleid dat de leer- en ontwikkelingskansen van alle leerlingen maximaliseert. In overeenstemming met de literatuur, die stelt dat evaluatiepraktijken het resultaat zijn van een complexe wisselwerking tussen factoren op verschillende niveaus (Ercikan & Solano-Flores, 2016), voelen we ons genoodzaakt om van start te gaan met één algemene, overkoepelende aanbeveling. Deze richt zich tot alle niveaus en betrokken actoren die (on)rechtstreeks invloed uitoefenen op het evaluatiebeleid van scholen en de evaluatiepraktijken van leerkrachten. Deze aanbeveling vormt als het ware het basisprincipe van waaruit de verschillende actoren best handelen bij het opzetten en ontwikkelen van initiatieven die gerelateerd zijn aan evaluatie.

Daarnaast formuleren we specifieke aanbevelingen ten aanzien van de overheid, de pedagogische begeleidingsdiensten, de lerarenopleidingen, de schoolleiders en de leerkrachten. We willen benadrukken dat de aanbevelingen voor leerkrachten bewust als laatste in de rij werden opgenomen. In navolging van de onderzoeksbevindingen willen we benadrukken dat de grootste uitdagingen zich situeren op de overige beleidsniveaus. De oplossing om tegemoet te komen aan het probleem dat evaluatie nu te veel een zaak van de individuele leerkracht is, zal dan ook vanuit die andere beleidsniveaus geïnitieerd moeten worden.

### **0. Algemene aanbeveling: Stel maximale leer- en ontwikkelingskansen centraal bij elk initiatief dat aan evaluatie raakt.**

Het staat buiten kijf dat een diverse leerlingenpopulatie vandaag de dag realiteit is in Vlaamse scholen en klassen. Dit uit zich bovendien in doelstellingen van internationale samenwerkingsverbanden, zoals bijvoorbeeld het Verdrag van de Verenigde Naties (2006) dat inclusief onderwijs als doel voorop stelt. Dit verdrag werd door België geratificeerd in 2009 wat betekent dat België sindsdien moet kunnen aantonen op welke wijze werk wordt gemaakt van het realiseren van zo kwaliteitsvol mogelijk onderwijs voor alle leerlingen. De meest recente definiëring van inclusief onderwijs (Operti et al.,

2013) legt bovendien de nadruk op een systeembenadering. Deze benadering doelt op het zodanig inrichten van het onderwijssysteem zodat álle leerlingen bereikt worden, ongeacht hun specifieke achtergrond of kenmerken. De idee van een gecompartmenteerde doelgroepenbenadering waarbij specifieke maatregelen nodig zijn voor specifieke (groepen van) leerlingen wordt hier opgeheven doordat het systeem gelijke kansen op kwaliteitsvol onderwijs creëert voor álle leerlingen, ook voor deze leerlingen die momenteel niet 'gelabeld' worden als leerlingen met specifieke onderwijsbehoeften.

Leerlingenevaluatie biedt heel wat mogelijkheden om ééniéders leer- en ontwikkelingskansen te maximaliseren. De resultaten van dit onderzoek tonen helaas dat het huidige evaluatiebeleid en de evaluatiepraktijken de kansen op kwaliteitsvol onderwijs voor alle leerlingen eerder dreigt te belemmeren dan te bevorderen. We besluiten dat het gebruikelijke concept van evaluatie én de uitwerking ervan in de praktijk grondige bijsturingen vereist. We schuiven daarom volgend principe naar voren: evaluatie moet in de eerste plaats het leren en ontwikkelen van álle leerlingen dienen. Dit uitgangspunt maximaliseert de leer- en ontwikkelingskansen van elke leerling en niet uitsluitend van 'risico-leerlingen' of leerlingen met specifieke onderwijsbehoeften. Het is cruciaal dat elk van de drie hieronder beschreven functies van evaluatie vanuit dit idee vorm en invulling krijgen. Zoniet, dreigt het streven om via evaluatie kansen te bevorderen alsnog te resulteren in het limiteren van kansen. Zo verdwijnen bijvoorbeeld dagdagelijkse brede vormen van evaluatie in het niets als uiteindelijk de studiekeuze bepaald wordt door het resultaat op één examen van wiskunde.

We lichten enkele aandachtspunten bij elk van deze functies toe.

### **1. Informerende functie**

- Evaluatie moet erop gericht zijn om het profiel van de leerling zo breed mogelijk in kaart te brengen. Specifiek houdt dit in dat er naast de cognitieve aspecten, ook aandacht gaat naar affectieve, motorische, motivationele en metacognitieve aspecten.
- Het brede profiel van de leerling besteedt niet alleen aandacht aan de actuele stand van zaken in het leerproces en de ontwikkeling, maar biedt ook informatie over de gerealiseerde groei in relatie tot vooropgestelde doelen.
- De informatie wordt verkregen op basis van een brede waaier aan evaluatie-instrumenten en evaluatie-vaardigheden. We stellen een weloverwogen mix voor van observaties, leerlingcontacten, zelf- en peerevaluatie, opdrachten, interacties, vraaggesprekken, toetsen, examens en gestandaardiseerde evaluatievormen. Binnen deze ruime waaier wordt best gewicht toegekend aan elke evaluatievorm en -tool.

### **2. Ondersteunende functie**

- Evalueren en leren zijn onlosmakelijk met elkaar verboden. Idealiter maakt deze functie bijgevolg integraal deel uit van de dagdagelijkse didactiek door evaluatieactiviteiten sterk in te bedden in de algemene onderwijspraktijk.
- De output van evaluatie wordt maximaal aangewend als input waarmee de onderwijspraktijk afgestemd wordt op de leer- en ontwikkelingsnoden van leerlingen (= output=input-principe). Dit resulteert in een onderwijspraktijk waarbij evaluatie het beginpunt vormt van de didactische cyclus.
- Het output=input-principe wordt nog krachtiger indien dit zich ook vertaalt in concrete handvatten voor de leerling zelf. Deze kunnen aangereikt worden via kwaliteitsvolle feedback die hen helpt om de vooropgestelde doelen te bereiken.

### **3. Beslissingnemende functie**

- Beslissingen over de schoolloopbaan van leerlingen moeten genomen worden op basis van *faire/billijke* evaluaties die aan de beslissingen voorafgaan.
- Het nemen van beslissingen moet beschouwd worden als een proces in plaats van als een moment.
- Het nemen van beslissingen over de schoolloopbaan is een gedeelde verantwoordelijkheid en daarom verloopt dit proces best in partnerschap met collega-leerkrachten, ouders én de leerlingen zelf.

De aanbevelingen die hierna weergegeven worden, richten zich naar specifieke actoren. De aanbevelingen hebben als doel om het huidige evaluatiebeleid en de evaluatiepraktijken zo bij te sturen zodat ze bijdragen aan de realisatie van de hierboven omschreven algemene aanbeveling. Belangrijk is dat elke actor hierin actief initiatief neemt opdat er op coherente wijze gewerkt kan worden aan een evaluatiebeleid dat de ontwikkeling naar een inclusiever onderwijs bevordert.

## **1. Aanbevelingen ten aanzien van de overheid**

Onderwijsvernieuwingen maken volgens Hopkins (2003) pas kans op slagen indien beleidsinitiatieven aandacht hebben voor zowel een systeem-brede als een systeem-diepe aanpak. De systeem-brede aanpak verwijst naar het voeren van een beleid dat coherent is en ook als coherent gepercipieerd wordt door de betrokken actoren. We adviseren om de plannen in de beleidsnota 2019-2024 om gevalideerde, gestandaardiseerde en genormeerde toetsen te introduceren best te kaderen binnen eerder genomen beleidsinitiatieven die de ontwikkeling naar meer inclusief onderwijs stimuleren. Hierbij is het van belang om de actuele invulling van 'inclusie' voor ogen te houden, namelijk de brede benadering waarbij het onderwijssysteem zo ingericht wordt opdat het voor alle leerlingen mogelijk wordt om hun potentieel ten volle te kunnen ontplooien en ontwikkelen, met oog op het minimaal bereiken van de eindtermen. Daarnaast heeft de overheid ook een rol om haar beleid vertaald te krijgen in de praktijk. Dit kan bevorderd worden door aandacht te besteden aan een systeem-diepe aanpak. Voor zowel de systeem-brede als de systeem-diepe aanpak formuleren we enkele concrete aanbevelingen.

### **Aanbeveling 1.1. Een systeem-brede aanpak realiseren via het bedachtzaam introduceren van gestandaardiseerde toetsen**

Hoewel gestandaardiseerde toetsen mogelijkheden bieden in het streven naar kwaliteitsvol onderwijs, pleiten we voor een bedachtzame introductie ervan. De resultaten van dit onderzoek, en meer specifiek de opvattingen van leerkrachten, lieten zien dat dergelijke toetsen soms als selectie-instrument dreigen aangewend te worden. De implementatie van gevalideerde, gestandaardiseerde en genormeerde toetsen wordt daarom best niet uitsluitend gekaderd binnen de internationale trend richting meer centrale toetsing om de onderwijskwaliteit op te volgen. We adviseren om bij deze implementatie eveneens te waken over het realiseren van de ratificatie van het Verdrag van de Verenigde Naties waartoe Vlaanderen zich sinds 2009 verbindt. Dit bevordert de systeem-brede aanpak – ook in de perceptie van de overige actoren in het onderwijsveld - aangezien dit aansluit bij bijvoorbeeld het M-decreet. Het streefdoel van inclusief onderwijs dat maximale leer- en ontwikkelingskansen bij alle leerlingen beoogt, dient bijgevolg het uitgangspunt te vormen bij de



implementatie van gevalideerde, gestandaardiseerde en genormeerde toetsen. De overheid wordt geadviseerd om hierbij concreet rekening te houden met volgende aandachtspunten:

- De implementatie van gevalideerde, gestandaardiseerde en genormeerde toetsen moet eerder gericht zijn op ontwikkeling van de school, de leerkracht en de leerling dan op verantwoording. We verwijzen hiervoor naar de *good practice* van case D in het basisonderwijs waarbij er op schoolniveau regelmatig gestandaardiseerde taaltoetsen afgenomen worden. De resultaten van deze toetsen worden uitsluitend aangewend voor het opvolgen en verbeteren van het taalonderwijs op school en het nemen van specifieke beleidskeuzes om de taalvaardigheid van leerlingen te bevorderen. Het gaat om gestandaardiseerde toetsen die dus vooral een reflexieve functie hebben. Het leidde concreet tot beslissingen op vlak van:
  - o Onderwijskundig beleid: taal werd niet langer als apart vak beschouwd, maar werd geïntegreerd binnen wereldoriëntatie.
  - o Professionaliseringsbeleid: leerkrachten lager onderwijs werden in een traject begeleid in hoe ze taalvaardigheid van leerlingen konden stimuleren en evalueren. Later werden ook leerkrachten van het kleuteronderwijs hierin geprofessionaliseerd.
  - o Beleid inzake kwaliteitsontwikkeling: de gestandaardiseerde toetsen worden regelmatig afgenomen en de evolutie van de resultaten wordt in verband gebracht met genomen acties van de school om na te gaan in welke mate deze de taalvaardigheid bevorderen/belemmeren.
- Verantwoording met betrekking tot de output op deze toetsen is uitermate onwenselijk. Mogelijke verantwoording dient zich te richten op de kwaliteitsontwikkeling van de school als organisatie. Anders gezegd: door scholen verantwoordelijk te houden voor de mate waarin men op basis van de resultaten op deze toetsen geïnformeerde beleidskeuzes maakt, worden scholen gestimuleerd om te ontwikkelen naar kwaliteitsvoller onderwijs ten aanzien van alle leerlingen.
- Een breed instrumentarium aan gevalideerde, gestandaardiseerde en genormeerde toetsen dat recht doet aan de ruime waaier aan na te streven leer- en ontwikkelingsdoelen is nodig.
- Scholen moeten hun beleid inzake kwaliteitsontwikkeling kunnen afstemmen op de contextspecifieke noden. Hiervoor moeten ze over een zekere mate van keuzevrijheid kunnen beschikken inzake de zelfevaluatie-instrumenten die ze hiervoor aanwenden.

### **Aanbeveling 1.2. Een systeem-brede aanpak realiseren door de autonomie van scholen te beklemtonen in het realiseren van een evaluatiebeleid dat maximale leer- en ontwikkelingskansen bij alle leerlingen nastreeft.**

Om op schoolniveau een evaluatiebeleid te ontwikkelen dat kan vertrekken vanuit de eigenheid en diverse compositie van de school en dat maximaal de kansen van alle leerlingen bevordert, vormt de autonomie van de school een essentieel aspect. Aangezien leren en evalueren onlosmakelijk met elkaar verbonden zijn, moet een school de ruimte waarover ze beschikt om het evaluatiebeleid af te stemmen op het onderwijskundig beleid ook benutten. Een krachtig evaluatiebeleid sluit aan bij de ambitie om het leren van alle leerlingen maximaal te bevorderen. Daarenboven moeten scholen verantwoordelijk gehouden worden om hun autonomie zinvol aan te wenden om gericht in te spelen op de specifieke noden van de context van de school. Op die manier kan bijvoorbeeld een school met een groot aandeel leerlingen wiens thuistaal niet Nederlands is, opteren voor een frequent gebruik van verschillende evaluatievormen die de taalvaardigheid van leerlingen in kaart brengen om zo de

evoluties te monitoren en de taaldidactiek hierop af te stemmen. De resultaten van het onderzoek lieten echter zien dat scholen over het algemeen heel wat kansen laten liggen om op basis van de ruime autonomie waarover men beschikt een schoolspecifiek evaluatiebeleid te ontwikkelen. We pleiten allerminst voor het vergroten van deze autonomie. We adviseren eerder om scholen verantwoordelijk te houden met betrekking tot de invulling van deze autonomie waarbij de nadruk ligt op het ontwikkelen van een evaluatiebeleid dat maximale leer- en ontwikkelingskansen nastreeft. Het raamwerk voor onderwijskwaliteit (Onderwijsinspectie, 2016) dat gehanteerd wordt bij schooldoorlichtingen legt hieromtrent de juiste accenten. Hoewel het casestudieonderzoek liet zien dat een doorlichting een stimulans vormt voor het invullen van deze autonomie, stellen we vast dat dit in de praktijk nog geen gemeengoed is.

### **Aanbeveling 1.3. Een systeem-diepe aanpak realiseren door professionalisering te voorzien**

We adviseren de overheid om sterk in te zetten op een *styeemdiepe* aanpak die ernaar streeft om het handelen van actoren op verschillende niveaus te stroomlijnen volgens dezelfde uitgangspunten. Een heldere communicatie vanuit de overheid waarbij de *styeembrede* aanpak geëxpliciteerd wordt, draagt hiertoe bij.

Verder is het van belang om professionalisering voor scholen mogelijk te maken opdat ze ondersteund worden in het ontwikkelen van een evaluatiebeleid dat maximale leer- en ontwikkelingskansen bij alle leerlingen nastreeft. De resultaten van het onderzoek laten namelijk een positieve samenhang zien tussen het deelnemen aan professionalisering en de mate waarin ingezet wordt op een geïntegreerde evaluatiepraktijk. Binnen dergelijke professionalisering is er naast specifieke aandacht voor de ontwikkeling en implementatie van het formele evaluatiebeleid en de evaluatiepraktijken ook aandacht voor de opvattingen omtrent evaluatie van betrokkenen.

De overheid wordt aanbevolen om hierbij professionaliseringsinitiatieven te voorzien die expliciet inzetten op het betrekken van verschillende actoren op school bij de ontwikkeling van dergelijk evaluatiebeleid. Steeds meer onderzoek toont namelijk aan dat gedeeld leiderschap bijdraagt tot organisatieperformantie en een positief effect heeft op leerlingresultaten (Harris & Spillane, 2008). Bovendien versterkt dit de ontwikkeling naar de school als een lerende organisatie die haar werking in vraag stelt en zo gericht bijstuurt. Hiermee kan tegemoet gekomen worden aan het vastgestelde gebrek aan reflecterend vermogen.

Bovendien moet dergelijke professionalisering ook het beleidsvoerend vermogen van scholen versterken door te streven naar een geïntegreerd beleid. Op deze manier wordt een gefragmenteerd of gesegmenteerd beleid, zoals in het onderzoek werd vastgesteld, voorkomen. De overheid wordt aanbevolen om professionaliseringsinitiatieven te voorzien die scholen ondersteunen in het expliciteren van verbanden tussen het evaluatiebeleid en andere aspecten van het schoolbeleid zoals bijvoorbeeld het taalbeleid, zorgbeleid of het beleid inzake de interne kwaliteitsontwikkeling.

## **2. Aanbevelingen ten aanzien van pedagogische begeleidingsdiensten**

### **Aanbeveling 2.1. Een pro-actief professionaliseringsaanbod voorzien dat scholen ondersteunt in de ontwikkeling van een evaluatiebeleid dat maximale leer- en ontwikkelingskansen vooropstelt.**

De resultaten van zowel het casestudie- als het surveyonderzoek wijzen op een positieve samenhang tussen scholen die inzetten op professionalisering en sterk geïntegreerde evaluatiepraktijken. Deze professionalisering heeft specifieke aandacht voor zowel het formele evaluatiebeleid, de evaluatiepraktijken zelf en de opvattingen omtrent evaluatie van betrokkenen. De professionalisering beoogt hiermee het evaluatiebeleid en de evaluatiepraktijken zo te ontwikkelen dat elk van de drie functies van evaluatie maximale leer- en ontwikkelingskansen bij alle leerlingen nastreeft.

Begeleidingsdiensten en nascholingsinitiatieven wordt aanbevolen om hierbij specifiek in te zetten op gedeeld leiderschap. Het casestudieonderzoek laat zien dat het evaluatiebeleid op een duurzamere wijze vertaald wordt in de praktijk indien meerdere actoren voor langere tijd betrokken worden bij de ontwikkeling hiervan. Het betrekken van verschillende actoren op school bij de ontwikkeling van dergelijk evaluatiebeleid draagt niet alleen bij tot organisatieperformantie maar vertaalt zich ook in een positief effect op leerlingenresultaten (Harris & Spillane, 2008).

### **Aanbeveling 2.2. Een pro-actief begeleidingsaanbod voorzien dat scholen specifiek ondersteunt in de ontwikkeling van een geïntegreerd evaluatiebeleid.**

Een evaluatiebeleid ontwikkelen is uitermate complex, precies omdat het op veel verschillende aspecten van het ruimere schoolbeleid betrekking heeft. Hoewel het referentiekader voor onderwijskwaliteit (Onderwijsinspectie, 2016) deze complexiteit weergeeft, zijn scholen zich niet altijd bewust van de afwezige of gebrekkige verbanden in het eigen evaluatiebeleid. De resultaten van dit onderzoek lieten zien hoe een krachtig evaluatiebeleid vaak afwezig is of sterk gefragmenteerd is. Pedagogisch begeleiders kunnen scholen pro-actief begeleiden om inzicht te verwerven in het eigen evaluatiebeleid om zo de leemtes te ontdekken en deze procesmatig aan te pakken. Zo kan er concreet gefocust worden op het verbinden van een evaluatiebeleid dat streeft naar het maximaliseren van leer- en ontwikkelingskansen en het zorgbeleid van de school, of het verband met het beleid inzake ouderbetrokkenheid, het professionaliseringsbeleid, het taalbeleid en of het beleid omtrent kwaliteitszorg.

### **Aanbeveling 2.3. Scholen pro-actief begeleiden in de ontwikkeling van een reflexieve cultuur onder andere door middel van gestandaardiseerde vormen van evaluatie.**

Een specifiek aspect van evaluatie omvat het aanwenden van evaluatie om te reflecteren over het beleid en de praktijken van de school, een cultuur die nog verre van gangbaar is in het Vlaamse onderwijs. De resultaten van het onderzoek laten zien hoe verplichte gestandaardiseerde vormen van evaluatie nauwelijks voor reflexieve doeleinden worden gebruikt. Met het vooruitzicht op het introduceren van netoverstijgende, gestandaardiseerde toetsen, wordt het des te belangrijker dat scholen de juiste ondersteuning aangereikt krijgen om dergelijke evaluatieresultaten onder de loep te nemen. Scholen zullen begeleid moeten worden in het ontwikkelen van expertise die hen in staat stelt om gestandaardiseerde vormen van evaluatie aan te wenden voor het monitoren van:

- de ontwikkeling van alle leerlingen met het oog op het minimaal bereiken van de eindtermen,
- de professionele ontwikkeling van het team,
- de ontwikkeling van de school als organisatie.

Het zwaartepunt van dergelijke ondersteuning ligt niet noodzakelijk op het monitoren op zich, maar op het ontwikkelen van een reflexieve cultuur. Dit stelt een school in staat om geïnformeerde beleidskeuzes te maken waaraan doelgerichte acties gekoppeld worden. Dergelijke cultuur is niet alleen noodzakelijk om het leren van leerlingen naar een hoger niveau te tillen, maar ook om als school te evolueren richting een lerende organisatie die zichzelf voortdurend bijstuurt.

### 3. Aanbevelingen ten aanzien van lerarenopleidingen

#### Aanbeveling 3.1. De kaart van diversiteit trekken

De onderzoeksbevindingen lieten zien dat de evaluatiepraktijk gekenmerkt wordt door een onderbenutting van de formatieve functie die evaluatie heeft. Daarenboven bleek ook dat de summatieve functie sterk aangestuurd wordt door het waterval-denken. Beide vaststellingen samen vormen een hindernis voor het creëren van maximale leer- en ontwikkelingskansen voor alle leerlingen. Meer zelfs: ze staan een evolutie richting meer inclusief onderwijs in de weg. In navolging van de evolutie naar meer inclusief onderwijs, wordt lerarenopleidingen daarom aanbevolen om de competenties inzake evaluatie aan te brengen via het trekken van de kaart van diversiteit.

Bij de competentieontwikkeling van student-leraren inzake evaluatie dient erover gewaakt te worden dat bij één van de drie functies van evaluatie de maximale leer- en ontwikkelingskansen van alle leerlingen het streefdoel zijn. Dit vertaalt zich idealiter in een accentverschuiving in het curriculum waarbij uitgegaan wordt van de idee dat evaluatie het beginpunt vormt in het didactisch proces en niet het eindpunt. Het aspect 'evaluatie' vroeg introduceren in het curriculum kan hier al toe bijdragen. Aangezien evalueren en leren onlosmakelijk met elkaar verbonden zijn, is het niet meer dan logisch dat evaluatie van bij het begin meegenomen wordt in de didactische leerlijn van leraren-in-opleiding. We pleiten voor een curriculum waarin evaluatie geen op zichzelf staand thema vormt, maar steeds terugkeert doorheen het curriculum waarbij telkens de verschillende functies van evaluatie geëxpliciteerd worden. Een curriculum van de lerarenopleiding waarbij evaluatie voortdurend aan de orde is, sluit bovendien aan bij het referentiekader voor onderwijskwaliteit (Onderwijsinspectie, 2016) waarbij evaluatie eveneens opduikt in meerdere kwaliteitsrubrieken.

Een dergelijke benadering waarbij de rol van evaluatie als dé ingang tot differentiatiepraktijken reeds vroeg beklemtoond wordt, zal student-leraren vaardiger maken in het ontwikkelen van een 'diversiteits-bril'. De onderzoeksresultaten wezen op dit vlak op een positief verband tussen het kunnen opmerken van diversiteit enerzijds en het afstemmen van zowel instructie- als evaluatiepraktijken op deze diversiteit. Via de bekwaamheid om diversiteit onder leerlingen op te merken, voelen leerkrachten zich competentier om hun onderwijs- en evaluatiepraktijken af te stemmen op de diversiteit onder leerlingen en in te zetten op hun talentontwikkeling.

#### Aanbeveling 3.2. Sterker inzetten op praktijkgerichte oefenkansen inzake evaluatie

Het vroeg introduceren van 'evaluatie' in het curriculum en het veelvuldig laten terugkeren hiervan, zal niet volstaan om de evaluatiecompetentie van toekomstige leraren ten volle te ontwikkelen. Leerkrachten geven aan dat ze op dit vlak het meest geleerd hebben via de praktijkervaring die ze

hieromtrent opdeden. Lerarenopleidingen wordt aanbevolen om deze verhoogde aandacht ook door te trekken naar praktijkgerichte opdrachten en stage-onderdelen, precies om toekomstige leraren al zoveel mogelijk kansen te geven tot praktijkervaring. Dit zal studenten alleen maar vaardiger maken in hun competentie om hun evaluatiepraktijk zo in te richten opdat de leer- en ontwikkelingskansen van leerlingen gemaximaliseerd worden. Hierbij is een sterke focus nodig op het ontwikkelen van de vaardigheid om evaluatiepraktijken sterk in te bedden in de algemene instructieactiviteiten (Cfr. geïntegreerde evaluatiepraktijk). Daarnaast zijn oefenkansen nodig om de vaardigheden inzake de informerende, ondersteunende en de beslissingnemende functie van evaluatie aan te scherpen.

## **4. Aanbevelingen ten aanzien van schoolleiders**

### **Aanbeveling 4.1. Een evaluatiebeleid dat streeft naar het maximaliseren van ieders leer- en ontwikkelingskansen.**

Een evaluatiebeleid waarin diversiteit een centrale plaats heeft, maximaliseert de leer- en ontwikkelingskansen van elke leerling. Het is hierbij cruciaal dat elk van de drie functies van evaluatie vanuit dit idee vorm en invulling krijgen (zie algemene aanbeveling). Zoniet, dreigt er een risico op het limiteren van de ontwikkelingskansen van (groepen van) leerlingen. Dit onderzoek stelde namelijk vast dat de evaluatiepraktijk gekenmerkt wordt door een overwegend gescheiden evaluatiepraktijk, wat veeleer een drempel vormt in het creëren van maximale leer- en ontwikkelingskansen voor alle leerlingen.

Via het centrale doel om bij het evaluatiebeleid maximale leer- en ontwikkelingskansen voor elke leerling steeds voor ogen te houden beogen we twee verschuivingen in de huidige evaluatiepraktijken van scholen. Enerzijds is er nood aan een sterkere klemtoon op de formatieve functie van evaluatie in plaats van de summatieve functie. Uit het onderzoek blijkt dit een noodzakelijke verschuiving aangezien het onderbenutten van de formatieve functie eerder de leer- en ontwikkelingskansen belemmert in plaats van bevordert. Via het streven naar maximale leer- en ontwikkelingskansen wordt eveneens een verschuiving beoogt met betrekking tot de summatieve functie van evaluatie. Dit houdt in dat de oriëntering van leerlingen niet langer uitgaat van tekorten bij leerlingen, maar veeleer aangestuurd wordt vanuit de vraag welke de talenten zijn van leerlingen en waar deze het best tot uiting kunnen komen. Op deze manier behoedt het evaluatiebeleid zich voor een oriënteringspraktijk die gebaseerd is op het watervaldenken en waarbinnen evaluatie een selecterende functie heeft.

### **Aanbeveling 4.2. Een sterk geïntegreerd evaluatiebeleid nastreven vanuit een reflexieve cultuur**

Om tot een sterk evaluatiebeleid te komen is het van belang om vanuit het ruimere schoolbeleid expliciet stil te staan bij elk van de drie functies van evaluatie. Dit draagt bij tot een sterkere integratie van het evaluatiebeleid en voorkomt een gesegmenteerd evaluatiebeleid wat nu doorgaans te vaak het geval is volgens de resultaten van het onderzoek. De verschillende functies krijgen bij voorkeur een plaats binnen het evaluatiebeleid waarbij de onderlinge samenhang met de overige beleidsdomeinen bewaakt wordt.

Zo haakt een evaluatiebeleid logischerwijs in op het *onderwijsbeleid* van de school aangezien evalueren en leren onlosmakelijk met elkaar verbonden zijn. Zowel het onderwijs- als het evaluatiebeleid van een school zijn gelinkt aan het *zorgbeleid* net omdat via evaluatie de noden en leerbehoeften van leerlingen vastgesteld worden. Ook vanuit het *communicatiebeleid* van een school

kan het evaluatiebeleid verder ingevuld worden. Vragen zoals waarom en hoe men leerlingen, ouders en collega's wil informeren inzake de aanpak en de uitkomsten van evaluatie kunnen van hieruit beantwoord worden.

Verder is er ook een samenhang tussen het evaluatiebeleid en het *kwaliteitsbeleid* van de school. Dit kwaliteitsbeleid krijgt idealiter vorm via een reflexieve cultuur waarbij bijvoorbeeld gestandaardiseerde vormen van evaluatie aangewend worden om de ontwikkeling te bevorderen van leerlingen, leerkrachten en de school als organisatie. Zelfevaluatie op organisatieniveau biedt zo informatie die kan aangewend worden om aan kwaliteitsontwikkeling te doen. Tegelijkertijd haakt dit mogelijks in op het *professionaliseringsbeleid* van de school om zo kwaliteitsvol onderwijs aan alle leerlingen te realiseren. Een reflexieve cultuur zal op termijn niet alleen bijdragen tot een meer geïntegreerd evaluatiebeleid, het stimuleert tegelijkertijd de school om te evolueren richting een lerende organisatie.

Ten slotte willen we er op wijzen dat het referentiekader voor Onderwijskwaliteit (Onderwijsinspectie, 2016) hierbij een bruikbaar raamwerk vormt. Het feit dat evaluatie via verschillende kwaliteitsverwachtingen in elk van de vier rubrieken van het referentiekader aan bod komt, draagt bij tot de uitwerking van een evaluatiebeleid dat aandacht besteedt aan enerzijds de verschillende functies van evaluatie en anderzijds de expliciete verbanden met verschillende beleidsaspecten van het ruimere schoolbeleid.

### **Aanbeveling 4.3. Faciliterende randvoorwaarden scheppen om het evaluatiebeleid te vertalen naar een afdoende evaluatiepraktijk**

Enkele randvoorwaarden faciliteren de vertaalslag van een evaluatiebeleid naar een gedragen beleid dat ook in de praktijk vertaald wordt. Bij dit alles is er een cruciale rol weggelegd voor de directeur die hierin een katalysatorfunctie heeft, zo bleek ook uit het casestudieonderzoek.

Een eerste randvoorwaarde is om de ontwikkeling van het evaluatiebeleid te realiseren via gedeeld leiderschap. Steeds meer onderzoek toont namelijk aan dat gedeeld leiderschap bijdraagt tot organisatieperformantie wat tevens resulteert in betere leerlingresultaten (Harris & Spillane, 2008). Het casestudieonderzoek omvat *good practices* waarbij via een kernteam een gedragen evaluatiebeleid geïnitieerd wordt. Belangrijk hierbij is dat een visie en een beleid omtrent evaluatie vakoverschrijdend ontwikkeld wordt, om pas in een volgende stap de concrete uitwerking te vertalen binnen een specifiek vak.

Naast het ontwikkelen in gedeeld leiderschap, wijst dit onderzoek ook op de rol van externe professionalisering. Deze kan betrekking hebben op het verruimen van kennis inzake kwaliteitsvolle evaluatie in het algemeen, het procesmatig ontwikkelen van een evaluatiebeleid, of het aanscherpen van specifieke evaluatievaardigheden van leerkrachten.

Ten slotte blijkt het van belang te zijn om voldoende mogelijkheden te voorzien waarbij leerkrachten ook aan interne professionalisering kunnen doen. Bijvoorbeeld door randvoorwaarden te creëren die onderlinge samenwerking tussen leerkrachten stimuleren om zo hun evaluatiepraktijken in co-creatie te ontwikkelen. Door in te zetten op gedeelde en gezamenlijke professionalisering wordt het mogelijk om gaandeweg naar een gedeelde evaluatiepraktijk te evolueren die in wisselwerking staat met het evaluatiebeleid.

## 5. Aanbevelingen ten aanzien van leerkrachten

### Aanbeveling 5.1. Een evaluatiepraktijk die maximale leer- en ontwikkelingskansen nastreeft bij álle leerlingen.

Een evaluatiebeleid die geënt is op diversiteit maximaliseert de leer- en ontwikkelingskansen van élke leerling. Het is hierbij cruciaal dat elk van de drie functies van evaluatie vanuit dit idee vorm en invulling krijgen (zie algemene aanbeveling). Zoniet, dreigt er een risico op het limiteren van de ontwikkelingskansen van (groepen van) leerlingen. Inzake het realiseren van een evaluatiepraktijk die maximale leer- en ontwikkelingskansen nastreeft, schuiven we drie specifieke aandachtspunten naar voren.

#### 5.1.1. Zet maximaal in op een sterke integratie van evaluatie en instructie

De onderzoeksresultaten laten zien hoe leer- en ontwikkelingskansen beperkt worden door een overwegend gescheiden evaluatiepraktijk. De kracht van evaluatie zit echter in het feit dat evalueren en leren onlosmakelijk met elkaar verbonden zijn. Dit komt het beste tot uiting in een geïntegreerde evaluatiepraktijk waarbij evaluatie voortdurend deel uitmaakt van de algemene onderwijspraktijk. Uit het onderzoek blijkt bovendien dat er een positief verband is tussen de mate waarin leerkrachten in staat zijn om diversiteit tussen leerlingen op te merken en de mate waarin ze een geïntegreerde evaluatiepraktijk aanwenden. Via dergelijke evaluatiepraktijk kunnen de leer- en ontwikkelingsbehoeften van leerlingen beter opgevangen worden. Een sterkere integratie van evaluatie en instructie resulteert enerzijds in het optimaal afstemmen van de instructieactiviteiten op de leerbehoeften van leerlingen. Anderzijds draagt dit bij tot het bevorderen van het leerproces van leerlingen aangezien een geïntegreerde evaluatiepraktijk veel kansen creëert om leerlingen directe feedback te geven. Op deze manier vormt evaluatie de motor van een inclusieve leeromgeving.

#### 5.1.2. Betrek zoveel mogelijk de leerlingen en hun ouders bij het evaluatieproces

Zowel met betrekking tot de ondersteunende functie van evaluatie als bij de functie van het nemen van beslissingen, loont het om de leerlingen en hun ouders te betrekken bij het evaluatieproces. Leerlingen kunnen betrokken worden voorafgaand bij het evaluatieproces door met hen heldere, transparante afspraken te maken over wanneer en wat er geëvalueerd wordt, en hoe dit gebeurt. Tijdens de evaluatie kunnen leerlingen nauw betrokken worden door hen een rol als (mede-)evaluator te geven via bijvoorbeeld zelf- of peerevaluatie. Het ondersteunende effect van evaluatie kan gemaximaliseerd worden door leerlingen ook na een evaluatie-activiteit volop te betrekken, bijvoorbeeld door ruime aandacht te besteden aan een feedbackmoment dat volgt na evaluatie. Uit het onderzoek blijkt namelijk dat een hogere betrokkenheid van leerlingen positief samenhangt met een geïntegreerde evaluatiepraktijk die overigens meer rekening houdt met de diversiteit onder leerlingen.

Met betrekking tot het evalueren om beslissingen te nemen over de schoolloopbaan is het eveneens van belang dat de leerling en zijn/haar ouders vroeg betrokken worden: dit om de leerling inzicht te laten krijgen in de eigen mogelijkheden en talenten, maar eveneens om de leerling en zijn/haar ouders te informeren over welke mogelijkheden de onderwijsstructuur biedt. Uit het casestudieonderzoek blijkt bijvoorbeeld dat de oriëntering van leerlingen sterk onderhevig is aan de kenmerken van het onderwijssysteem en gestuurd wordt door het watervaldenken. Een procesmatige aanpak waarbij de verschillende partners betrokken worden stimuleert daarentegen een meer leerlinggerichte

oriëntering. Ook na het nemen van een beslissing blijft het van belang om samen met leerling en ouders de vinger aan de pols te houden om zo gemaakte keuzes eventueel te herzien.

### **5.1.3. Maak van je evaluatiepraktijk een gedeelde praktijk**

Leerkrachten wordt geadviseerd om hun evaluatiepraktijk in te richten via co-creatie. Door nauw samen te werken met collega-leerkrachten wordt het evaluatiebeleid van de school ook bottom-up ontwikkeld, zo bleek uit de *good practices* van het casestudieonderzoek. Bovendien draagt het samenwerken en het uitwisselen van praktijkervaringen bij tot het intern professionaliseren van leerkrachten.



# Bibliografie

- Anthoons, I., Deprez, N., Thyssen, S., & Verbruggen, M. (2004). Worden leerlingen in verschillende scholen op gelijke wijze gedelibereerd? Onderzoek naar de invloed van de deliberatiepraktijk op de doorstroming van leerlingen. (Vol. Pedagogisch didactisch seminarie): Academische lerarenopleiding, schooljaar 2003-2004.
- Banks, J., & al., e. (2007). Teaching Diverse Learners. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing Teachers for a Changing World*. San Fransisco.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, 5(1), 7-74.
- Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.
- Bol, T., Witschge, J., Van de Werfhorst, H. G., & Dronkers, J. (2014). Curricular tracking and central examinations: Counterbalancing the impact of social background on student achievement in 36 countries. *Social Forces*, 92(4), 1545-1572.
- Bonner, S. M. (2016). Teachers' Perceptions About Assessment: Competing Narratives. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of Human and Social Conditions in Assessment*. London: Routledge.
- Brookhart, S. M. (2007). Expanding views about formative classroom assessment: A review of the literature. In J. H. Mc Millan (Ed.), *Formative classroom assessment: Theory into practice* (pp. 43-62). New York: Teachers College Press.
- Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in higher education*, 36(4), 395-407.
- Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review*, 24(2), 205-249.
- Dann, R. (2014). Assessment as learning: blurring the boundaries of assessment and learning for theory, policy and practice. *Assessment in Education: Principles, Policy & Practice*, 21(2), 149-166.
- De Backer, F., Van Avermaet, P., & Slembrouck, S. (2016). Schools as laboratories for exploring multilingual assessment policies and practices. *Language and Education*, 1-14.
- De Boer, A., Pijl, S. J., & Minnaert, A. (2011). Regular primary schoolteachers' attitudes towards inclusive education: A review of the literature. *International journal of inclusive education*, 15(3), 331-353.
- De Meyer, I., Janssens, R., Warlop, N., Van Keer, H., De Wever, B., & Valcke, M. (2019). *Leesvaardigheid van 15-jarigen. Vlaams rapport PISA2018*. Retrieved from [http://www.pisa.ugent.be/uploads/assets/157/1582105796291-RAPPORT\\_2018\\_ZonderFoto.pdf](http://www.pisa.ugent.be/uploads/assets/157/1582105796291-RAPPORT_2018_ZonderFoto.pdf)
- Denessen, E. (2000). *Opvattingen over onderwijs. Leerstof- en leerlinggerichtheid in Nederland*. Leuven-Apeldoorn: Garant.
- Devos, G., Vandenberghe, P., & Verhoeven, J. (1989). *Schoolbeleid; mogelijkheden en grenzen; een empirisch onderzoek*: Acco; Leuven/Amersfoort.
- Earl, L. M. (2003). *Assessment As Learning: Using Classroom Assessment to Maximize Student Learning*. London: Sage Publications.

- Ercikan, K., & Solano-Flores, G. (2016). Section Discussion: Assessment and Sociocultural Context: A Bidirectional Relationship. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of Human and Social Conditions in Assessment*. London: Routledge.
- Ertmer, P. A., & Newby, T. J. (2013). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly*, 26(2), 43-71.
- Goosen, K., Boone, S., Dehertogh, B., Kavadias, D., Mahieu, P., Van Avermaet, P., Van Kerckhove, C. (2017). *Is dat iets voor mij, juf? Leerlingen versterken in het keuzeprocess van basis naar secundair*. Leuven: LannooCampus.
- Harris, A., & Spillane, J. (2008). Distributed leadership through the looking glass. *Management in education*, 22(1), 31-34.
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*: Routledge.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112.
- Havnes, A., Smith, K., Dysthe, O., & Ludvigsen, K. (2012). Formative assessment and feedback: Making learning visible. *Studies in Educational Evaluation*, 38(1), 21-27.
- Hopkins, D. (2003). *School improvement for real*: routledge.
- James, M., & Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *The Curriculum Journal*, 17(2), 109-138.
- Johnson, B., & Turner, L. A. (2003). Data collection strategies in mixed methods research. *Handbook of mixed methods in social and behavioral research*, 297-319.
- Jones, N., Saville, N., & Salamoura, A. (2016). *Learning oriented assessment* (Vol. 45): Cambridge University Press.
- Mattheeuws, K. (z.d.). Het evaluatieproces verrijken en leerlingen eigenaar maken van eigen leren. Het leerling-contact. *Zorgbreed*, 67.
- McNamara, T. (2001). Language assessment as social practice: challenges for research. *Language Testing*, 18(4), 333-349.
- Nagy, P. (2000). The three roles of assessment: Gatekeeping, accountability, and instructional diagnosis. *Canadian Journal of Education/Revue canadienne de l'éducation*, 262-279.
- Nayir, F., Brown, M., Burns, D., Joe, O. H., McNamara, G., Nortvedt, G., Wiese, E. F. (2019). Assessment with and for Migration Background Students-Cases from Europe. *Eurasian Journal of Educational Research*, 19(79), 39-68.
- Nicaise, I., Spruyt, B., Van Houtte, M., & Kavadias, D. (2014). *Het onderwijsdebat. Waarom de hervorming van het secundair onderwijs broodnodig is*. Antwerpen: EPO.
- Onderwijsinspectie. (2016). Het referentiekader voor onderwijskwaliteit (OK). Retrieved from [https://www.onderwijsinspectie.be/nl/het-referentiekader-voor-onderwijskwaliteit-ok#Anch\\_05](https://www.onderwijsinspectie.be/nl/het-referentiekader-voor-onderwijskwaliteit-ok#Anch_05)
- Onderwijsinspectie. (2020). *Onderwijsspiegel 2020: jaarlijks rapport van de onderwijsinspectie* Retrieved from Brussel: <https://www.onderwijsinspectie.be/sites/default/files/atoms/files/Onderwijsspiegel%202020%20-%20webversie.pdf>
- Onderwijsraad. (2001). *Advies: Wat scholen vermogen. Autonomie, beleidsvoerend vermogen en bestuurlijke inrichting in het primair en voortgezet onderwijs*. Retrieved from Den Haag:
- Opertti, R., Walker, Z., & Zhang, Y. (2013). Inclusive education: From targeting groups and schools to achieving quality education as the core of EFA. *The SAGE Handbook of Special Education: Two Volume Set*, 1, 149.
- Penninckx, M., Vanhoof, J., & Van Petegem, P. (2011). *Evaluatie in het Vlaamse onderwijs. Beleid en praktijk van leerling tot overheid*. Antwerpen-Apeldoorn: Garant.
- Petry, K., Struyf, E., Verschuere, K., Adriaensens, S., Bodvin, K., Bogaerts, L., ....Verachtert, P. (2020). *M-decreet in de Vlaamse scholen: praktijken en achterliggende processen en mechanismen. Syntheserapport met aanbevelingen voor beleid en praktijk*. Retrieved from Gent:

Philips, I., Seghers, M., Versteden, P., & Ysenbaert, J. (2013). *Toolkit Competenties Nederlands Breed Evalueren*. Retrieved from Leuven

Gent:

<https://onderwijs.vlaanderen.be/sites/default/files/atoms/files/Toolkit%20breed%20evalueren%20competenties%20Nederlands%20in%20het%20lager%20onderwijs.pdf>

- Sach, E. (2015). An exploration of teachers' narratives: what are the facilitators and constraints which promote or inhibit 'good' formative assessment practices in schools? *Education 3-13*, 43(3), 322-335.
- Sahlberg, P. (2006). Education reform for raising economic competitiveness. *Journal of Educational Change*, 7(4), 259-287.
- Schildkamp, K., Lai, M. K., & Earl, L. (2012). *Data-based decision making in education: Challenges and opportunities* (Vol. 17): Springer Science & Business Media.
- Schunk, D. H. (2012). *Learning theories an educational perspective sixth edition*: Pearson.
- Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39-83). Chicago: Rand McNally Shavelson R.
- Shepard, L. A. (2000). The Role of Assessment in a Learning Culture. *Educational Researcher*, 29(7), 4-14.
- Shohamy, E. (2001). *The Power of Tests. A critical perspective on the uses of language tests*. Harlow, England: Longman.
- Struyf, E., Verschueren, K., Petry, K., & Van Mieghem, A. (2018). *Dynamieken achter de implementatie van het M-decreet: Een casestudieonderzoek*. Retrieved from Gent:
- Struyven, K., Gheyssens, E., Coubergs, C., De Doncker, H., & De Neve, D. (2019). *Binnenklasdifferentiatie in de praktijk. Ieders leerkracht realiseren*. Leuven: Acco.
- Teijssen, E., De Fraine, B., & Wouters, R. (2017). Het functioneren van delibererende klassenraden. Stand van zaken in het Vlaamse secundair onderwijs. *Impuls*, 48(2), 69-77.
- Thys, S. (2018). *The tertiary effect of social class: multilevel studies on the role of the primary school teacher in educational decision-making*. Ghent University.
- Tierney, R. (2013). Fairness in Classroom Assessment. In J. H. Mc Millan (Ed.), *Handbook of Research on Classroom Assessment* (pp. 125-141). Thousand Oaks: Sage.
- Tomlinson, & Moon. (2013a). *Assessment and student success in a differentiated classroom*: ASCD.
- Tomlinson, & Moon. (2013b). Differentiation and classroom assessment. In M. Millan (Ed.), *Research on Classroom Assessment*. London: Sage
- Tomlinson, C. A. (2001). *How to differentiate instruction in mixed-ability classrooms*: ASCD.
- Van Avermaet, P., & Sierens, S. (2010). Diversiteit is de norm. Er mee leren omgaan de uitdaging. Een referentiekader voor omgaan met diversiteit in onderwijs. *De Coen, D. ea (Red.), Handboek beleidsvoerend vermogen (Doelgerichte visie, Aflevering 4)*, 1-48.
- Van Houtte, M. (2016). Lower-track students' sense of academic futility: Selection or effect? *Journal of Sociology*, 52(4), 874-889.
- Van Houtte, M., & Stevens, P. A. (2010). The culture of futility and its impact on study culture in technical/vocational schools in Belgium. *Oxford Review of Education*, 36(1), 23-43.
- Van Houtte, M., & Stevens, P. A. (2015). Tracking and sense of futility: The impact of between-school tracking versus within-school tracking in secondary education in Flanders (Belgium). *British educational research journal*, 41(5), 782-800.
- Van Petegem, P., & Vanhoof, J. (2002). *Evaluatie op de testbank. Een handboek voor het ontwikkelen van alternatieve evaluatievormen*. Mechelen: Wolters Plantyn.
- Vanbuel, M., Vandommele, G., & Van den Branden, K. (2020). *Taalvaardigheid screenen aan de start. De implementatie en implicaties van een verplichte low-stakes taalscreening in lagere en secundaire scholen*. Retrieved from Gent:
- Vandenbergh, R. (2005). Beleidsvoerend vermogen van scholen: verantwoording en controle *Beleidsvoerend vermogen van scholen ontwikkelen. Een verkenning*. (pp. 175-192). Antwerpen-Apeldoorn: Vlaamse Onderwijsraad & Garant-Uitgevers.

- Vanhoof, J., Deneire, A., & Van Petegem, P. (2011). *Waar zit beleidsvoerend vermogen in (ver)scholen?* Mechelen: Plantyn.
- Vantieghem, W., Van Avermaet, P., Groenez, S., & Lambert, M. (2018). Diversiteitsbarometer Onderwijs Vlaamse Gemeenschap *Diversiteitsbarometer onderwijs* (pp. 47-193): Els Keytsman.
- William, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37(1), 3-14.
- Ysenbaert, J., Van Avermaet, P., & Van Houtte, M. (2017). Literatuurstudie evaluatie en diversiteit. . Steunpunt Onderwijsonderzoek, Gent.
- Ysenbaert, J., Van Avermaet, P., & Van Houtte, M. (2018). *Evaluatie & diversiteit in het basis- en secundair onderwijs. Deelrapport: casestudieonderzoek.*, Steunpunt Onderwijsonderzoek, Gent.
- Ysenbaert, J., Van Avermaet, P., & Van Houtte, M. (n.d.). *Evaluatie & diversiteit in het basis- en secundair onderwijs. Deelrapport: survey-onderzoek.*, Steunpunt Onderwijsonderzoek, Gent.
- Ysenbaert, J., Van Houtte, M., & Van Avermaet, P. (2020). Assessment policies and practices in contexts of diversity: unravelling the tensions. *Educational Assessment, Evaluation and Accountability*. doi:10.1007/s11092-020-09319-7
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational psychologist*, 25(1), 3-17.