

Kwaliteitsmonitoring in het Vlaamse onderwijs

Onderzoek naar het samenspel van het OK-kader, de doorlichting en gestandaardiseerde toetsen

Beleidsamenvatting OBPWO-project 21.01

Mei 2023

Jan Vanhoof (Universiteit Antwerpen, promotor-woordvoerder)

Randi Buvens (Universiteit Antwerpen, onderzoeker)

Lore Pelgrims (KU Leuven, onderzoeker)

Koen Aesaert (KU Leuven, promotor)

Jonas Dockx (KU Leuven, promotor)

Peter Van Petegem (Universiteit Antwerpen, promotor)

Inhoud

Abstract	3
1. Inleiding	4
2. Leeswijzer bij het rapport.....	5
3. Overzicht gehanteerde onderzoeksmethoden	7
3.1 De doorlichting en het OK-kader als hefboomen voor onderwijskwaliteit.....	7
3.1.1 Literatuuronderzoek.....	7
3.1.2 Gevalstudies naar proces en effecten van Inspectie 2.0.....	8
3.1.3 Diepte-interviews naar OK-kader en Inspectie 2.0	8
3.1.4 Survey naar OK-kader en Inspectie 2.0	9
3.2 Bronnen van externe kwaliteitszorg als hefboom voor onderwijskwaliteit	9
3.2.1 Documentanalyse.....	9
3.2.2 Cross-instrumentele studies.....	10
3.2.3 Focusgroepen	11
4. Conclusie: lessen uit de verschillende deelstudies	12
4.1 De doorlichting en het OK-kader als hefboomen voor onderwijskwaliteit.....	12
4.1.1 Over processen en effecten van Inspectie 2.0	13
4.1.2 Over het Referentiekader Onderwijskwaliteit	20
4.2 Bronnen van externe kwaliteitszorg als hefboom voor onderwijskwaliteit	24
4.2.1 Gelijkenissen en verschillen tussen bronnen van externe kwaliteitszorg.....	25
4.2.2 Mate van convergentie of divergentie tussen de resultaten van verschillende bronnen ...	27
4.2.3 Complementariteit van de verschillende bronnen van externe kwaliteitszorg	29
5. Beleidsaanbevelingen.....	35
6. Referentielijst	43

Abstract

Aanleiding van dit onderzoek is een vernieuwde doorlichtingsaanpak 'Inspectie 2.0' en de introductie van het Referentiekader voor Onderwijskwaliteit (OK-kader) met bijhorende kwaliteitsverwachtingen, dat de krijtlijnen uittekent voor kwaliteitsvol onderwijs. Naast doorlichtingen volgens de principes van Inspectie 2.0, doet de Vlaamse overheid ook beroep op verschillende andere bronnen van externe kwaliteitszorg om de kwaliteit van het Vlaamse onderwijs te monitoren, namelijk peilingsonderzoek en internationaal vergelijkend onderzoek. Elk van deze bronnen van externe kwaliteitszorg biedt specifieke informatie over de Vlaamse onderwijskwaliteit, toch blijken de resultaten en bijhorende conclusies van de verschillende bronnen niet altijd in eenzelfde lijn te liggen. We gaan daarom op zoek naar antwoorden op de vragen of en hoe (1) de doorlichting en het OK-kader en (2) de verschillende bronnen van kwaliteitszorg (kunnen) functioneren als hefboomen voor onderwijskwaliteit. In een eerste deelrapport werd een gevalstudieaanpak uitgerold, in combinatie met 24 diepte-interviews en een surveyluik. Dit deelrapport focust op de doorlichting en het OK-kader als hefboomen voor onderwijskwaliteit. Een tweede deelrapport omvat een documentanalyse, cross-instrumenteel onderzoek en focusgroepgesprekken in teken van het samenspel van de verschillende externe bronnen van kwaliteitszorg.

Op basis van de onderzoeksresultaten van het eerste deelrapport stellen we vast dat de vernieuwde doorlichtingsaanpak goed onthaald werd. Inspectie 2.0 kan echter maar fungeren als hefboom voor onderwijskwaliteit als bepaalde randvoorwaarden vervuld zijn. Zo zijn de professionaliteit en geloofwaardigheid van de betrokken inspecteur(s), de mate van onderwijskundig leiderschap van de schoolleider en het innovatief vermogen van het schoolteam belangrijke voorspellers die zorgen dat er met doorlichtingsresultaten aan de slag gegaan wordt. De (f)actoren die het ondernemen van actie het meest in de weg staan zijn een gebrek aan continuïteit in samenstelling van schoolteams en een gebrek aan (gepaste) ondersteuning. Wat het OK-kader betreft stellen we slechts een matige vertrouwdheid vast. Schoolleiders zijn er beduidend meer mee vertrouwd zijn dan leerkrachten. De interviews leren ons dat vertrouwdheid niet per definitie impliceert dat het OK-kader ook een weg gevonden heeft binnen de schoolwerking. Daar waar dat wel het geval is, wordt het OK-kader omschreven als een instrument dat goede houvast, structuur en richting biedt. Om het OK-kader in meer scholen te doen landen, vormen het compenseren van een gebrek aan gebruiksvriendelijkheid en het inzetten op de gepercipieerde relevantie van het kader door de schoolleider de belangrijkste randvoorwaarden.

In het tweede deelrapport stellen we zowel gelijkenissen als verschillen vast wanneer we het onderzoekproces van de externe kwaliteitsbronnen in kaart brengen. Een vergelijking van de resultaten van die bronnen wijst op een relatief convergent beeld op systeemniveau, zeker na controle voor achtergrondkenmerken. Wel zijn er ook steeds individuele scholen waarvoor de resultaten van verschillende bronnen voor externe kwaliteitszorg niet overeenkomen. Dit onderzoek toont ook dat de diversiteit onder de verschillende bronnen een meerwaarde is. Elke bron heeft zijn eigen doel en onderzoeksproces dat leidt tot bepaalde resultaten en conclusies. Door meerdere bronnen te combineren en een goed beeld te hebben wat elke bron wel en niet meet, kunnen de verschillende bronnen van externe kwaliteitszorg samen fungeren als een hefboom voor onderwijskwaliteit.

1. Inleiding

Het debat over dalende onderwijskwaliteit is de laatste jaren niet weg te denken uit ons onderwijslandschap. Om die onderwijskwaliteit te monitoren en kwaliteitsontwikkeling te stimuleren, is in Vlaanderen het gebruik van schooldoorlichtingen goed ingeburgerd. Net als in vele andere landen, maakt het gebruik van schooldoorlichtingen het mogelijk om scholen te beoordelen, stimuleren en verantwoordelijk te houden voor het bewaken en verbeteren van hun eigen onderwijskwaliteit (Gärtner, Wurster, & Pant, 2014). Hoewel ingebed in ruimere internationale tendensen, kent Vlaanderen daartoe een unieke aanpak. Tot januari 2018 gebruikte de Vlaamse onderwijsinspectie het CIPO-model (acroniem voor context, input, proces en output; Scheerens, 2011) als leidraad voor haar doorlichtingen. Sinds begin 2018 wordt evenwel vertrokken van het zogenaamde 'Referentiekader voor Onderwijskwaliteit (OK-kader)'. Dit OK-kader is het resultaat van een unieke samenwerking tussen de Vlaamse onderwijsaanbieders, het Vlaamse onderwijsveld en de Vlaamse Onderwijsinspectie. De kern van het referentiekader bestaat uit kwaliteitsverwachtingen die verdeeld zijn over vier rubrieken, met name 'resultaten en effecten', 'ontwikkeling van de lerenden stimuleren', 'kwaliteitsontwikkeling' en 'beleid' (Bronnendocument OK-kader, 2017). Het is dit nieuwe raamwerk met geëxpliciteerde verwachtingen omtrent hoogwaardig kwalitatief onderwijs dat wordt gehanteerd in de nieuwe doorlichtingsaanpak die door de Vlaamse Onderwijsinspectie begin 2018 gelanceerd werd, genaamd 'Inspectie 2.0'.

Naast doorlichtingen doet de Vlaamse overheid ook beroep op verschillende andere bronnen van externe kwaliteitszorg om de kwaliteit van het Vlaamse onderwijs te monitoren, met name het peilingsonderzoek en het internationaal vergelijkend onderzoek. Elk van deze bronnen van externe kwaliteitszorg biedt eigen specifieke informatie over de Vlaamse onderwijskwaliteit. Zo vergelijkt het internationaal vergelijkend onderzoek (zoals PISA, PIRLS, TIMSS) de prestaties van Vlaamse leerlingen met die van leerlingen uit andere onderwijssystemen (Mullis et al., 2017, 2020; OECD, 2019). De peilingen daarentegen beschrijven het aandeel leerlingen dat de onderwijsdoelen bereikt (AHOVOKS, 2021). Vanaf het schooljaar 2023-2024 worden de peilingen vervangen door de zogenaamde Vlaamse toetsen.

Tot op heden weten we weinig over de ervaringen van het werkveld met de nieuwe manier van doorlichten en de effecten die Inspectie 2.0 met zich meebrengt. Daarnaast is ook de mate waarin het onderwijsveld vertrouwd is met het OK-kader en welke impact de invoering van dit kader gehad heeft op het beleid van scholen tot dusver niet de focus geweest van systematisch onderzoek. Wat de verschillende bronnen van externe kwaliteitszorg betreft, blijken de resultaten van deze verschillende bronnen en de daarmee samenhangende conclusies over onderwijskwaliteit niet steeds in dezelfde lijn te liggen. De vraag rijst dan ook in welke mate deze resultaten overeenkomen of verschillen en hoe deze verschillende bronnen van externe kwaliteitszorg complementair ingezet kunnen worden. Zorgt het samenspel van de doorlichting, het OK-kader en de gestandaardiseerde toetsen in een Vlaamse onderwijscontext voor gepaste instrumenten om onderwijskwaliteit te monitoren en kwaliteitsontwikkeling te stimuleren? En welke randvoorwaarden zijn daarbij nodig?

Deze bevindingen leiden tot een tweeledige set van onderzoeksvragen, zodat ook de onderzoeksresultaten gepresenteerd worden in twee deelrapporten. In het eerste deelrapport zoomen we specifiek in op de percepties van het werkveld op de nieuwe manier van doorlichten en het nieuwe Referentiekader Onderwijskwaliteit. We gaan daarbij niet op zoek naar antwoorden op de mate van kwaliteitszorg aanwezig in scholen, wel naar vertrouwdheid met en mate van gebruik van het Referentiekader Onderwijskwaliteit. Het dient daarbij expliciet te worden benadrukt dat kennis of gebruik van het OK-kader niet gelijk staat aan een kwaliteitsvolle schoolwerking. Het tweede

deelrapport focust op de verschillende bronnen van externe kwaliteitszorg. Voor een uitgebreidere inleiding op beide thema's verwijzen we graag naar de desbetreffende deelrapporten.

Deelrapport 1: Inspectie 2.0 en het Referentiekader Onderwijskwaliteit als hefboomen voor onderwijskwaliteit.

OV 1.1 Welke zijn effecten van de doorlichting in scholen?

OV 1.2 Welke (f)actoren (inspectie-, school-, gebruikers-en contextgebonden) spelen een rol in het verklaren van effecten van de doorlichting in scholen?

OV 2.1 Hoe competent zijn Vlaamse leerkrachten en schoolleiders in het gebruik van het OK-kader?

OV 2.2 Heeft de invoering van het OK-kader impact (gehad) op het beleid van scholen? Zo ja, op welke manier?

OV 2.3 Welke (f)actoren (OK-kader-, school-, gebruikers-en contextgebonden) spelen een rol in de ontwikkeling en aanwending van competenties inzake OK-kadergebruik?

Deelrapport 2: Bronnen van externe kwaliteitszorg als hefboom voor onderwijskwaliteit

OV 3.1 Welke gelijkenissen en verschillen zijn er tussen de doorlichting, het peilingsonderzoek en het internationaal vergelijkend onderzoek inzake de evaluatiedoelen, de getoetste kwaliteitsindicatoren, de gehanteerde maatstaven en rapporteringsniveaus, de gemeten inhoud, de getoetste steekproef/populatie en de gehanteerde instrumenten?

OV 3.2 Uitgaande van de resultaten van deze verschillende bronnen, in welke mate convergeren of divergeren de uitspraken over onderwijskwaliteit het voorbije decennium?

OV 3.3 In welke mate duiden de vastgestelde gelijkenissen en verschillen tussen de bronnen (zie OV 3.1) de convergerende en divergerende uitspraken over onderwijskwaliteit (zie OV 3.2)?

OV 3.4 Op welke manier kunnen de verschillende bronnen en de daarbij horende data worden gecombineerd om tot meer betrouwbare en accurate uitspraken over onderwijskwaliteit te komen op systeem- en schoolniveau?

2. Leeswijzer bij het rapport

Deze beleidssamenvatting rapporteert over zowel het eerste als het tweede deelrapport van het OBPWO-project 'Kwaliteitsmonitoring in het Vlaamse onderwijs'. De samenvatting kan gelezen worden zonder de afzonderlijke deelrapporten te bekijken waarop ze gebaseerd is. Voor meer achtergrondinformatie over de methodologie, een uitgebreide beschrijving van de resultaten en een beschrijvende conclusie verwijzen we graag naar de deelrapporten en desbetreffende modules.

Figuur 1 toont de structuur van het volledige OBPWO-rapport. Het eerste deelrapport omvat vier verschillende op zichzelf staande modules die samen een antwoord bieden op de onderzoeksvragen (OV 1.1 - OV 2.3):

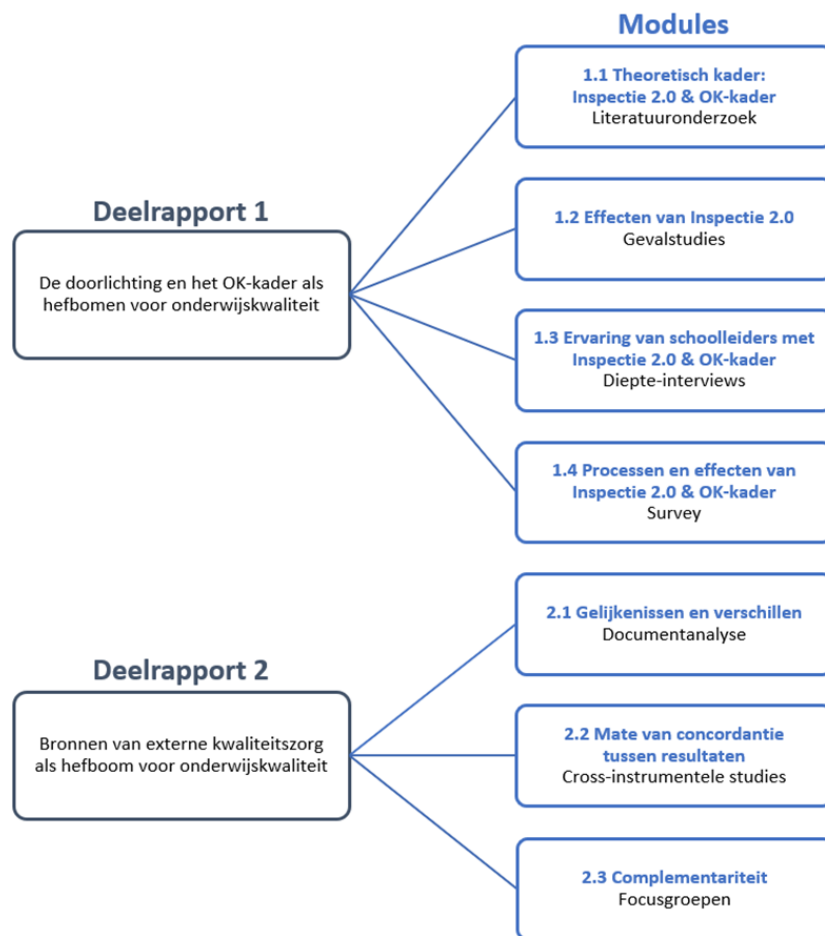
- In Module 1.1 schetsen we een conceptueel kader op basis van een literatuurstudie die focust op schooldoorlichtingen, Inspectie 2.0 en het Referentiekader Onderwijskwaliteit. Deze module vormt de basis voor de daaropvolgende modules van het eerste deelrapport.

- In Module 1.2 worden op basis van tien gevalstudies de korte- en lange termijneffecten van Inspectie 2.0 beschreven.
- Module 1.3 beschrijft de resultaten van 24 diepte-interviews afgenomen met schoolleiders van recent doorgelichte en niet doorgelichte scholen, waarbij opnieuw processen en effecten van de doorlichting in kaart gebracht worden en we tevens een blik werpen op de mate van vertrouwdheid van schoolleiders met het OK-kader en of de invoering van dit kader mogelijke effecten met zich meebracht.
- Module 1.4 bevat de resultaten van een surveyonderzoek, waarbij zowel schoolleiders als leerkrachten bevestigd werden naar hun ervaringen met en de effecten van Inspectie 2.0 en het OK-kader.

De onderzoeksvragen van het tweede deelrapport (OV 3.1- OV 3.4) worden beantwoord in modules 2.1 tot en met 2.3:

- In Module 2.1 beschrijven we de gelijkenissen en verschillen in het onderzoeksopzet tussen de doorlichtingen, het peilingsonderzoek en het internationaal vergelijkend onderzoek (zie OV 3.1).
- Module 2.2 omvat de resultaten van drie cross-instrumentele studies die nagaan in welke mate de uitspraken over aspecten van onderwijskwaliteit van verschillende bronnen convergeren of divergeren (zie OV 3.2).
- Module 2.3 heeft een tweeledig doel. We onderzoeken in welke mate de gelijkenissen en verschillen – uit Module 2.1 – een verklaring vormen voor de resultaten uit Module 2.2 (zie OV 3.3). Daarnaast belichten we op basis van de resultaten van de focusgroepen hoe de verschillende bronnen van externe kwaliteitszorg gecombineerd kunnen worden (OV 3.4).

Figuur 1
Structuur van het OBPWO-rapport



In wat volgt, lichten we beknopt de methodologie voor beide deelrapporten toe. Daarna volgt voor beide deelrapporten een conclusie met daarin de antwoorden op de onderzoeksvragen en de belangrijkste bevindingen en inzichten overheen de verschillende modules. Vanuit de synthese van de deelrapporten reiken we ten slotte enkele concrete aandachtspunten of aanbevelingen voor het beleid aan.

3. Overzicht gehanteerde onderzoeksmethoden

3.1 De doorlichting en het OK-kader als hefboomen voor onderwijskwaliteit

Om de processen en effecten van een doorlichting en het gebruik van het OK-kader te beschrijven en verklaren, werd een beroep gedaan op verschillende vormen van gegevensverzameling. Het eerste deel van het onderzoeksrapport bestaat dan ook uit vier afzonderlijk te raadplegen modules, die -waar relevant- telkens opnieuw voorzien worden van een inleiding, een eigen methodologische uiteenzetting, een rapportage van de belangrijkste resultaten en een conclusie.

3.1.1 Literatuuronderzoek

Dit deelrapport start met een literatuuronderzoek (Module 1.1) dat in kaart brengt wat er over de te onderzoeken thema's reeds beschreven is, internationaal en in Vlaamse literatuur. Door middel van een systematisch literatuuronderzoek werd een update gemaakt van de bestaande (inter)nationale kennisbasis omtrent effecten, processen en randvoorwaarden van een doorlichting. Daarnaast bevat deze module ook een toelichting bij het Referentiekader Onderwijskwaliteit. De literatuur die werd

doorgenomen betreft zowel theoretische inzichten, empirische studies als cases van ‘best practices’. Behalve door middel van systematische raadpleging van databanken, werd ook de sneeuwbalmethode gebruikt in de zoektocht naar relevante literatuur. Op basis van deze literatuurstudie werd een conceptueel kader uitgewerkt op basis waarvan de lijnen voor het verder kwantitatief en kwalitatief onderzoek werden uitgezet. Het multimethode onderzoeksdesign (Tashakkori & Teddlie, 2003) omvat immers een procedure voor het verzamelen, analyseren en combineren van kwantitatieve en kwalitatieve gegevens in één studie. De verschillende onderzoeksmethoden op elkaar laten verder bouwen (cf. sequentieel) en deze ondertussen concurrent met elkaar verbinden stelt ons in staat de onderzoeksvragen breder en diepgaander te beantwoorden (Creswell, 2005).

3.1.2 Gevalstudies naar proces en effecten van Inspectie 2.0

Module 1.2 presenteert de resultaten van een gevalstudie-aanpak, uitgerold in tien basisscholen die doorgelicht werden in schooljaar 2018-2019. Dit is de (eerste) groep van scholen die volledig volgens de aanpak van Inspectie 2.0 doorgelicht werd (zonder corona-implicaties) en die vervolgens drie schooljaren de tijd en verantwoordelijkheid kreeg met de resultaten van deze doorlichting aan de slag te gaan. Onderzoeksgroep Edubron heeft bij een steekproef uit deze populatie van in schooljaar 2018-2019 doorgelichte basisscholen twee à drie maanden na de doorlichting een survey-onderzoek gedaan naar processen van de doorlichting en naar (intenties bij) gebruik van de ontvangen feedback (zie Quintelier, 2020). Deze groep van scholen is een toevallige selectie uit de scholen die dat schooljaar doorgelicht werden. Data van in totaal 80 basisscholen werden verzameld, met telkens de inbreng van de schoolleider en een groep van leraren. Tien van deze scholen werden in de gevalstudie betrokken, wat leidde tot een unieke kans om data van kort na de doorlichting te koppelen aan de kwalitatieve data verzameld 3 jaar na de doorlichting. Met het uitvoeren van deze gevalstudies willen we aandacht besteden aan de complexe sociale interacties en aan de betekenis die de deelnemers in de bestudeerde systemen geven aan elkaar en aan de gebeurtenissen binnen en na het doorlichtingsproces. Swanborn (2010) geeft aan dat er in gevalstudies aandacht besteed wordt aan de zogenaamde ‘meervoudige realiteiten’. Dat doen we door in deze scholen telkens de directeur en (gezamenlijk) twee leraren te bevragen. De ‘eenvoudige’ modellen die ten grondslag liggen aan statistische analyses doen de complexiteit van het functioneren van scholen immers te kort. De keuze van deze 10 cases baseerden we op theoretische overwegingen (bv. hoge/lage tevredenheid bij het inspectieverloop, hoge/lage intentie om met resultaten aan de slag te gaan) die geïnformeerd werden door de reeds uitgevoerde survey. De voornaamste bron van data waren de mondelinge diepte-interviews met schoolleiders en leerkrachten (Hutjes & Van Buuren, 1992; Denzin & Lincoln, 2000). De casestudies werden verder ook geïnformeerd door een oriënterende analyse van relevante documenten zoals bijvoorbeeld de doorlichtingsverslagen.

3.1.3 Diepte-interviews naar OK-kader en Inspectie 2.0

Aan de hand van 24 afgenomen diepte-interviews (Module 1.3) geven we op een zodanige manier antwoord op OV1 en OV2 dat de specifieke eigenheid van verschillende contexten in kaart wordt gebracht. Eigen aan semigestructureerde diepte-interviews is dat formulering, inhoud en antwoordkeuze van de vragen niet volledig vastliggen (Mortelmans, 2007). Dit maakt dat, naargelang het gebruik van het OK-kader en het specifieke verloop van de doorlichting en de opvattingen van de respondent daaromtrent bepaalde topics meer of minder diepgaand aan bod komen en een zowel brede als diepgaande dataverzameling gefaciliteerd wordt. Alle interviews werden digitaal opgenomen en getranscribeerd met het oog op analyse aan de hand het softwareprogramma Nvivo. Daarbij hanteerden we de principes van Miles en Huberman (1994). Dat betekent dat de analyse zowel geïnspireerd werd door het conceptueel kader uit Module 1.1 (deductief) als door de inbreng van de respondenten zelf (inductief). Naast de semigestructureerde interviews is er ook sprake van documentanalyse van bijvoorbeeld doorlichtingsverslagen. In totaal werden 24 interviews afgenomen binnen 12 basisscholen

en 12 secundaire scholen. We brengen de onderzoeksvragen in dit kwalitatieve luik in kaart vanuit het perspectief van de schoolleider, omdat deze van alle betrokkenen het best geplaatst is de onderzoekers te informeren. Schoolleiders zijn de verbindingsfiguren tussen de leraren, het schoolbestuur en pedagogische begeleidingsdiensten. Zij hebben van alle personeelsleden in scholen de meest brede kijk op de te bestuderen fenomenen. In de interviews met schoolleiders bevroegen we niet alleen hun eigen perspectief, maar ook hoe zij de impact van het OK-kader en de doorlichting op hun leraren, schoolbestuur en hun pedagogische begeleidingsdienst ervaren (hebben). Met het oog op een rijke en diepgaande beschrijving van de onderzoeksthematiek werd gestreefd naar een heterogene steekproef, waarbij ook schoolleiders van scholen uit het buitengewoon basis- en secundair onderwijs betrokken werden, alsook schoolleiders waarvan de doorlichting resulteerde in een ongunstig advies.

3.1.4 Survey naar OK-kader en Inspectie 2.0

Het eerste deelrapport van de studie eindigt met een kwantitatief luik dat bestaat uit een surveyonderzoek (Module 1.4). Aan de hand van reeds bestaande en zelf ontwikkelde instrumenten bracht een digitale vragenlijst de verschillende aspecten uit het theoretisch kader bij OV1 en OV2 in kaart. De onderzoekspopulatie bestond uit alle Vlaamse reguliere basis- en secundaire scholen. Daaruit werd een steekproef getrokken waarbij uiteindelijk 612 respondenten deelnamen, verspreid over 97 verschillende basis- en secundaire scholen. De stratificatie gebeurde in eerste instantie naar al dan niet doorgelicht geweest zijn sinds januari 2018. Deze verhouding bedroeg respectievelijk 60% /40%. Op die manier kunnen we de link tussen al dan niet doorgelicht worden en competenties inzake OK-kadergebruik in de analyses meenemen. We kozen verhoudingsgewijs voor een grotere groep doorgelichte scholen. Dat deden we om de onderzoeksinhouden bij het thema doorlichting in een voldoende grote set van scholen te kunnen bevragen (merk ook op dat vragen m.b.t. het OK-kader in alle scholen aan bod kwamen). De stratificatie gebeurde verder ook naar onderwijsnet (en dus aard van schoolbestuur en ondersteunende pedagogische begeleidingsdienst). Het opzet was om in deze scholen de schoolleider en daarnaast zoveel mogelijk leden van –indien aanwezig –het middenkader te bevragen. Voor de analyse van de survey-gegevens deden we een beroep op beschrijvende statistiek, maar ook op regressieanalyses. De veronderstelling bij dit opzet is dat de kwantitatieve onderzoeksbevindingen uit het survey-onderzoek het beeld dat verkregen werd uit de eerdere fasen (gevalstudies en diepte-interviews) zowel verbreden als verdiepen (Creswell, 2005).

3.2 Bronnen van externe kwaliteitszorg als hefboom voor onderwijskwaliteit

Om een antwoord te formuleren op de onderzoeksvragen met betrekking tot de verschillende bronnen van externe kwaliteitszorg, werd ook in het tweede deelrapport beroep gedaan op verschillende onderzoeksmethoden. Dit rapport start met een documentanalyse waarin de gelijkenissen en verschillen tussen de bronnen van externe kwaliteitszorg in kaart gebracht worden. In een tweede fase werden er drie cross-instrumentele studies uitgevoerd, die nagaan in welke mate de resultaten van verschillende bronnen voor externe kwaliteitszorg convergeren of divergeren. Dit deelrapport eindigt met een kwalitatief luik, waarin twee focusgroepen georganiseerd werden. Dat luik richt zich op mogelijke verklaringen voor de concordantie tussen de resultaten van de verschillende bronnen enerzijds en op hoe die bronnen complementair ingezet kunnen worden anderzijds.

3.2.1 Documentanalyse

In Module 2.1 van het tweede deelrapport beschrijven we op basis van een documentanalyse de gelijkenissen en verschillen tussen de doorlichting, het peilingsonderzoek en het internationaal vergelijkend onderzoek (zie Onderzoeksvraag 3.1). Om tot die gelijkenissen en verschillen te komen, werden de onderstaande bronnen vergeleken:

- peiling basisonderwijs Nederlands 2018 en wiskunde 2021;
- peiling wiskunde A-stroom 2018 en B-stroom 2019;
- PIRLS 2016;
- TIMSS 2019;
- PISA 2018;
- doorlichtingen van de onderwijsinspectie volgens de aanpak bij Inspectie 2.0.

In een eerste stap werden de documenten van elke bron afzonderlijk gecodeerd aan de hand van de volgende dimensies die gebaseerd zijn op het analysekader van Treviño en Órdenes (2017):

- design, bestaande uit: (1) doel, (2) domein en construct, (3) populatie, (4) steekproef en (5) modus;
- standaardbepaling
- analytische methode;
- besluitvorming.

Dat leidde per bron tot een verticale analyse en een bijhorende ankertekst, waarin voor elke bron elke dimensie van het analysekader toegelicht wordt. In een tweede stap hebben we de gelijkenissen en verschillen tussen de bronnen in kaart gebracht via een horizontale analyse. Concreet hebben we voor elke dimensie uit het analysekader een vergelijking gemaakt tussen de informatie die opgenomen is in de verticale analyses van de verschillende bronnen. Verder bouwend op de documentanalyse ontwikkelden we een matrix. Daarin wordt voor elke bron van externe kwaliteitszorg aangegeven wat de mogelijkheden en beperkingen zijn om betrouwbare en valide uitspraken te doen over specifieke aspecten van onderwijskwaliteit op een specifiek aggregatieniveau (systeem-, school- en leerlingniveau).

3.2.2 Cross-instrumentele studies

Module 2.2 omvat de resultaten van de drie cross-instrumentele studies die nagaan in welke mate de uitspraken over onderwijskwaliteit van verschillende bronnen convergeren of divergeren (zie Onderzoeksvraag 3.2). In Studie 1 onderzoeken we de concordantie tussen de resultaten op de PIRLS-schaal begrijpend lezen en peilingschaal lezen aan de hand van de data van PIRLS Repeat. Meer specifiek schatten we het schooleffect van alle PIRLS Repeat scholen op zowel de PIRLS-schaal als de peilingschaal, waarna we de schooleffecten op die twee schalen kunnen vergelijken. Om rekening te houden met factoren waarover scholen geen controle hebben (bijvoorbeeld kenmerken van hun leerlingpubliek), maken we daarbij gebruik van toegevoegde waarde modellen. Ook gaan we voor de deelnemende scholen na in welke mate de resultaten op de PIRLS- en peilingschaal convergeren of divergeren met de resultaten uit de doorlichtingen voor het leergebied Nederlands.

Studie 2 focust op de concordantie op schoolniveau tussen de verschillende bronnen van externe kwaliteitszorg. Meer specifiek beantwoorden we de vraag in welke mate het resultaat van een bepaalde school op de ene bron overeenkomt met het resultaat van dezelfde school bij een andere bron. Om die vraag te beantwoorden, hebben we de schooleffecten geschat en vergeleken voor scholen die deelnamen aan een van de volgende combinaties:

Basisonderwijs:

- PIRLS 2016 en de doorlichtingen voor het leergebied Nederlands;
- TIMSS 2019 en de doorlichtingen voor het leergebied wiskunde;
- Peiling Nederlands 2018 en de doorlichtingen voor het leergebied Nederlands;
- Peiling wiskunde 2021 en de doorlichtingen voor het leergebied wiskunde;
- PIRLS 2016 en peiling Nederlands 2018;

- TIMSS 2019 en peiling wiskunde 2021.

Secundair onderwijs:

- PISA 2018 (leesvaardigheid) en de doorlichtingen voor het vak Nederlands;
- PISA 2018 (wiskundige geletterdheid) en de doorlichtingen voor het vak wiskunde;
- Peiling wiskunde A-stroom 2018 en de doorlichtingen voor het vak wiskunde;
- Peiling wiskunde B-stroom 2019 en de doorlichtingen voor het vak wiskunde;
- PISA 2018 (wiskundige geletterdheid) en peiling A-stroom 2018;
- PISA 2018 (wiskundige geletterdheid) en peiling B-stroom 2019.

In Studie 3 gaan we na of bronnen die een gelijkaardig domein of construct meten ook dezelfde trends op systeemniveau vinden. Meer bepaald maken we in die studie een vergelijking tussen de volgende trends:

- Doorlichtingsronde 3 en Inspectie 2.0;
- PIRLS en de peiling Nederlands basisonderwijs;
- TIMSS en de peiling wiskunde basisonderwijs;
- PISA (wiskundige geletterdheid) en de peiling wiskunde A- en B-stroom.

3.2.3 Focusgroepen

In Module 2.3 willen we in de eerste plaats tot mogelijke verklaringen komen voor de concordantie tussen de verschillende bronnen van externe kwaliteitszorg (Onderzoeksvraag 3.3). In de tweede plaats willen we een beter beeld krijgen van hoe de verschillende bronnen complementair ingezet kunnen worden (Onderzoeksvraag 3.4).

Om bovenstaande onderzoeksvragen te beantwoorden, hebben we focusgroepgesprekken georganiseerd. Daarvoor selecteerden we 15 participanten die rechtstreeks in contact komen met de ontwikkeling en rapportage van minstens één van de drie bronnen voor externe kwaliteitszorg. In totaal organiseerden we twee focusgroepen die online afgenomen werden. Op basis van de digitale opnames werden de focusgroepen volledig getranscribeerd. Vervolgens werden die transcripten gelezen, herlezen en gecodeerd. De codering gebeurde zowel inductief als deductief. In eerste instantie werd elke focusgroep afzonderlijk geanalyseerd (within-case analysis). Vervolgens werden twee afzonderlijke analyses systematisch met elkaar vergeleken (cross-case analysis). Doorheen het hele proces van data-analyse werd er steeds gezocht naar patronen en verbanden tussen de data (Miles & Huberman, 2014).

4. Conclusie: lessen uit de verschillende deelstudies

In wat volgt lichten we de conclusies uit beide deelrapporten toe. Hoewel de verschillende thema's op elkaar inhaken presenteren we de conclusies toch per set van onderzoeksvragen. Een complex systeem van kwaliteitszorg overheen thema's en gaande van micro- tot macroprocessen in onderwijs vergt steeds een keuze van perspectief. Elk perspectief biedt mogelijkheden en beperkingen. Hieronder blijven wij doelbewust dicht bij de onderzoeksresultaten om aan de lezer zo zuiver mogelijk de bevindingen mee geven. Elk onderzoeksthema komt daarom afzonderlijk aan bod. Een ander, meer geïntegreerd, perspectief zou met zich meebrengen dat leemtes, mogelijke discussiepunten en te maken keuzes in het samenspel van het OK-kader, de doorlichting en gebruik van gestandaardiseerde toetsen door de onderzoekers ingevuld dienen te worden op een manier die niet volledig door empirische evidentie geschraagd kan worden. Het is aan de gebruiker van deze studie om geïnformeerd door de onderzoeksconclusies vanuit de eigen kijk op onderwijskwaliteit en vanuit de eigen opdracht binnen het onderwijsveld verder invulling te geven aan kwaliteitszorgprocessen in het Vlaamse onderwijs. Een aantal aanzetten daartoe, vertrekkend vanuit de lessen uit dit onderzoek, willen we evenwel wel graag meegeven. Op basis van de bevindingen uit beide deelrapporten formuleren we daarom aan het eind een aantal specifieke en thema-overstijgende beleidsaanbevelingen.

4.1 De doorlichting en het OK-kader als hefboomen voor onderwijskwaliteit

Dit deel van de conclusie brengt de resultaten samen van de verschillende modules die deel uitmaken van het eerste deelrapport van het OBPWO-rapport *Kwaliteitsmonitoring in het Vlaamse Onderwijs*. Centraal staat de vraag of de doorlichting en het OK-kader momenteel functioneren als hefboomen voor onderwijskwaliteit. We gaan daarbij op zoek naar antwoorden op de onderzoeksvragen omtrent Inspectie 2.0 en het OK-kader met aandacht voor de thema's kwaliteitsontwikkeling, leerlingenbegeleiding en leerlingevaluatie. We benadrukken hierbij dat dit onderzoek de percepties van schoolleiders en leraren in kaart brengt. Daarnaast is omwille van haalbaarheidscriteria het perspectief van andere stakeholders (zoals deze van inspecteurs of pedagogisch begeleiders) in dit onderzoek niet kunnen meegenomen worden.

We starten deze conclusie met de beschrijving van de effecten die zich voordoen in scholen naar aanleiding van een doorlichting, gevolgd door de (f)actoren die een rol spelen in het verklaren van de aan- of afwezigheid van deze effecten. In een tweede deel beschrijven we hoe vertrouwd het werkveld momenteel is met het Referentiekader Onderwijskwaliteit, tot welke effecten de invoering van dit OK-kader geleid heeft en welke (f)actoren daarbij een belemmerende of bevorderende rol spelen. Afhankelijk van de onderzoeksvragen baseren we ons op de gecombineerde resultaten van de tien gevalstudies (Module 1.2), de 24 diepte-interviews met de schoolleiders (Module 1.3) en het surveyonderzoek (Module 1.4) zoals weergegeven in Tabel 1.

Tabel 1

Overzicht van de modules uit Deelrapport 1 en de bijhorende onderzoeksvragen

	Module 1.2	Module 1.3		Module 1.4
Methodologie	Gevalstudies	Diepte-interviews		Survey
Respondenten	10 schoolleiders en leerkrachtenduo's	12 doorgelichte schoolleiders	12 niet doorgelichte schoolleiders	612 schoolleiders en leerkrachten
Onderzoeksvragen:				
Hoe vertrouwdheid is het werkveld met het OK-kader?		X	X	X
Welke zijn de effecten van het OK-kader?		X	X	X
Welke (f)actoren verklaren de effecten van het OK-kader?		X	X	X
Welke zijn de effecten van Inspectie 2.0?	X	X		X
Welke (f)actoren verklaren de effecten van Inspectie 2.0?	X	X		X

Een verduidelijking van de belangrijkste theoretische concepten gehanteerd in deze conclusie is terug te vinden in Tabel 2. Het volledige conceptueel kader bij dit deel van het onderzoek is terug te vinden in Module 1.1.

Tabel 2

Overzicht van de belangrijkste theoretische concepten

Concept	Omschrijving
Conceptueel effect	De invloed van de doorlichting op 'opvattingen over' en 'inzicht in'.
Instrumenteel effect	Feitelijk genomen beslissingen als een resultaat van de doorlichting en de acties die op deze beslissingen gebaseerd zijn.
Strategisch effect	Gebruk van de resultaten uit het doorlichtingsverslag vanuit een verantwoordingsperspectief (bv. promoten van de school).
Symbolisch effect	Gebruk van de resultaten uit het doorlichtingsverslag met onderliggende redenen (bv. legitimerend of overtuigend van aard).
Procedurele rechtvaardigheid	De eerlijkheid van het besluitvormingsproces waarin informatie verzameld werd om het doorlichtingsresultaat te bepalen.
Interactionele rechtvaardigheid	De percepties van de medewerkers in de doorgelichte scholen van de eerlijkheid en interpersoonlijke behandeling die ze krijgen van de inspecteurs.
Distributieve rechtvaardigheid	De eerlijkheid van de uitkomsten van de doorlichting of gevolgen van de inspectiebeslissing.

4.1.1 Over processen en effecten van Inspectie 2.0

Op basis van de onderzoeksresultaten van de verschillende modules stellen we vast dat Inspectie 2.0 eerder leidt tot instrumentele dan conceptuele effecten. Strategische effecten en misleidende praktijken om inspecteurs op het verkeerde been te zetten blijken eerder beperkt. Symbolische effecten, waarbij het doorlichtingsverslag veelal fungeert als 'stok achter de deur', worden vooral door schoolleiders erkend. Daarnaast percipiëren respondenten een doorlichting nog steeds als een periode die stress en spanning met zich meebrengt, alsook de nodige planlast. Voornaamste (f)actoren die statistisch voorspellen dat er met de resultaten uit een doorlichtingsverslag aan de slag gegaan wordt,

zijn de professionaliteit en geloofwaardigheid van de betrokken inspecteur(s), de mate van onderwijskundig leiderschap van de schoolleider en het innovatief vermogen van het schoolteam. De (f)actoren die het ondernemen van actie het meest in de weg staan zijn volgens de resultaten van dit onderzoek een gebrek aan continuïteit in samenstelling van de schoolteams - mede door het lerarentekort, de nasleep van de coronapandemie en een gebrek aan (gepaste) ondersteuning. In wat volgt gaan we dieper in op de effecten die een doorlichting met zich teweegbrengt en de belangrijkste randvoorwaarden die ervoor zorgen dat een doorlichting een hefboom voor onderwijskwaliteit kan zijn.

4.1.1.1 Welke zijn effecten van de doorlichting in scholen?

We stellen vast dat conceptuele effecten zich initieel vooral voordoen op korte termijn naar aanleiding van de aankondiging van de doorlichting, tijdens en meteen na afloop van de doorlichting. We onthouden dat deze eerder algemene conceptuele effecten vooral leiden tot een reflectie op de klas- en schoolwerking. In tegenstelling tot eerdere bevindingen waarbij volgens Penninckx (2015) de vorige doorlichtingsaanpak wel reflectie bevorderde, maar niet resulteerde in inzichten in de sterktes en zwaktes van de school, stellen we vast dat Inspectie 2.0 volgens de respondenten leidt tot inzichten in zowel de eigen sterktes en zwaktes als die van de school. Enkele schoolleiders spreken ook over een niet te onderschatten gevoel van 'bevestiging' na een doorlichting. In mindere mate geven respondenten aan dankzij de doorlichting stil te staan bij de kwaliteit van het aangeboden onderwijs en slechts een enkele keer is er ook sprake van anders denken inzake leerlingenbegeleiding of leerlingevaluatie.

De instrumentele effecten die zich voordoen naar aanleiding van de doorlichting zijn zeer uiteenlopend, doch meestal in relatie tot de werkpunten uit het doorlichtingsverslag. In elke school die deelnam aan de gevalstudies en diepte-interviews werd volgens de respondenten actie ondernomen op basis van het doorlichtingsverslag. De resultaten van het survey-onderzoek bevestigen dat op het merendeel van de scholen actie ondernomen wordt naar aanleiding van een doorlichting. Daarnaast werd volgens sommige respondenten ook actie ondernomen op basis van de mondelinge feedback die tijdens het synthesesgesprek werd gegeven. Hoewel instrumentele effecten van een schooldoorlichting zich in internationale context eerder op school- dan op klasniveau situeren, treden we in dit onderzoek de bevindingen van Penninckx (2015) bij die stelt dat een schooldoorlichting in Vlaamse context ook tot klasgerelateerde instrumentele effecten leidt. Zo stellen we bijvoorbeeld vast dat de meeste leerkrachten van de scholen uit de gevalstudies aan de slag gingen met de feedback die zij geven aan leerlingen en gaven enkele schoolleiders aan in te zetten op het creëren van krachtige leeromgevingen door de zorgwerking door te voeren tot op de klasvloer. Op schoolniveau leidt Inspectie 2.0 tot onder andere proactieve en preventieve acties die gepland en geïmplementeerd worden op vlak van psychisch en sociaal functioneren, onderwijsloopbaan en leren en studeren. Zo is er onder meer sprake van het opstellen van een talentenplan of pestactieplan, maar ook het uitschrijven en implementeren van een volledig nieuw zorgbeleid wordt genoemd als een instrumenteel effect van Inspectie 2.0. Dat volgens eerder onderzoek de invloed van een doorlichting snel verweekeld geraakt met andere externe en interne invloeden op de school en het bijgevolg een methodologische uitdaging is om de precieze effecten van een doorlichting na te gaan, blijkt ook hier uit het feit dat sommige respondenten er niet helemaal van overtuigd zijn dat alle ondernomen acties waarover sprake is in de interviews het gevolg zijn van de doorlichting (De Maeyer, Penninckx, Vanhoof, & Van Petegem, 2014; Ehren, Perryman, & Shackleton, 2015). Daar waar de respondenten bepaalde aspecten van de schoolwerking zelf als werkpunt opgaven bij de doorlichting, zijn ze ervan overtuigd dat deze acties ook ondernomen geweest zouden zijn zonder de komst van de doorlichting. Vaak erkennen respondenten wel dat de komst van de doorlichting het implementeren van acties als deze versneld heeft. Ondanks de uitdagingen die de

afgelopen jaren (en nog steeds) niet weg te denken waren uit de Vlaamse onderwijscontext, stellen we toch vast dat bij het merendeel van de respondenten Inspectie 2.0 leidt tot structurele (langetermijn)effecten. We sluiten hierbij aan bij de bevindingen van Elmore en Fhurman (2001) die stellen dat scholen die geconfronteerd worden met een 'high-stake' systeem van doorlichten, meestal opteren voor een 'quick fix'. Inspectie 2.0 stelt kwaliteitsontwikkeling en kwaliteitszorg echter centraal en koppelt relatief weinig gevolgen aan een gunstig of ongunstig advies. Een doorlichting volgens de principes van Inspectie 2.0 beoogt bovendien om een schoolteam vertrouwen te geven en vertrekt van het uitgangspunt dat de school de eerste verantwoordelijke is voor haar kwaliteit. Scholen verbeteren hun onderwijskwaliteit bijgevolg naar eigen zeggen vanuit een eerder autonome motivatie, zij het door de onderwijsinspectie gestimuleerd.

Op gebied van leerlingenbegeleiding stellen we vooral instrumentele effecten vast bij de respondenten die doorgelicht werden in schooljaar 2018-2019. We merken hierbij op dat 'leerlingenbegeleiding' ten tijde van hun doorlichting nog een afzonderlijk kwaliteitsgebied was dat bovendien bij alle deelnemende scholen van het gevalstudie-onderzoek deel uitmaakte van de doorlichtingsfocus. De doorlichting leidde in deze scholen onder meer tot een verhoging van de ouderbetrokkenheid, ondersteuning van leraren - meer bepaald op vlak van inzetten op preventieve basiszorg - en het plannen van proactieve en preventieve acties op vlak van psychisch en sociaal functioneren van leerlingen. Om zicht te krijgen op de specifieke onderwijsbehoeften van leerlingen ligt er voor veel respondenten nog een uitdaging weggelegd in het invoeren of optimaliseren van hun leerlingvolgsysteem.

Hoewel 'evaluatie van de leerling' bij een doorlichting als afzonderlijk aspect onderzocht wordt bij het onderzoek naar de onderwijsleerpraktijk, namen we aangaande 'leerlingevaluatie' in dit onderzoek ook effecten mee die zich situeren binnen de ontwikkelingsschalen 'effectieve feedback', 'afstemming van het aanbod op het gevalideerd doelenkader' en 'rapportering'. Zoals eerder vermeld werden in meerdere scholen acties ondernomen aangaande het feedbackbeleid. Zo wordt in sommige scholen, naar aanleiding van de doorlichting, volgens onze respondenten feedback aan leerlingen nu doel- en ontwikkelingsgericht geformuleerd. In andere scholen bracht de komst van de doorlichting dan weer meer aandacht voor procesevaluatie of het herdenken van de inhoud en vorm van leerlingenrapporten, met een streven naar een evenwichtige verdeling tussen punten-en doelenrapporten en meer aandacht voor zelfreflectie. Effecten op gebied van leerlingevaluatie blijken volgens de respondenten meestal per definitie langetermijneffecten aangezien aanpassingen aan een evaluatie-, rapporterings- of feedbackbeleid volgens de respondenten vaak enige tijd nodig hebben om in de schoolwerking geïmplementeerd te geraken.

Datzelfde geldt voor effecten op vlak van kwaliteitsontwikkeling. Voor het merendeel van de respondenten leidde Inspectie 2.0 tot aanpassingen op gebied van onderwijskundig beleid, zoals het realiseren van een doeltreffend taalbeleid of het inzetten op professionalisering van de leerkrachten. De meeste bevroegde schoolleiders geven daarnaast ook aan om op basis van de aanbevelingen uit het doorlichtingsverslag actie te willen ondernemen om de onderwijskwaliteit nog meer op systematische en betrouwbare wijze te willen evalueren. Er heersen wel degelijk ondersteuningsnoden om zich daarbij te kunnen baseren op diverse kwalitatieve en kwantitatieve bronnen en daarnaast ook aandacht te besteden aan de resultaten en effecten bij leerlingen. We stellen vast dat de komst van de doorlichting, vooral in de scholen doorgelicht in 2018-2019, ook een effect uitoefende op het werken met het OK-kader. Mogelijke verklaring hiervoor is dat deze scholen bij de start van Inspectie 2.0 als één van de eerste scholen in contact kwamen met het toen nog nieuwe Referentiekader Onderwijskwaliteit. Veel van deze schoolleiders geven dan ook aan het OK-kader te hebben leren kennen door de komst van de

doorlichting. Deze schoolleiders geven vaak aan dat ze op termijn hoogstwaarschijnlijk met het kader aan de slag gegaan zouden zijn, maar dat de komst van de doorlichting dit versneld heeft. In het merendeel van de bevroegde scholen is dit echter een korte termijneffect en is de aandacht voor het kader na afloop van de doorlichting ook weer verslapt. Reden hiervoor is volgens de schoolleiders dat er momenteel andere prioriteiten zijn en dat het vaak ontbreekt aan tijd om het kader eigen te maken. Er is naar eigen zeggen vooral sprake van een zeker strategisch effect van het OK-kader zelf, waarbij de respondenten werken met het OK-kader vaak zien als een goede voorbereiding op een volgende doorlichting. Hierop komen we in het tweede deel van de conclusie nog terug.

We stellen in mindere mate strategische en symbolische effecten vast naar aanleiding van de komst van de doorlichting. Deze bevinding ligt in lijn met eerdere onderzoeksresultaten, waarbij ook Penninckx (2015) concludeerde dat symbolische en strategische effecten als gevolg van de vorige doorlichtingsaanpak eerder klein bleken. Hoewel Kelchtermans (2007) stelt dat ook leerkrachten van het symbolisch effect van een doorlichting gebruik maken om hun persoonlijke doelen na te streven, besluiten we op basis van deze onderzoeksresultaten dat het doorlichtingsverslag vooral door de schoolleiders gebruikt wordt als 'een stok achter de deur' om bepaalde teamleden aan te zetten tot actie of om bepaalde acties (versneld) geïmplementeerd te krijgen. De resultaten uit het surveyonderzoek bevestigen dit.

Wat ongewenste neveneffecten van een doorlichting betreft toont internationaal onderzoek vooral window dressing activiteiten als neveneffect, waarbij leden van het schoolteam zich anders voordoen dan ze werkelijk zijn, of de school afspraken maakt over wat leraren wel mogen vertellen en waarover ze moeten zwijgen tijdens een doorlichting (Brimblecombe & Ormston, 1995). In de Vlaamse context en op basis van de vorige doorlichtingsaanpak stellen Penninckx et al. (2015) slechts een beperkte mate van onwenselijke strategische voorbereiding vast. Indien dit toch het geval was, betrof het vooral het vervaardigen van extra en nieuwe documenten. Op basis van de gesprekken met schoolleiders en leerkrachten en de afgenomen survey stellen we ook in dit onderzoek vast dat misleidende praktijken zoals window dressing of het fabriceren van decor- of documenten om inspecteurs bewust op het verkeerde been te zetten zeer beperkt blijven. De surveyresultaten treden in dit geval eerder internationale bevindingen bij, met slechts in zeer beperkte mate praktijken als window dressing, waarbij binnen het team afspraken gemaakt werden over wat wel en niet te vertellen tijdens de doorlichting. Op basis van de interviews voegen we daar nog aan toe dat respondenten vooral de rondleiding door de leerlingen op een bepaalde manier proberen te sturen, waarbij meestal benadrukt werd welke route gevolgd moest worden en wat zeker wél aan bod moest komen tijdens de rondgang. Toch zijn ook deze vormen van neveneffecten eerder beperkt en gaat het meestal over het kiezen van zelfzekere en mondige leerlingen en het benoemen van reeds aanwezige praktijken, eerder dan misleidende praktijken. Zowel de bevroegde schoolleiders als de leerkrachten staan er naar eigen zeggen op een correct en volledig beeld te tonen van de school(werking). Ze zijn er immers van overtuigd dat de inspecteurs over genoeg expertise beschikken om deze vormen van misleiding te doorprikken.

Een doorlichting brengt nog steeds heel wat gevoelens teweeg in het werkveld. Zo spreken de meeste respondenten van een zekere mate van stress en spanning ten gevolge van de doorlichting. Dat een doorlichting nog steeds een stressvolle periode is, blijkt ook uit het surveyonderzoek. De respondenten van de gevalstudies en diepte-interviews nuanceren dit echter door toe te lichten vooral stress te ervaren bij aankondiging en aanvang van de doorlichting. Deze spanning verdwijnt bij de meeste respondenten naargelang de doorlichting vordert. Na afloop van de doorlichting is er eerder sprake van opluchting en een gevoel van bevestiging. Dit weerspiegelt zich in de surveyresultaten waarbij de

respondenten uit de doorgelichte scholen aangeven iets minder stress te ervaren binnen het team ten opzichte van de respondenten uit de niet doorgelichte scholen, hoewel het verschil eerder klein is. In een enkele school is er ook sprake van een zekere teleurstelling, vooral omdat de doorlichting aanvankelijk hoop bracht, maar op korte termijn niet leidde tot de gewenste verandering binnen de schoolwerking. Daarnaast geven enkele schoolleiders van scholen met ongunstig advies ook aan dat de doorlichting na afloop zorgde voor een zekere verdeeldheid binnen het team en het uiten van beschuldigingen naar bepaalde teamleden waarvan ze wisten dat deze verantwoordelijk waren voor bepaalde tekortkomingen in het doorlichtingsverslag.

4.1.1.2 Welke (f)actoren spelen een rol in het verklaren van effecten van de doorlichting in scholen?

De uitgangspunten van Inspectie 2.0 worden gewaardeerd

We stellen vast dat bij het merendeel van de bevroegde respondenten vooral positieve signalen weerklinken aangaande inspectie 2.0, vooral wat betreft de interactionele rechtvaardigheid, het professionalisme en de geloofwaardigheid van de inspecteur en de kenmerken van de mondelinge feedback. Respondenten geven vaker aan akkoord te gaan met de ontvangen feedback en zich te kunnen vinden in het bijhorend eindoordeel, wanneer ze het gevoel hebben dat de doorlichting op een eerlijke, betrouwbare en transparante manier verlopen is. We treden hier de bevindingen bij van Quintelier (2020) die stelt dat een inspectieproces dat ervaren wordt als valide en betrouwbaar leidt tot een hogere mate van feedbackacceptatie. Kenmerkend voor Inspectie 2.0 zijn de gesprekken met ouders, leerkrachten en leerlingen, die volgens de schoolleiders naar eigen zeggen mee bijdragen tot een hoge interactionele rechtvaardigheid en een positieve beleving van het inspectieproces. Daar waar de respondenten toch minder fijne ervaringen linken aan de doorlichting is dit vooral te wijten aan een lage procedurele en distributieve rechtvaardigheid. Vanuit de surveyresultaten lijkt een ongunstig doorlichtingsadvies vaak samen te gaan met een gepercipieerde lage procedurele en distributieve rechtvaardigheid. Vanuit de interviews voegen we daar het gebruik van datatriangulatie om tot een eindoordeel te komen aan toe, wat volgens de respondenten niet altijd ten volle gerealiseerd is. Dit gebruik van datatriangulatie is een kenmerk van Inspectie 2.0 dat door de respondenten als positief beschouwd wordt en dat volgens hen de norm zou moeten zijn. Toch zijn enkele bevroegde schoolleiders van mening dat het passende evenwicht tussen de verschillende databronnen naar hun gevoel niet altijd gegarandeerd was. In enkele scholen legde volgens de directeurs de desk- en documentstudie te veel gewicht in de schaal. Volgens anderen klonk de stem van enkele teamleden dan weer te hard door in het eindoordeel. Tenslotte lag voor enkele schoolleiders ook een ongunstig advies ten gevolge van werkpunten op vlak van woonbaarheid, veiligheid en hygiëne mede aan de basis van een lage distributieve rechtvaardigheid omdat ze van mening waren dat zij hier niet altijd de eindverantwoordelijkheid voor konden dragen.

Inspecteurs worden als professioneel omschreven

De betrokken inspecteurs worden over het algemeen ervaren als zijnde professioneel en betrouwbaar. Omschrijvingen aangaande de betrokken inspecteurs als 'menselijk', 'een luisterend oor' en 'open voor feedback' komen geregeld ter sprake. Respondenten geven ook aan dat dit bijdraagt tot een hoge gepercipieerde interactionele rechtvaardigheid en een invloed uitoefent op de algehele beleving van de doorlichting. Ehren en Visscher (2008) stelden vast dat een positieve relatie tussen de inspecteurs en de leerkrachten de acceptatie van de feedback verhoogt als voorloper van verdere verbetering. Allen en Burgess (2012) erkennen daarnaast ook het belang van de professionaliteit van de betrokken inspecteurs om duurzame en kwalitatieve gewenste doorlichtingseffecten teweeg te brengen. Ook in de Vlaamse context lijken deze stellingen op te gaan, wanneer we op basis van de surveyresultaten

vaststellen dat de geloofwaardigheid en professionaliteit van de betrokken inspecteurs wel degelijk statistisch significante voorspellers zijn voor de instrumentele effecten van een schooldoorlichting.

Inspectiefeedback: verba volant, scripta manent

Vanuit de interviews blijkt de ontvangen feedback een belangrijk kenmerk in het verklaren van effecten van een doorlichting. De ontvangen mondelinge feedback doorheen de doorlichtingsweek is volgens de respondenten doorgaans helder en constructief, maar niet altijd relevant vanwege de gepercipieerde (on)haalbaarheid van de feedback. Wat betreft de schriftelijke feedback die we terugvinden in het eindverslag zijn de meeste respondenten van oordeel dat deze eerder complex en summier is. Audia en Locke (2003) stellen vast dat moeilijke woordenschat en globale beschrijvingen van tekortkomingen kunnen resulteren in slecht geïnformeerde feedbackontvangers. In het verlengde hiervan is de mondelinge toelichting bij het doorlichtingsverslag en bijhorend eindoordeel volgens de respondenten dan ook een grote meerwaarde. In de uitgangspunten van inspectie 2.0 hanteert de onderwijsinspectie de begrippen 'controlerende' en 'stimulerende' rol. Wat het 'stimulerende' aspect betreft, duidt ze daarbij expliciet op het verschil tussen stimuleren en begeleiden. Inspecteurs in het Vlaams onderwijsveld worden niet geacht om scholen advies te geven over hoe ze hun huidige praktijken kunnen verbeteren, deze taak is weggelegd voor de pedagogische begeleidingsdienst. De onderwijsinspectie analyseert en rapporteert over de sterktes en zwaktes van de scholen, maar mag niet overgaan tot het formuleren van aanbevelingen om blootgelegde zwaktes te counteren (Onderwijsinspectie 'Het concept Inspectie 2.0', 2021). Volgens de respondenten van de gevalstudies en diepte-interviews bevat de mondelinge feedback die tijdens het synthesegesprek wordt gegeven ter verduidelijking van het doorlichtingsverslag adviezen en aanbevelingen om met werkpunten aan de slag te gaan. Veel van deze informatie gaat naar eigen zeggen echter vaak verloren door bijvoorbeeld de afwezigheid van een notulist of het niet mogen maken van opnames. Dit oefende in sommige gevallen een rechtstreekse invloed uit op verdere effecten van de doorlichting.

Het schoolteam en hun schoolleider als sleutelfiguren

Persoonskenmerken zoals feedbacknut van leraren en schoolleiders blijken minder relevant te zijn bij het voorspellen van instrumentele effecten van een doorlichting, doch wel een zekere invloed uit te oefenen op de conceptuele effecten ervan. We treden Anseel en Lievens (2006) hierbij deels bij door te stellen dat op basis van onze onderzoeksresultaten iemands globale houding ten aanzien van feedback het denken over bepaalde aspecten substantieel beïnvloedt. Om effectief actie te ondernemen op basis van het doorlichtingsverslag zijn het volgens de respondenten eerder de kenmerken van het schoolteam die een rol spelen. Innovatief en reflectief vermogen, de aanwezigheid van ondersteunende professionele relaties en gedeeld leiderschap blijken op basis van de onderzoeksresultaten een positieve impact te hebben op de doorlichtingseffecten. Door het merendeel van de respondenten wordt daar nog gezamenlijke doelgerichtheid aan toegevoegd. Volgens de schoolleiders zorgen gedeeld leiderschap en gezamenlijke doelgerichtheid ervoor dat een doorlichting een gedeelde verantwoordelijkheid wordt en zo het gevoel van 'samen school maken' versterkt. Ze erkennen daarbij ook het belang van vertrouwen binnen het team. Dit gevoel van gedeelde verantwoordelijkheid bevordert volgens hen na afloop van de doorlichting de wil om met de doorlichtingsresultaten aan de slag te gaan. Daar waar bepaalde acties weer stopgezet werden, gaat het dan ook vaak om acties die niet gedragen werden door het hele team.

Dat de schoolleider een belangrijke rol speelt in een doorlichtingsproces wordt in meerdere onderzoeken en ook hier bevestigd (Matthews & Sammons, 2004; Sun, Creemers, & De Jong, 2007). Net zoals King (2011) reeds stelde, oefent de houding van de schoolleider ten aanzien van de doorlichting een grote invloed uit op het enthousiasme en de moraal van het schoolteam. Bekijken we de verschillende leiderschapstijlen van de bevraagde schoolleiders, stellen we vast dat op basis van onze

onderzoeksresultaten in een school met een sterk onderwijskundig beleid met een schoolleider die meer de kenmerken vertoont van onderwijskundig leiderschap de kans op instrumentele effecten van een doorlichting groter is.

Factoren die volgens de respondenten maken dat de doorlichting niet tot waardevolle effecten leidde zijn onder meer directiewissels, de uitval van teamleden en het niet vinden van een passende vervanging. In dezelfde lijn draagt ook het lerarenkort niet bij tot een optimale gang van zaken om met werkpunten uit een doorlichtingsverslag aan de slag te gaan. Daarnaast stond de algemene schoolwerking op de deelnemende scholen gedurende een tweetal schooljaren min of meer 'on hold' omwille van de coronapandemie. Een niet te onderschatten belemmerende factor, die ervoor zorgde dat het aanpakken van de werkpunten uit het doorlichtingsverslag gedurende deze tijd geen prioriteit was. Andere prioriteiten kregen de bovenhand en de nadruk lag meer op initiatieven op korte termijn en de organisatorische aspecten van de schoolwerking. In de bevroegde scholen gaven de schoolleiders vaak wel aan van plan te zijn om de draad nadien weer op te pikken en alsnog aan de slag te gaan met bepaalde werkpunten uit hun doorlichtingsverslag om zo toch te werken aan de eerder genoemde structurele langetermijneffecten.

Ondersteuning werkt, in een verscheidenheid van vormen, ...

Tenslotte blijkt ook de aangeboden ondersteuning en de mogelijk gemiste ondersteuning essentieel in het doorlichtingsverhaal. Door het merendeel van de respondenten wordt vooral steun gevonden in het team en bij de schoolleider. Scholen waar volgens de respondenten minder sprake is van ondersteunende professionele relaties vertonen naar eigen zeggen meer moeilijkheden om gericht met de doorlichtingsresultaten aan de slag te gaan. Op basis van de interviews concluderen we dat de schoolleider een zeer bepalende rol speelt in het doorlichtingsverhaal, niet alleen als ondersteuner en begeleider, maar ook als initiatiefnemer om na afloop van de doorlichting effectief over te gaan tot actie. Door de schoolleiders, maar ook in mindere mate door de leerkrachten worden ook de betrokken inspecteurs als een meerwaarde genoemd op gebied van ondersteuning. Een belangrijke vorm van ondersteuning volgens de schoolleiders is bijvoorbeeld het online voorbereidend gesprek met de teamcoördinator met daarbij de mogelijkheid tot het stellen van vragen. De schoolleiders zelf erkennen daarnaast ook collega-directeurs of coördinerend directeurs als bron van ondersteuning voor, tijdens en na de doorlichting. Wanneer gevraagd wordt naar bepaalde ondersteuningsnoden benoemen bijna alle schoolleiders 'nood aan tijd' en 'nood aan leerkrachten' als grootste ondersteuningsnoden. Zoals eerder vermeld zijn ook veel schoolleiders nog zoekende om op systematische en betrouwbare wijze hun onderwijskwaliteit te evalueren en zich daarbij te baseren op diverse kwalitatieve en kwantitatieve bronnen. Daarnaast blijven sommige schoolleiders met enkele zeer concrete ondersteuningsvragen zitten zoals bijvoorbeeld het gepast in kaart brengen van (de studievoortgang van) leerlingen of effecten van leerlingbegeleiding, waarbij ze niet weten waar ze nog ten rade kunnen gaan, wat maakt dat bepaalde initiatieven zoals het vernieuwen van een leerlingvolgsysteem bijgevolg stopgezet worden of gaandeweg uitdoven.

De meningen over de (kwaliteit van de) ondersteuning aangeboden door de pedagogische begeleidingsdiensten zijn zeer uiteenlopend. We stellen vast dat de ondersteuning van de pedagogische begeleidingsdiensten volgens de bevroegde respondenten vooral na afloop van de doorlichting plaatsvindt en slechts in mindere mate ter voorbereiding op of tijdens de doorlichting. Dit ligt in lijn met de vooropgestelde bewuste keuzes van de PBD om ter voorbereiding op een doorlichting niet intens te begeleiden, teneinde neveneffecten als *window dressing* te vermijden. De respondenten van de gevalstudies zijn over het algemeen minder tevreden over de mate en kwaliteit van ondersteuning

vergeleken met de respondenten van de diepte-interviews. Mogelijke verklaring hiervoor is dat de scholen uit de gevalstudies bij de eerste scholen waren die volgens de nieuwe doorlichtingsaanpak werden doorgelicht. Na afloop van de doorlichting speelt de pedagogische begeleidingsdienst vaak wel een essentiële rol in het opvolgen van werkpunten. Enkele bevroegde leerkrachten uit de gevalstudies percipiëren deze vorm van ondersteuning echter vaak als weinig relevant en de oplossingen die worden aangereikt als niet haalbaar in de betreffende school- en/of klascontext. De schoolleiders zijn vaak wel overwegend positief. Zo zijn velen voorstander van de aanwezigheid van de pedagogisch begeleider bij het synthesegesprek. De resultaten uit het surveyonderzoek bevestigen dat de schoolleiders net iets meer ondersteuning ervaren vanwege de pedagogische begeleidingsdiensten dan de leerkrachten, hoewel beide groepen respondenten eerder neutraal blijven.

... maar kent tevens bezorgdheden

Over het algemeen zijn respondenten te vinden voor de vooropgestelde doorlichtingstermijn van zes jaar. De scholen met een gunstig advies en slechts enkele of geen werkpunten in hun doorlichtingsverslag uiten echter een zekere bezorgdheid wat betreft deze vooropgestelde doorlichtingstermijn. Zo wordt in de interviews door het merendeel van de respondenten voldoende aandacht gevraagd voor de begeleiding van het schoolteam na afloop van een doorlichting. Vele respondenten zijn vragende partij om mede door de onderwijsinspectie sneller, maar meer vrijblijvend opgevolgd te worden, waarbij ondersteuning ook de vorm aanneemt van een procesgericht traject voorzien van tussentijdse feedback. Hoewel deze opvolging een mogelijke taak is van de pedagogische begeleidingsdienst, zien scholen dit toch ook weggelegd als een taak voor de onderwijsinspectie omdat er naar eigen zeggen nog 'onvoldoende afstemming' is tussen beide onderwijspartners.

4.1.2 Over het Referentiekader Onderwijskwaliteit

Over het algemeen zijn de respondenten slechts matig vertrouwd met het OK-kader en beschikken eerder enkel over feitenkennis dan over conceptuele- en procedurele kennis en vaardigheden om met het kader aan de slag te gaan. Het merendeel van de respondenten is zich wel bewust van het bestaan van het Referentiekader. Op basis van de onderzoeksresultaten stellen we vast dat schoolleiders beduidend meer vertrouwd zijn met het kader dan leerkrachten. De interviews leren ons echter dat vertrouwd zijn met het OK-kader niet per definitie impliceert dat het OK-kader ook een weg gevonden heeft binnen de schoolwerking. Of een school al dan niet werd doorgelicht volgens de principes van inspectie 2.0 blijkt niet van tel bij het verklaren van de mate van vertrouwdheid met het OK-kader. De diepte-interviews leren ons dat op korte termijn aan de slag gaan met het OK-kader naar aanleiding van de doorlichting niet impliceert dat er ook op lange termijn sprake is van een zekere vertrouwdheid met het kader. De onderwijскоepel waartoe de respondent behoort blijkt dat wel: respondenten van OVSG voelen zich naar eigen zeggen meer vertrouwd met het OK-kader dan respondenten van KOV en GO!.

4.1.2.1 Hoe vertrouwd is het werkveld met het OK-kader?

Hoe en wanneer het kader tot stand kwam is bij het merendeel van de respondenten geen parate kennis. Respondenten geven aan dat ze zich bewust zijn van het bestaan van het kader en dat ze weten waar het OK-kader te kunnen raadplegen. De begrippen uit het OK-kader zijn vooral gekend door de schoolleiders. Ook wat de conceptuele en procedurele kennis betreft en de nodige vaardigheden om met het OK-kader aan de slag te gaan, zijn het vooral de schoolleiders die aangeven hierover te beschikken. Hoewel ook hier minder dan de helft van deze respondentengroep zichzelf naar eigen zeggen in staat vindt de werking van het OK-kader aan een collega uit te leggen.

Affectieve attitudes in de zin van het OK-kader dat angst inboezemt of doet twijfelen aan het eigen kunnen zijn zeer beperkt. Volgens zowel de survey als de diepte-interviews geeft een derde van de respondenten wel aan een zekere druk van bovenaf te ervaren om met het OK-kader aan de slag te gaan. Op basis van de interviews leren we bijvoorbeeld dat het dan gaat het om gepercipieerde druk vanuit het schoolbestuur of vanuit de onderwijsinspectie. Bekijken we de aanwezige cognitieve attitudes, merken we op dat het merendeel van de schoolleiders uit de diepte-interviews het OK-kader percipieert als weinig helder en weinig functioneel. Toch omschrijven diezelfde schoolleiders het OK-kader ook als zinvol en relevant dat een meerwaarde zou kunnen vormen voor de schoolwerking. De resultaten van het surveyonderzoek bevestigen dit. Op basis van de interviews stellen we vast dat een deel van de schoolleiders die het kader percipieert als zinvol en relevant bijgevolg zelf aan de slag ging om een vertaalslag van het OK-kader te maken om zo het gebrek aan helderheid en functionaliteit te compenseren.

4.1.2.2 Welke zijn de effecten van de invoering van het OK-kader?

Hoewel er in se een duidelijk onderscheid is en kan gemaakt worden tussen het OK-kader en het toezichtkader (TOK-kader, gebaseerd op het OK-kader en gehanteerd door de onderwijsinspectie), blijkt op basis van de interviews dat er in de praktijk toch een zekere verwarring wordt gepercipieerd. Dit maakt het niet evident om uitspraken te doen over de specifieke en unieke impact van het OK-kader, aangezien vele ondernomen acties vaak ook voortkomen uit het inschalen van de eigen onderwijskwaliteit op basis van de ontwikkelingsschalen uit het TOK. Naast het TOK gingen enkele van de bevragede schoolleiders ook aan de slag met een eigen gemaakte vertaalslag van het OK-kader of alternatieve kaders, (deels) gebaseerd op het OK-kader, aangereikt door de schoolgemeenschap of onderwijskoepel. Daarnaast blijkt uit de interviews dat de invoering van het OK-kader bij enkele schoolleiders op korte termijn tot bepaalde conceptuele of instrumentele effecten leidde, maar dat op lange termijn het OK-kader toch (soms letterlijk) weer in de kast belandde. Op basis van deze bevindingen stellen we vast dat het originele OK-kader slechts in beperkte mate een weg gevonden heeft naar de schoolwerking, wat niet impliceert dat ook de ideeën uit het OK-kader geen weg vonden naar het werkveld. Wat wel zeker is, is dat het effect van het OK-kader zicht hoofdzakelijk manifesteert op beleidsniveau en slechts in beperkte mate bij de leerkrachten op de klasvloer.

Op het eerste zicht lijken de resultaten van het surveyonderzoek en de bevindingen uit de diepte-interviews elkaar wat tegen te spreken. Daar waar volgens de schoolleiders uit de interviews werken met het OK-kader vooral tot conceptuele effecten leidt, wijzen de resultaten van het surveyonderzoek vooral op instrumentele effecten van het OK-kader. Bekijken we de verschillen tussen de schoolleiders en leerkrachten, lijkt de stelling op te gaan dat het OK-kader vooral bij die eerste groep toch ook al een verandering in het denken teweegbracht. Het OK-kader zet hen onder meer aan tot reflectie over het schoolbeleid en helpt hen te borgen wat goed loopt. Dat blijkt ook uit de interviews waarin de schoolleiders spreken over 'een vollediger beeld van wat kwaliteitsvol onderwijs inhoudt'. Het OK-kader biedt volgens hen structuur en richting en legt waar nodig blinde vlekken bloot. Het zorgt er volgens de schoolleiders voor dat ze zich weer bewust worden van wat hun taak als schoolleider inhoudt en met welke aspecten van de schoolwerking ze best rekening houden. Het biedt voor sommigen ook inzichten in de eigen sterktes en zwaktes. Daarnaast spreken sommigen ook over een gevoel van geruststelling en een zeker houvast, waarbij het OK-kader voor hen welgekomen sturing en houvast geeft, al dan niet naar aanleiding van een doorlichting. Het OK-kader creëert vooral voor de schoolleiders een gemeenschappelijke taal om over onderwijskwaliteit te spreken. Om opnieuw de complexiteit van het OK-kaderverhaal te schetsen, merken we op dat deze conceptuele effecten zich niet enkel voordoen bij

de schoolleiders die vertrouwd zijn met het kader, maar ook bij enkele van de bevroegde schoolleiders die zich ooit vertrouwd noemden met het kader, maar bij wie het ondertussen op de achtergrond is beland.

Zowel wat conceptuele als instrumentele effecten betreft stellen we in de interviews vast dat werken met het kader vooral een invloed uitoefent op aspecten van kwaliteitsontwikkeling, eerder dan op gebied van leerlingenbegeleiding of leerlingevaluatie. Ook uit de surveyresultaten blijkt dat het OK-kader net iets vaker leidt tot conceptuele en instrumentele effecten op vlak van kwaliteitsontwikkeling, dan op vlak van leerlingevaluatie of leerlingenbegeleiding. Slechts anekdotisch vermeldt een schoolleider in een interview conceptuele effecten op vlak van leerlingenbegeleiding of leerlingevaluatie, dit vooral naar aanleiding van het werken met de ontwikkelingsschalen. In enkele scholen werd wel gereflecteerd over bepaalde aspecten van de schoolwerking op basis van het OK-kader of het TOK, maar vervolgens leidden deze conceptuele effecten niet tot instrumentele effecten en in enkele scholen verdween de output van deze reflecties in een kaft, vaak ter voorbereiding op een (mogelijke) doorlichting.

In de scholen waar de schoolleiders vertrouwd zijn met het OK-kader en waar het OK-kader een weg gevonden heeft in de schoolwerking en soms zelfs een stuk tot op de klasvloer, zien we dat de komst van het OK-kader tot enkele gevolgen geleid heeft die volgens de desbetreffende scholen wel degelijk een hefboom vormen voor onderwijskwaliteit. Zo is er bijvoorbeeld sprake van het stimuleren van de aanwezigheid van een kwaliteitscultuur en het inzetten op het kwaliteitsdenken bij het schoolteam. De mate waarin het kader dit effect teweegbrengt hangt echter nauw samen met de manier waarop het kader gehanteerd wordt binnen de schoolwerking. Op basis van de interviews stellen we vast dat enkele schoolleiders het kader immers hanteren als een referentiekader in de strikte zin van het woord, waarbij op personeelsvergaderingen bijvoorbeeld enkel de link gelegd wordt naar de desbetreffende (deel)rubriek(en) waar op dat moment mee aan de slag gegaan wordt. Enkele schoolleiders hanteren het OK-kader ook als een werkinstrument, waarbij het volgens de betreffende schoolleiders van belang is de nodige vaardigheden te ontwikkelen om niet enkel bepaalde aspecten van de schoolwerking binnen het kader te kunnen situeren. Gebruikers van het OK-kader dienen daarnaast ook over de nodige reflectievaardigheden te beschikken om de nodige acties te implementeren, te borgen of bij te sturen waar nodig om aan de desbetreffende kwaliteitsverwachtingen te voldoen.

Symbolische effecten van het OK-kader blijken in de bevroegde scholen zeer beperkt. Daar waar toch sprake is van een zeker symbolisch effect zijn het vooral de schoolleiders die aangeven het OK-kader te hanteren om het schoolteam aan te zetten tot actie, of om persoonlijke standpunten sterker te onderbouwen of gemaakte keuzes beter te verantwoorden. In de interviews komt wel een zeker strategisch gebruik van het OK-kader ter sprake, soms in combinatie met een symbolisch effect. De schoolleiders die de meerwaarde van het OK-kader inzien vinden het een bijkomstig voordeel dat ze het gevoel hebben ook 'een beter beeld van de school te kunnen schetsen' bij de onderwijsinspectie. Verschillende bevroegde schoolleiders die geen uitgesproken meerwaarde zien in het OK-kader geven expliciet aan het OK-kader nog eens uit de kast te zullen halen wanneer een doorlichting zich aankondigt, ook hier vanuit de overtuiging om bij de onderwijsinspectie op een beter blaadje te staan. Het merendeel van de bevroegde schoolleiders ziet er wel de meerwaarde van in om (al dan niet vanuit een strategisch perspectief) de leden uit hun schoolteam vertrouwd te maken met de terminologie uit het OK-kader, om als het ware een gemeenschappelijke taal te creëren om met de onderwijsinspecteurs in gesprek te kunnen gaan. Dit vormt voor vele bevroegde schoolleiders echter een uitdaging. Neveneffecten hangen op basis van de interviews nauw samen met de mate van vertrouwdheid van de

schoolleiders met het OK-kader. De bevroegde schoolleiders die in mindere mate vertrouwd zijn met het OK-kader omschrijven het kader eerder als ‘planlast’ en nemen begrippen in de mond als ‘nutteloze stress’ en ‘onzekerheid’.

4.1.2.3 Welke (f)actoren spelen een rol in het verklaren van de effecten van het OK-kader?

Op basis van de surveyresultaten stellen we vast dat zowel voor de conceptuele- als instrumentele effecten van het OK-kader de procedurele kennis en vaardigheden van de gebruiker alsook de percepties op de gebruiksvriendelijkheid en zinvolheid van het kader belangrijke voorspellers zijn. Voor de instrumentele effecten van het OK-kader komen daar nog de aangeboden ondersteuning bij en kenmerken van onderwijskundig leiderschap. Deze bevindingen liggen in lijn met de informatie die we vergaarden vanuit de interviews.

Dat de procedurele kennis en vaardigheden en de kenmerken van het kader belangrijke voorspellers zijn, wordt door het merendeel van de schoolleiders in de interviews erkend. Zij verduidelijken dat er een zekere wisselwerking is tussen beide. Het merendeel van deze schoolleiders vindt het OK-kader geen helder, noch een functioneel instrument, vervolgens splitst deze groep zich op in twee subgroepen. Enerzijds zijn er die schoolleiders die door gebrek aan helderheid en functionaliteit geen meerwaarde voor het OK-kader weggelegd zien in de schoolwerking. Bij hen verdwijnt het kader naar de achtergrond. Anderzijds is er een groot deel van de schoolleiders dat aangeeft het kader ondanks het gebrek aan helderheid en functionaliteit wel te percipiëren als ‘zinvol’ en een ‘meerwaarde’, waardoor ze zich op autodidactische wijze het OK-kader eigen maken, of het OK-kader verder concretiseren of er een vertaalslag van maken om het al dan niet tot bij de rest van het schoolteam te krijgen. We stellen vast dat de percepties van de schoolleider op de relevantie van het OK-kader hierbij doorslaggevend zijn.

Ook de aangeboden of gemiste ondersteuning blijkt een rol te spelen in het verklaren van de effecten van het OK-kader. We merken allereerst op dat leerkrachten zich minder ondersteund voelen om vertrouwd te geraken met het OK-kader dan schoolleiders. Het merendeel van de schoolleiders uit de interviews geeft aan dat de ondersteuning van de pedagogische begeleidingsdienst voor hen doorslaggevend was om met het OK-kader vertrouwd te geraken en er vervolgens mee aan de slag te gaan in de schoolwerking. We stellen vast dat de pedagogische begeleidingsdiensten inspanningen leverden om het OK-kader te concretiseren, bijgevolg spreken deze schoolleiders dan ook over de ondersteuning van de pedagogische begeleidingsdienst als zinvol en als een meerwaarde. Deze ondersteuning neemt vaak de vorm aan van introductiesessies en workshops waarbij gevraagd werd inschalingen te maken. Daarnaast noemt een groot deel van de schoolleiders ook het aanbieden van alternatieve kaders als een veelvoorkomende vorm van ondersteuning. Deze alternatieve kaders, zijn gebaseerd op het OK-kader, maar zijn volgens de bevroegde gebruikers vaak functioneler dan het OK-kader zelf. Dat ondersteuning ook een belemmerende factor kan zijn voor een succesvolle introductie van het OK-kader blijkt uit het feit dat in enkele scholen volgens de schoolleiders het OK-kader naar de achtergrond verdwenen is omwille van de verwarrende en vaak overweldigende wijze waarop het OK-kader door externe onderwijspartners bij hen geïntroduceerd werd. Naast de schoolleiders zijn het volgens de surveyresultaten vooral de respondenten uit recent doorgelichte scholen die zich gemiddeld genomen iets meer ondersteund voelen bij het werken met het OK-kader. Ook in de diepte-interviews wordt door enkele respondenten de betrokken inspecteur genoemd als bron van ondersteuning om zo bijvoorbeeld een duidelijker beeld te krijgen van wat er onder bepaalde kwaliteitsverwachtingen verstaan wordt. Respondenten uit niet doorgelichte scholen geven aan gemiddeld iets meer behoefte te hebben aan bijkomende ondersteuning. Tenslotte blijkt ook de onderwijskoepel waartoe de respondent ertoe te doen. Respondenten uit GO! en KOV voelen zich gemiddeld minder ondersteund

om met het OK-kader aan de slag te gaan dan respondenten uit OVSG, wat tevens in lijn ligt met de mate van vertrouwdsheid met het OK-kader.

Overige randvoorwaarden om volgens de bevroagde schoolleiders het OK-kader ook tot bij de rest van het schoolteam te krijgen zijn een sterk beleidsvoerend vermogen, met specifiek aandacht voor innovatief en reflectief vermogen en de aanwezigheid van een leerklimaat, meer bepaald het organiseren van collaboratieve leeractiviteiten. Toch zijn de aanwezigheid van een leerklimaat an sich en het organiseren van collaboratieve leeractiviteiten omtrent het OK-kader geen garantie om het OK-kader ingebed te krijgen in de schoolwerking. Bij de schoolleiders uit de steekproef die spreken van een school met een sterk beleidsvoerend vermogen én aanwezigheid van een leerklimaat zien we vaak op korte termijn wel een zekere vertrouwdsheid met het OK-kader, maar op lange termijn verdwijnt het OK-kader echter vaker wel dan niet weer naar de achtergrond. Scholen waar het kader toch een weg vond vertonen vaak een structurele inbedding van het OK-kader in de schoolwerking waarbij de schoolleiders naar eigen zeggen vooral inzetten op collaboratieve leeractiviteiten om dit te bewerkstelligen. Hoe dan ook, ook deze randvoorwaarden blijken enkel van tel wanneer aan de voorwaarde voldaan is dat de schoolleider zelf het kader percipieert als een zinvol kader dat een meerwaarde biedt voor de schoolwerking.

4.2 Bronnen van externe kwaliteitszorg als hefboom voor onderwijskwaliteit

Zoals eerder aangegeven, focust het tweede deelrapport van dit OBPWO-project op de verschillende bronnen van externe kwaliteitszorg, namelijk de doorlichtingen, het peilingsonderzoek en het internationaal vergelijkend onderzoek. In het eerste deel van deze conclusie worden de gelijkenissen en verschillen tussen de doorlichtingen, het peilingsonderzoek en het internationaal vergelijkend onderzoek besproken. In het tweede deel ligt de nadruk op de mate waarin de resultaten van de verschillende bronnen convergeren of divergeren. Het derde deel focust enerzijds op de mogelijke verklaringen voor de concordantie tussen de resultaten van de verschillende bronnen en anderzijds op hoe de resultaten van de verschillende bronnen van externe kwaliteitszorg gecombineerd kunnen worden. Tabel 3 biedt een overzicht van de verschillende onderzoeksvragen uit het tweede deelrapport en de daar bijhorende modules.

Tabel 3

Overzicht van de modules uit Deelrapport 2 en de bijhorende onderzoeksvragen

	Module 2.1	Module 2.2	Module 2.3
Methodologie	Documentanalyse	Cross-instrumentele studies	Focusgroepen
Onderzoeksvragen:			
Welke gelijkenissen en verschillen zijn er tussen de doorlichtingen, het peilingsonderzoek en het internationaal vergelijkend onderzoek?	X		
In welke mate convergeren of divergeren de resultaten van de verschillende bronnen?		X	
Welke mogelijke verklaringen zijn er voor de concordantie in de resultaten van de verschillende bronnen?			X
Hoe kunnen de verschillende bronnen complementair ingezet worden?			X

4.2.1 Gelijkenissen en verschillen tussen bronnen van externe kwaliteitszorg

De gelijkenissen en verschillen tussen de bronnen van externe kwaliteitszorg situeren zich op verschillende dimensies. Eerst en vooral zijn er gelijkenissen en verschillen tussen de bronnen in de doelen die ze vooropstellen. Concreet onderscheiden we daarbij drie verschillende doelen: systeemmonitoring, kwaliteitsbewaking en schoolfeedback. Systeemmonitoring omvat het opvolgen en evalueren van de Vlaamse onderwijskwaliteit doorheen de tijd. Dat staat centraal bij PIRLS, TIMSS, PISA en de peilingen. Die bronnen hebben namelijk als doel om op systeemniveau de Vlaamse onderwijskwaliteit doorheen de tijd op te volgen en te evalueren. Een tweede doel is kwaliteitsmonitoring. Daaronder verstaan we het toezicht om te garanderen dat zowel elke school als het Vlaamse onderwijs in zijn geheel aan de minimale verwachtingen voor kwaliteitsvol onderwijs voldoet. Kwaliteitsmonitoring staat voorop bij doorlichtingen en de peilingen. Die bronnen trachten immers te garanderen dat zowel elke school als het Vlaamse onderwijs in zijn geheel aan de minimale verwachtingen voor kwaliteitszorg onderwijs voldoet. Tot slot zet elke bron in op schoolfeedback. Door middel van een doorlichtingsverslag of schoolfeedbackrapport bieden die bronnen concrete handvatten aan scholen om hun interne kwaliteitszorg te verbeteren of te versterken.

Wat betreft de dimensie 'domein en construct' zijn er zowel gelijkenissen als verschillen te vinden. De doorlichtingen, het peilingsonderzoek en het internationaal vergelijkend onderzoek richten zich namelijk op de volgende constructen en domeinen:

- De onderwijsinspectie richt zich hoofdzakelijk op proceskenmerken die zich situeren op de klasvloer en op het niveau van het schoolbeleid.
- De peilingen meten kennis, inzicht, vaardigheden en in mindere mate attitudes over verschillende domeinen, zoals gedefinieerd in de Vlaamse onderwijsdoelen.
- PIRLS meet de vaardigheid voor begrijpend lezen.
- TIMSS richt zich op het meten van vaardigheden binnen het domein wiskunde en wetenschappen.
- PISA focust op functionele kennis en vaardigheden bij drie domeinen: leesvaardigheid, wiskundige geletterdheid en wetenschappelijke geletterdheid.

Hoewel de bronnen verschillende constructen en domeinen meten, is inhoudelijke overlap toch mogelijk, bijvoorbeeld tussen de peiling wiskunde basisonderwijs en TIMSS.

Om de vooropgestelde doelen te bereiken, verzamelen de bronnen data voor een bepaalde populatie. Eerst en vooral is er een duidelijk verschil in populatie tussen de doorlichtingen en de andere bronnen. De populatie van de doorlichtingen bestaat immers uit scholen en niet uit leerlingen (zoals wel het geval is bij de andere bronnen). Daarnaast zijn er zowel gelijkenissen als verschillen te vinden tussen de populatie van de peilingen en het internationaal vergelijkend onderzoek: (1) PISA definieert zijn populatie als 15-jarige leerlingen, (2) PIRLS en TIMSS richten zich op het vierde leerjaar en (3) de peilingen focussen op leerlingen die in het zesde leerjaar basisonderwijs of in het tweede leerjaar van de eerste graad secundair onderwijs zitten.¹

De steekproeftrekking verloopt vrij gelijkaardig bij de peilingsonderzoeken en de internationaal vergelijkende onderzoeken, terwijl de steekproef van de doorlichtingen er totaal anders uitziet. Bij PIRLS, TIMSS, PISA en de peilingen wordt er eerst een steekproef van scholen geselecteerd, waarna er in een tweede stap een specifiek aantal leerlingen of klassen geselecteerd wordt. De onderwijsinspectie

¹ Er zijn ook peilingen afgenomen op het einde van de tweede en derde graad. In dit onderzoek wordt er echter niet gefocust op die peilingen.

daarentegen moet rekening houden met de voorwaarde dat elke school minstens één keer om de zes jaar doorgelicht moet worden.

Onder modus verstaan we de manier waarop de data verzameld worden. Vooraleer er data verzameld kunnen worden, moeten er bij de peilingen en het internationaal vergelijkend onderzoek eerst toetsen en achtergrondvragenlijsten ontwikkeld worden. Die toetsontwikkeling verloopt bij die verschillende bronnen op een gelijkaardige manier. Ook de toetsafname verloopt gelijkaardig bij PIRLS, TIMSS, PISA en de peilingen. De gestandaardiseerde afname gebeurt immers steeds onder begeleiding van een leerkracht en/of toetsbegeleider. Bovendien worden er gedetailleerde handleidingen voorzien waarin de te volgen procedures stap voor stap beschreven worden.

De manier van dataverzameling verloopt anders bij de doorlichtingen dan bij de andere bronnen. Zo maakt de onderwijsinspectie gebruik van verschillende dataverzamelmethode, bijvoorbeeld gesprekken, observaties enzoverder. Daarnaast gebeurt de inschaling op de ontwikkelingsschalen tijdens een teamoverleg van de inspecteurs, wat een verschil is met PIRLS, TIMSS, PISA en de peilingen. Opvallend is wel dat de onderwijsinspectie – net zoals de andere bronnen – gebruik maakt van uitgebreide leidraden en/of instrumentenbundels.

Elke bron geeft aan hoe een bepaalde school of het Vlaams onderwijs zich verhoudt ten opzichte van een vastgelegde standaard. Met standaardbepaling refereren we naar de specifieke standaard die de bron daarbij hanteert. Dit onderzoek toont dat de standaardbepaling op diverse en uiteenlopende manieren gebeurt bij de verschillende bronnen van externe kwaliteitszorg.

Het peilingsonderzoek geeft aan hoe een bepaalde school of het Vlaams onderwijs zich verhoudt ten opzichte van de *eindtermen*. De peilingen maken daarvoor gebruik van een criteriumgerichte meting. Dat houdt in dat de onderwijsdoelen zelf het uitgangspunt vormen voor de cesuur die bepaalt of de leerlingen de onderwijsdoelen behalen. Daarenboven wijzigt die cesuur niet over herhalingspeilingen heen.

Het internationaal vergelijkend onderzoek geeft aan hoe het Vlaams onderwijs zich situeert ten opzichte van de internationaal bepaalde *benchmarks* (bij PIRLS en TIMSS) of de *vaardigheidsniveaus* (bij PISA). In tegenstelling tot bij de peilingen wordt er pas na de dataverzameling een interpretatie gegeven aan de verschillende benchmarks en vaardigheidsniveaus. Daarenboven kan de interpretatie die gegeven wordt aan een bepaald benchmark of vaardigheidsniveau beperkt variëren over cycli heen.

De onderwijsinspectie geeft aan hoe een bepaalde school zich verhoudt ten opzichte van de *ontwikkelingsschalen*. Voor elk proceskenmerk dat de onderwijsinspectie onderzoekt, is er een ontwikkelingsschaal opgesteld die bestaat uit vier ontwikkelingsniveaus: beneden, benadert, volgens en overstijgt de verwachting. Door de verzamelde data te vergelijken met kritische kenmerken bepaalt de onderwijsinspectie voor elk proceskenmerk het ontwikkelingsniveau van de school.

Wanneer we de analytische methode bij de peilingen en het internationaal vergelijkend onderzoek bekijken, zijn er zowel gelijkenissen als verschillen. Eerst en vooral maken de peilingen, PIRLS, TIMSS en PISA allemaal gebruik van modellen. Al verschilt het van bron tot bron welke specifieke modellen er gehanteerd worden. Daarenboven maken de peilingen en de internationale onderzoeken gebruik van een andere procedure om de vaardigheid van leerlingen te schatten. Bij een vergelijking van de analytische methode van de doorlichtingen en bovenstaande bronnen vallen er twee verschillen op. Ten eerste maken de doorlichtingen gebruik van ontwikkelingsschalen om tot een advies te komen en niet van vaardigheidsscores. Ten tweede hanteert de onderwijsinspectie geen modellen, maar

vooropgestelde criteria om op basis van inschalingen op de ontwikkelingsschalen tot een advies te komen.

Tot slot ontwikkelden we een matrix die aangeeft wat de mogelijkheden en beperkingen zijn van elke bron om betrouwbare en valide uitspraken te doen over specifieke aspecten van onderwijskwaliteit op een specifiek aggregatieniveau. Die matrix toont dat geen enkele bron uitspraken kan doen over alle aspecten van onderwijskwaliteit. Elke bron is namelijk ontwikkeld om uitspraken te doen op een specifiek aggregatieniveau. Omwille van keuzes die gemaakt zijn bij het design, de standaardbepaling en de analytische methoden gaan uitspraken op andere aggregatieniveaus immers samen met beperkingen. Bijvoorbeeld, de internationaal vergelijkende onderzoeken zijn gericht op het systeemniveau en focussen op leerlingen in een bepaald leerjaar of met een specifieke leeftijd. Bijgevolg worden uitspraken op basis van de internationale onderzoeken op schoolniveau steeds gekenmerkt door dezelfde beperking: ze zijn gebaseerd op slechts één leerjaar (PIRLS en TIMSS) of een beperkt aantal 15-jarige leerlingen (PISA). De resultaten in de feedbackrapporten kunnen daardoor niet veralgemeend worden naar de hele school.

Hoe de resultaten van verschillende bronnen – gegeven de beperkingen van elke bron – samengelegd kunnen worden, wordt besproken in '4.2.3 Complementariteit van de verschillende bronnen van externe kwaliteitszorg'. In wat volgt, wordt er eerst in kaart gebracht in welke mate de resultaten van verschillende bronnen convergeren of divergeren.

4.2.2 Mate van convergentie of divergentie tussen de resultaten van verschillende bronnen

Aan de hand van drie cross-instrumentele studies hebben we onderzocht in welke mate de resultaten van de doorlichtingen, het peilingsonderzoek en het internationaal vergelijkend onderzoek in Vlaanderen convergeren of divergeren. In wat volgt, maken we een onderscheid tussen conclusies met betrekking tot het systeem- en schoolniveau.

4.2.2.1 Mate van convergentie of divergentie op schoolniveau

Uit dit onderzoek blijkt dat de resultaten van de doorlichtingen, het peilingsonderzoek en het internationaal vergelijkend onderzoek op schoolniveau voor de meeste scholen bij elkaar aansluiten. Wanneer we controleren voor de achtergrondkenmerken van leerlingen en scholen sluiten de vaststellingen op schoolniveau beter bij elkaar aan en stijgt de concordantie.

Hoewel de resultaten van de verschillende bronnen bij de meeste scholen leiden tot een relatief convergent beeld – zeker wanneer we controleren voor achtergrondkenmerken – , zijn er ook steeds scholen waarbij de resultaten van meerdere bronnen leiden tot een divergent beeld. Die divergerende resultaten kunnen een grote impact hebben voor de scholen in kwestie. Dat zou bijvoorbeeld het geval zijn als er gevolgen gekoppeld worden (bijvoorbeeld een begeleidingstraject, een bijkomende doorlichting...) aan het resultaat van een school. Die impact zou bovendien nog groter worden als de resultaten van slechts één bron bepalend zijn voor de toekenning van dergelijke gevolgen. Daarom is het belangrijk om steeds beroep te doen op meerdere bronnen van externe kwaliteitszorg (zie '4.2.3 Complementariteit van de verschillende bronnen van externe kwaliteitszorg').

Over het algemeen stellen we dus vast dat de resultaten van de verschillende bronnen op schoolniveau voor de meeste – maar niet alle – scholen convergeren wanneer we een eerlijke vergelijking maken door te controleren voor achtergrondkenmerken van de leerlingen en scholen. Bij dat algemene antwoord op de onderzoeksvraag is het echter van belang om de volgende twee kanttekeningen te maken:

- Welke groepen leerlingen of scholen we met elkaar vergelijken, maakt een verschil: Zoals eerder aangehaald, richt elke bron zich op een specifieke populatie en zijn er heel wat verschillen in de populaties die de verschillende bronnen vooropstellen. Wanneer we twee

bronnen met elkaar vergelijken, moeten we rekening houden met die verschillen. Uit dit onderzoek blijkt dat welke groepen leerlingen of scholen we met elkaar vergelijken weldegelijk een verschil maakt voor de resultaten. Bijvoorbeeld, wanneer we de schooleffecten op basis van de peiling B-stroom vergelijken met de PISA-schooleffecten – die berekend zijn op basis van leerlingen uit de A-stroom, B-stroom, ASO, TSO en BSO – leidt dat voor een groot aantal scholen tot een divergent beeld. Echter, wanneer we een vergelijking maken met enkel de PISA-leerlingen uit de B-stroom en het BSO is er een relatief grote overeenkomst tussen het resultaat op PISA en de peiling wiskunde B-stroom.

- Een absoluut criterium vergelijken met een relatief criterium:
De onderwijsinspectie werkt met een absoluut criterium. Elke school wordt vergeleken met een criterium – namelijk de verwachting – en voldoet bijgevolg wel of niet aan dat criterium. Het peilingsonderzoek en het internationaal vergelijkend onderzoek daarentegen maken (bij het schatten van schooleffecten) gebruik van een relatief criterium. Daarbij wordt elke school vergeleken met het gemiddelde. Dat gemiddelde is echter onderhevig aan fluctuaties in leerprestaties doorheen de tijd, terwijl een absoluut criterium constant blijft doorheen de tijd.

Hoe het absoluut criterium bij de doorlichtingen zich juist verhoudt ten opzichte van het relatief criterium bij de verschillende peilingen en internationale onderzoeken is onduidelijk. Wel stellen we vast dat voldoen aan de verwachting bij de doorlichting niet noodzakelijk overeenkomt met het Vlaams gemiddelde. Om te kunnen besluiten of een school – of leerling – voldoet aan de minimale verwachting – of eindtermen – is een vergelijking met een absoluut criterium dus noodzakelijk. Omgekeerd is een vergelijking met het Vlaamse gemiddelde – het relatieve criterium – essentieel om te kunnen concluderen of een school (onder of boven)gemiddeld presteert.

4.2.2.2 Mate van convergentie of divergentie op systeemniveau

Op systeemniveau hebben we een vergelijking gemaakt tussen de trends voor doorlichtingsronde 3 en de trends voor Inspectie 2.0. Ten eerste toont die vergelijking dat het percentage scholen dat een specifieke inschaling ontvangt van de onderwijsinspectie relatief sterk varieert over schooljaren heen. Zo varieert het percentage scholen dat aan de verwachting voldoet bijvoorbeeld sterk over verschillende schooljaren heen. Die vaststelling gaat op voor zowel doorlichtingsronde 3 als Inspectie 2.0. Ten tweede stellen we vast dat de uiterste beoordelingscategorieën vaker gebruikt worden bij doorlichtingsronde 3. Dat is een duidelijk verschil met de resultaten bij Inspectie 2.0 waarbij voornamelijk de middelste twee categorieën – benadert en volgens de verwachting – gebruikt worden. Ook hier is het van belang om enkele kanttekeningen te maken:

- Doorlichtingsronde 3 is een andere doorlichtingsmethode dan Inspectie 2.0:
Er zijn heel wat verschillen tussen de doorlichtingsmethodes ‘doorlichtingsronde 3’ en ‘Inspectie 2.0’. Meer bepaald gaat het om verschillen in (1) beoordeelde proceskenmerken, (2) aantal inschalingen per proceskenmerk en (3) beoordelingsniveaus. Die verschillen zorgen ervoor dat het niet eenvoudig is om de trends voor de doorlichtingen in kaart te brengen. De resultaten moeten dus met enige voorzichtigheid geïnterpreteerd worden.
- Er zijn inhoudelijke aanpassingen mogelijk binnen eenzelfde doorlichtingsmethode:
Op systeemniveau is er heel wat variatie in de resultaten van de doorlichtingen over verschillende schooljaren heen, zelfs binnen eenzelfde doorlichtingsmethode. Die variatie is mogelijk deels te wijten aan de inhoudelijke aanpassingen die de onderwijsinspectie doet binnen eenzelfde doorlichtingsmethode. De onderwijsinspectie kan namelijk de inhoudelijke omschrijving van de proceskenmerken bij doorlichtingsronde 3 en/of de kritische kenmerken

die horen bij de ontwikkelingsschalen van Inspectie 2.0 bijsturen en aanpassen. Om zeker te zijn dat die variatie niet te wijten is aan inhoudelijke aanpassingen is het noodzakelijk dat dezelfde constructen ongewijzigd blijven en steeds op dezelfde manier gemeten worden binnen eenzelfde doorlichtingsmethode.

Daarnaast worden de resultaten van het peilingonderzoeken en het internationaal vergelijkend onderzoek op systeemniveau vergeleken. Meer specifiek wordt er een vergelijking gemaakt tussen (1) de peiling Nederlands en PIRLS, (2) de peiling wiskunde basisonderwijs en TIMSS en (3) de peilingen wiskunde A- en B-stroom en PISA. Over het algemeen vinden we een dalende trend in de leerprestaties van leerlingen. Al zijn er ook enkele peilingschalen die stabiel blijven of stijgen doorheen de tijd.

Wanneer we de trends – berekend door de IEA, OESO en STEP – van de peilingen en het internationaal onderzoek dat hetzelfde construct beoogt vergelijken, is er eenzelfde richting in de trends. De sterkte van de trends is wel verschillend. Echter, wanneer we voor zowel de internationale studies als peilingsonderzoeken een herkalibratie doen aan de hand van hetzelfde model op basis van enkel de Vlaamse gegevens, dan wordt ook de sterkte van de trends meer gelijkaardig. Anders gezegd, de resultaten van het peilingsonderzoek en het internationaal vergelijkend onderzoek convergeren meer op systeemniveau wanneer we een gelijkaardige schattingsprocedure hanteren. Daarbij is het echter van belang om de volgende opmerkingen te maken:

- Welke groepen leerlingen of scholen we met elkaar vergelijken, maakt een verschil: Net zoals bij het schoolniveau moeten we ook bij de vergelijking van trends rekening houden met die verschillen in populaties. Bijvoorbeeld, bij PISA worden de trends berekend op basis van een diverse groep leerlingen die niet allemaal dezelfde onderwijsvorm volgen. Bij het peilingsonderzoek zijn er aparte peilingen voor de A- en B-stroom. Dit onderzoek bevestigt dat het weldegelijk een verschil maakt welke groepen leerlingen of scholen we met elkaar vergelijken.
- Er zijn aanpassingen mogelijk aan de gemeten constructen bij het internationaal vergelijkend onderzoek: Bij elke nieuwe cyclus doen de internationale onderzoeken aanpassingen aan de constructen die ze meten. Bijgevolg zijn de trends bij internationaal vergelijkende onderzoeken mogelijk (deels) toe te schrijven aan veranderingen in het gemeten construct. Om zeker te zijn dat de trends niet te wijten zijn aan veranderingen in het gemeten construct is het essentieel dat de constructen ongewijzigd blijven en steeds op dezelfde manier gemeten worden over cycli heen.

4.2.3 Complementariteit van de verschillende bronnen van externe kwaliteitszorg

Om de onderwijskwaliteit van scholen en onderwijssystemen in te schatten wordt het gebruik van meerdere bronnen sterk aanbevolen. Nochtans tonen de bovenstaande resultaten dat het eenvoudigweg gebruik maken van meerdere bronnen niet automatisch leidt tot eenduidige en betrouwbare besluitvorming over onderwijskwaliteit. Zo illustreert de matrix dat geen enkele bron valide uitspraken kan doen over alle aspecten van onderwijskwaliteit en tonen de resultaten uit de cross-instrumentele studies dat informatie uit meerdere bronnen ook tot een divergerend beeld kan leiden.

Verder bouwend op de resultaten uit de eerste twee delen focussen we in dit derde deel op de complementariteit van de verschillende bronnen. We onderzochten enerzijds de mogelijke verklaringen voor de concordantie tussen de resultaten van de verschillende bronnen en anderzijds hoe de verschillende bronnen van externe kwaliteitszorg gecombineerd kunnen worden. Daarvoor organiseerden we twee online focusgroepen die vervolgens getranscribeerd, gecodeerd en geanalyseerd werden.

4.2.3.1 Verklaringen voor de mate van concordantie tussen de resultaten van verschillende bronnen

De leden van de focusgroepen geven eerst en vooral aan dat er geen concreet antwoord is op wat nu dé oorzaak is van die convergerende en divergerende resultaten. Daarom beperken de focusgroepen zich tot mogelijke verklaringen die een rol kunnen spelen. Hieronder bespreken we de mogelijke verklaringen afzonderlijk. We benadrukken dat er in werkelijkheid hoogstwaarschijnlijk meerdere verklaringen tegelijkertijd spelen.

De vastgestelde gelijkenissen en verschillen tussen bronnen van externe kwaliteitszorg

Volgens de leden van de focusgroepen duiden de gelijkenissen en verschillen tussen de bronnen voor een stuk de gevonden resultaten. Zoals eerder omschreven, situeren die gelijkenissen en verschillen zich op verschillende dimensies: doel, domein en construct, populatie en steekproef, modus, standaardbepaling en analytische methode.

De leden van de focusgroepen bevestigen dat elke bron ontwikkeld wordt met oog op een specifiek *doel*. Bijgevolg is niet elke bron even geschikt om uitspraken te doen op elk aggregatieniveau. De toetsen van de peilingen en de internationaal vergelijkende onderzoeken zijn bijvoorbeeld niet ontworpen om op schoolniveau een betrouwbaar resultaat te geven. Een doorlichting van de onderwijsinspectie daarentegen heeft net wel als doel om betrouwbare uitspraken op schoolniveau te doen. Volgens de focusgroepen kunnen die gelijkenissen en verschillen in het beoogde doel aanleiding geven tot zowel convergerende als divergerende resultaten.

Daarnaast is wat er gemeten wordt (*domein en construct*) gewoonweg anders bij de verschillende bronnen van externe kwaliteitszorg. De deelnemers in de focusgroepen veronderstellen meer convergerende resultaten wanneer er meer overeenstemming is in wat er gemeten wordt. Daarnaast kan ook het moment waarop bepaalde leerinhouden aan bod komen bepalend zijn. In Vlaanderen zijn er immers verschillende leerplannen. Ook dat kan op systeemniveau in het basisonderwijs een verklarende factor zijn. Zo is het mogelijk dat inhouden die gemeten worden in TIMSS of PIRLS (in het vierde leerjaar) volgens een bepaald leerplan pas in de derde graad aan bod komen, terwijl die inhouden wel al in de tweede graad aan bod komen bij een ander leerplan. In dat geval kan het verschil in leerplannen een mogelijke verklaring zijn voor de mate van concordantie tussen de resultaten van PIRLS/TIMSS en de peilingen (afgenomen in het zesde leerjaar).

Ook de *populatie* en de bijhorende *steekproef* kunnen een mogelijke verklaring zijn van zowel convergerende als divergerende resultaten. Ten eerste kan de selectie van (een groep van) specifieke leerlingen voor een stuk de resultaten op schoolniveau verklaren. Zo doet het er bijvoorbeeld toe welke klas deelneemt aan een peiling of in welk schooljaar een bron wordt afgenomen. Als in dit schooljaar de leerlingen op het einde van de eerste graad erg sterk zijn, haalt die school dit jaar misschien een beter resultaat op de peiling A-stroom of B-stroom dan in andere jaren.

Ten tweede is er op systeemniveau een groot verschil wat betreft de verplichting tot deelname. Een doorlichting kunnen scholen niet weigeren, terwijl de vraag tot deelname aan een peilings- of internationaal onderzoek wel geweigerd kan worden. Dat heeft als gevolg dat er voor elke school een doorlichtingsverslag beschikbaar is. Dat is anders bij het peilings- en internationaal onderzoek, aangezien scholen zelf ervoor kunnen kiezen om niet deel te nemen. De groep scholen die geen resultaten hebben op de peilingen en/of het internationaal vergelijkend onderzoek, zijn bijgevolg niet opgenomen in de analyses. Volgens de focusgroepen verwachten scholen die ervoor kiezen om niet deel te nemen vaak van zichzelf dat ze niet echt goed gaan scoren. Dat kan mogelijks een invloed hebben op de gevonden concordantie tussen de resultaten van verschillende bronnen.

Volgens de focusgroepen speelt ook de manier waarop de data verzameld worden (*modus*) een rol. Wanneer de twee bronnen exact hetzelfde meten, kunnen de resultaten toch divergeren omwille van een verschil in de manier waarop de data verzameld werden. Zo is het mogelijk dat de leerkracht in de vragenlijst sociaal wenselijke antwoorden invult en dat die antwoorden daarom sterk verschillen van de beoordeling door de externe inspecteur. Ook de *modus* kan dus een mogelijke verklaring zijn voor de mate van concordantie.

Ook de *standaardbepaling* gebeurt op uiteenlopende manieren bij de verschillende bronnen van externe kwaliteitszorg. Dat kan eveneens bijdragen aan de gevonden resultaten. In de focusgroep wordt er aangehaald dat het niveau dat volgens eindtermen verwacht wordt – en dus gemeten wordt bij de peilingen – eigenlijk niet zo hoog is in vergelijking met wat er verwacht wordt in internationaal vergelijkend onderzoek. In het internationaal onderzoek zijn de items daardoor wellicht moeilijker dan de items in de peilingstoetsen, wat kan leiden tot divergerende resultaten tussen die bronnen. De input van de focusgroepen is in lijn met de resultaten van PIRLS Repeat (Dockx et al., 2019).

Tot slot is de *analytische methode* gewoonweg verschillend bij de bronnen van externe kwaliteitszorg. Bij de peilingen en de internationaal vergelijkende onderzoeken wordt er gewerkt met scoring en verschillende modellen. Dat is iets totaal anders en kan nooit helemaal hetzelfde zijn als de manier waarop de onderwijsinspectie de verzamelde data analyseert. De analyse van kwantitatieve data bij de peilingen en de internationale onderzoeken enerzijds en de analyse van voornamelijk kwalitatieve data bij de doorlichtingen kunnen op zich al aanleiding geven tot convergerende of divergerende resultaten – afhankelijk van welke bronnen je vergelijkt.

Andere mogelijke verklaringen

De vastgestelde gelijkenissen en verschillen tussen de bronnen van externe kwaliteitszorg zijn echter niet de enige mogelijke verklaringen. Er zijn volgens de focusgroepen heel wat andere mogelijke oorzaken, zowel op school- als op systeemniveau, die (gedeeltelijk) kunnen bijdragen tot convergerende en divergerende resultaten.

Ten eerste scoren de meeste scholen in Vlaanderen gemiddeld, zoals blijkt uit de voorgaande resultaten. Volgens de leden van de focusgroepen kan dat ook een verklaring zijn voor de gevonden convergerende resultaten. Wanneer de meeste scholen gemiddeld presteren, gebeurt het vaak dat scholen ook bij verschillende bronnen in die groep ‘gemiddeld’ vallen.

Ten tweede blijkt uit de resultaten van de cross-instrumentele studies dat de resultaten het sterkst convergeren wanneer er gecontroleerd wordt voor de achtergrondkenmerken van leerlingen en scholen. Bijgevolg duiden die achtergrondkenmerken van leerlingen en scholen ook een deel van de gevonden resultaten. Ook de participanten van de focusgroepen bevestigen dat. De achtergrondkenmerken van een school en de wijze waarop een school met haar leerlingenpubliek omgaat, kunnen voor een stuk de gevonden resultaten verklaren.

Ten derde is elke meting door een bron van externe kwaliteitszorg een momentopname en worden die metingen bovendien niet op exact hetzelfde moment uitgevoerd. Er kan in een school heel wat veranderen tussen het tijdstip waarop de ene meting afgenomen wordt en het tijdstip waarop de volgende meting plaatsvindt. Dergelijke verandering kan ook plaatsvinden naar aanleiding van een doorlichting of het ontvangen van een schoolfeedbackrapport. Opvallend is wel dat een doorlichting – in tegenstelling tot de andere bronnen van externe kwaliteitszorg – volgens de leden van de focusgroepen vaker leidt tot verandering.

Ten vierde gaat het niet alleen op schoolniveau steeds om momentopnames. Ook op systeemniveau zijn algemene tendensen en veranderingen tussen de tijdstippen van twee metingen mogelijk. Daar spelen

er mogelijks enkele algemene tendensen die een verklaring kunnen zijn voor de mate van concordantie op systeemniveau, bijvoorbeeld het lerarentekort, de coronapandemie...

Ten vijfde kunnen scholen die deelnemen aan een peiling, internationaal vergelijkend onderzoek of doorlichting ook strategisch gedrag vertonen. Zo kan het gebeuren dat een school bepaalde leerlingen niet laat deelnemen aan een internationale toets, omdat ze bang zijn om gezichtsverlies te lijden. Dergelijk strategisch gedrag kan ook een invloed hebben op de resultaten en bijgevolg ook op de mate van concordantie op het einde van de rit. De leden van de focusgroepen geven vanuit hun ervaring echter aan dat het aantal scholen dat strategisch gedrag vertoont op dit moment nog beperkt is.

Tot slot meet een toets van de peilingen of een internationaal vergelijkend onderzoek niet alleen de vaardigheid van de leerlingen, maar ook de houding van de leerlingen ten aanzien van die toets. Volgens de focusgroepen is dat van belang, omdat toetsmotivatie in sterke mate samenhangt met de resultaten die leerlingen behalen. Hoe hoger de leerlingen het belang van de toets percipiëren, hoe meer ze zich gaan inzetten, aldus de focusgroepen.

4.2.3.2 Resultaten van verschillende bronnen combineren

Het werd reeds herhaaldelijk vermeld dat wat de verschillende bronnen van externe kwaliteitszorg meten en hoe ze dat meten, verschilt van bron tot bron. Die diversiteit beschouwen de leden van de focusgroepen net als de sterkte van die verschillende bronnen. Elke bron vormt slechts één stukje van de puzzel. Het is daarom belangrijk dat zowel het Vlaamse onderwijs als de scholen beroep kunnen doen op meerdere bronnen die elk een verschillend perspectief belichten.

Drie sleutelfactoren: duiding, meerdere bronnen en vertaling naar de onderwijspraktijk

Om de verschillende bronnen van externe kwaliteitszorg te combineren, zijn drie sleutelfactoren cruciaal op zowel het school- als systeemniveau. Concreet gaat het om: (1) duiding, (2) meerdere bronnen gebruiken en (3) vertaling naar de onderwijspraktijk. Hoewel de sleutelfactoren hieronder afzonderlijk toegelicht worden, hangen ze in realiteit onderling sterk samen. Ze kunnen in de praktijk dus niet strikt van elkaar onderscheiden worden.

De eerste sleutelfactor is *duiding*. Daaronder verstaan we het juist begrijpen wat elke bron wel en niet meet. Om de meerwaarde van elke bron te kunnen zien, is duiding immers essentieel. Enkel wanneer de gelijkenissen en verschillen in het onderzoeksopzet op voorhand helder zijn, is het mogelijk om de resultaten van verschillende bronnen in verhouding tot elkaar te interpreteren.

De tweede sleutelfactor is *meerdere bronnen* gebruiken. Eén bron zegt ook niet alles en geeft scholen weinig inzicht in mogelijke verbeterstrategieën. Elk van die bronnen is immers slechts een puzzelstukje. Bijgevolg is het waardevol om zowel op school- als systeemniveau beroep te doen op meerdere bronnen. De kunst bestaat erin om die verschillende bronnen of puzzelstukjes vanuit het juiste perspectief naast elkaar te leggen. Zowel op school- als systeemniveau is duiding (Sleutelfactor 1) dus essentieel om de resultaten van verschillende bronnen naast elkaar te leggen en in verhouding tot elkaar te interpreteren.

De derde en laatste sleutelfactor is de *vertaling naar de onderwijspraktijk*. Die sleutelfactor verwijst naar de vraag: wat betekenen de verzamelde resultaten voor onze school of ons onderwijssysteem? Louter de beschikbaarheid van data leidt namelijk niet tot een sterkere onderwijskwaliteit. Een school of onderwijssysteem moet aan de slag gaan met die resultaten om tot mogelijke verbeterstrategieën te komen. Er bestaat echter geen stappenplan dat scholen of onderwijssystemen kunnen volgen om die vertaling te maken. Die vertaling vraagt ruimte en tijd om in dialoog te gaan, dingen uit te proberen, te evalueren, waar nodig bij te sturen en goede elementen te borgen.

Uitdagingen op schoolniveau

Uit de focusgroepen blijkt dat het voor scholen niet evident is om de resultaten van verschillende bronnen te combineren en interpreteren. Er zijn heel wat uitdagingen die scholen tegenkomen wanneer ze aan de slag gaan met de resultaten van verschillende bronnen.

Zo is er om te beginnen de interpretatie van de schoolfeedbackrapporten. Verschillende participanten geven aan dat de interpretatie van schoolfeedbackrapporten niet evident is voor scholen. Vooral de interpretatie van kwantitatieve data blijkt niet eenvoudig voor scholen. Nochtans is die interpretatie essentieel om op schoolniveau aan de slag te kunnen met de resultaten van verschillende bronnen.

Daarnaast is er een gebrek aan tijd en ruimte om bronnen te combineren en te vertalen. Zoals eerder aangehaald, is de vertaling naar de onderwijspraktijk een proces dat tijd nodig heeft. In de realiteit komen de resultaten van bronnen van externe kwaliteitszorg bovendien op zeer gevarieerde tijdstippen binnen bij een school. Op dat moment wordt vaak niet meteen die puzzel gelegd wegens gebrek aan tijd en ruimte.

Een volgende uitdaging is het omgaan met divergerende resultaten. In de praktijk gebeurt het wel degelijk dat scholen divergerende of tegenstrijdige resultaten krijgen van verschillende bronnen. Het kan voor scholen heel verwarrend zijn als ze in het ene onderzoek een bepaald resultaat behalen en in een ander onderzoek een totaal ander resultaat behalen. Toch is het mogelijk om aan die divergerende resultaten een interpretatie te geven door de sterke en minder sterke kanten van de verschillende bronnen te kennen (Sleutelfactor 1: Duiding).

De hoeveelheid data die ter beschikking gesteld wordt aan scholen is aanzienlijk. De bronnen van externe kwaliteitszorg zijn namelijk slechts een fractie van de hele berg aan beschikbare data voor scholen. Die grote hoeveelheid maakt het erg complex voor scholen en bemoeilijkt de interpretatie en combinatie van de resultaten.

Tot slot is er nood aan begeleiding en datageletterdheid. Er zijn immers heel wat uitdagingen die scholen tegenkomen bij de combinatie van verschillende bronnen. Het begrijpen en lezen van de informatie in een schoolfeedbackrapport of doorlichtingsverslag is slechts de eerste stap en vaak loopt het al mis bij die eerste stap. Het is daarom noodzakelijk om te werken aan datageletterdheid bij leraren en schoolleiders. Meer specifiek kan een professionaliseringstraject datageletterdheid zich richten op het aanreiken van concrete handvatten om aan de slag te gaan met data en opleiding geven over statistische terminologie en het lezen van visueel weergegeven resultaten. Daarnaast benadrukken de focusgroepen dat scholen beroep moeten kunnen doen op pedagogische begeleidingsdiensten om hen te ondersteunen bij de interpretatie en combinatie van verschillende bronnen enerzijds en bij de vertaling naar de praktijk anderzijds.

Uitdagingen op systeemniveau

Ook op systeemniveau situeren zich enkele gelijkaardige uitdagingen met betrekking tot de combinatie van verschillende bronnen van externe kwaliteitszorg. Niet alleen op schoolniveau, maar ook op systeemniveau zijn divergerende resultaten mogelijk. Door een goed zicht te hebben op het doel en bijhorend onderzoeksopzet van elke bron is het toch mogelijk om die divergerende resultaten te interpreteren. Ook hier is het dus essentieel om de sterke en minder sterke kanten van de verschillende bronnen te kennen (Sleutelfactor 1: Duiding).

De leden van de focusgroep wijzen bovendien op enkele ontbrekende puzzelstukken. Zo zijn er volgens hen een aantal zaken in Vlaanderen waar we voorlopig nog geen zicht op hebben: de stabiliteit van indicatoren enerzijds en de link tussen de toetsen en de processen in de klas anderzijds.

Tot slot is datageletterdheid en het gebrek aan duiding een uitdaging. Zo gaven enkele leden van de focusgroepen aan dat ze voorafgaand aan dit onderzoek nog nooit een overzicht kregen van wat elke bron van externe kwaliteitszorg exact meet en hoe ze dat meet. Nochtans blijkt uit de analyses dat duiding over het onderzoeksopzet een cruciaal element is om die verschillende bronnen te kunnen combineren. Bijgevolg is professionalisering die zich daarop richt niet alleen op schoolniveau, maar ook op systeemniveau aangewezen.

Een toekomst met Vlaamse toetsen

Tijdens de focusgroepen stond de combinatie van resultaten van de doorlichtingen, de peilingen en het internationaal vergelijkend onderzoek centraal. Vanaf het schooljaar 2023-2024 wordt het peilingsonderzoek echter vervangen door de Vlaamse toetsen. Uit de analyse van de focusgroepen blijkt dat bovenstaande resultaten ook relevant zijn voor een toekomst waarin het peilingsonderzoek vervangen wordt door de Vlaamse toetsen:

- Elke bron blijft slechts één stukje van de puzzel.
- De drie sleutelfactoren zijn cruciaal om verschillende bronnen van externe kwaliteitszorg te combineren, ongeacht of het gaat om de resultaten van de peilingen of de resultaten van de Vlaamse toetsen.
- De besproken uitdagingen op school- en systeemniveau blijven ook in een toekomst met Vlaamse toetsen pertinent.

Daarnaast wijzen de leden van de focusgroepen op enkele aandachtspunten. Ten eerste verliezen we ook indicatoren doordat de peilingen verdwijnen. Bij de peilingen werden er meerdere domeinen bevraagd, terwijl de Vlaamse toetsen enkel focussen op wiskunde en Nederlands. Daarom pleiten de focusgroepen ervoor om de puzzel zo rijk en divers mogelijk te houden in de toekomst. Ten tweede geven de leden van de focusgroepen aan dat duiding (Sleutelfactor 1) nog belangrijker zal worden met de komst van de Vlaamse toetsen. Het moet voor scholen duidelijk zijn wat ze wel en niet kunnen leren van de Vlaamse toetsen. Ten derde benadrukken de focusgroepen dat het belangrijk is om ook bij de Vlaamse toetsen steeds te controleren voor achtergrondkenmerken. Het gaat immers over hoeveel extra leerwinst een school creëert en niet over welke instroom van leerlingen die school heeft. Tot slot wijzen de focusgroepen op enkele kansen. Zo wordt het in de toekomst eenvoudiger om de concordantie tussen bronnen van externe kwaliteitszorg in kaart te brengen, aangezien elke school zal deelnemen aan de Vlaamse toetsen. Dat schept mogelijkheden voor verder onderzoek en verdere samenwerking tussen onderzoekers van de Vlaamse toetsen, de internationaal vergelijkende onderzoeken en de onderwijsinspectie.

5. Beleidsaanbevelingen

Onderwijsonderzoek kan niet in de plaats van het beleid treden. Het beleid neemt beslissingen en treft maatregelen. Daarbij laat het zich inspireren en leiden door een geheel van kennis, waarden en opvattingen inzake onderwijs, individu en samenleving. Wat beleidsondersteunend onderzoek doet is op een wetenschappelijk verantwoorde wijze gegevens verzamelen, ordenen en in samenhang presenteren. In dit onderzoek betreft het gegevens die beleidsmakers kunnen inbrengen in hun beslissing of, en zo ja volgens welke modaliteiten, in het Vlaamse onderwijs een (verder) uitgebouwd systeem van monitoring van aspecten van onderwijskwaliteit opzetten wenselijk en haalbaar is. Vertrekkende vanuit deze positionering willen we desalniettemin de kans nemen om ook als onderzoekers standpunt in te nemen. Elk onderzoeksteam neemt daarbij het voortouw vanuit het eigen deelrapport. Dat standpunt hebben we vertaald in principes die in het uitdenken van een beleid inzake ontwikkeling en monitoring van onderwijskwaliteit van wezenlijk belang zijn. In wat volgt beschrijven we vanuit het voorliggende onderzoek een aantal principes op basis waarvan (het samenspel van) het OK-kader, de doorlichting en gestandaardiseerde toetsen ons inziens nog sterker van betekenis kunnen zijn in het Vlaamse onderwijs. Het realiseren van deze uitgangspunten is wat wij aanbevelen. Daarom hebben we ze als imperatief geformuleerd. We richten ons daarbij op actoren op verschillende niveaus in het onderwijsveld. We vertrekken van richtinggevende principes die vervolgens door middel van concrete aanbevelingen verder uitgewerkt worden. De principes vloeien voort vanuit empirische evidentie met betrekking tot ontwikkeling en monitoring van onderwijskwaliteit uit verschillende onderwijscontexten, dit onderzoek inclusief. De aanbevelingen zijn toekomstgericht. Hoewel dit onderzoek geen betrekking had op de concrete intenties inzake uitrol van centrale toetsen, gaat het om adviezen die ons inziens best meegenomen worden bij de verdere implementatie van een systeem van kwaliteitszorg waar ook de Vlaamse toetsen deel van uitmaken.

Aanbevelingen bij het Referentiekader Onderwijskwaliteit

[Blijf – in de geest van het Referentiekader Onderwijskwaliteit – overheen onderwijspartners verder werken aan een gemeenschappelijke taal over wat gedeeld wordt in de kijk op onderwijskwaliteit](#)

Dit onderzoek leert dat de bevroegde respondenten de relevantie en zinvolheid van een kader als het Referentiekader Onderwijskwaliteit sterk onderschrijven. De mogelijkheid om overheen contexten, instanties en gebruikers gebruik te maken van een gedeelde taal wordt sterk gewaardeerd. Deelaspecten van goed onderwijs worden benoemd en in een ruimer kader samengebracht. Lovenswaardig daarbij is ook dat het OK-kader vanuit een gemeenschappelijk zorg voor onderwijskwaliteit bij verschillende kernpartners ontstond. Wanneer met voldoende respect voor elkaars opdracht en eigenheid welwillend gezocht wordt naar wat bindt overheen overheidsinstanties, onderwijsinspectie en onderwijsverstrekker (met hun begeleidingsdiensten) kunnen kaders ontstaan die tot onderlinge afstemming leiden en die gebruikers in scholen helpen om coherentie te zien in de verschillende beleidskanalen die als een brandpunt bij hen samenkomen. We adviseren alle onderwijspartners daarom om (ook in andere contexten) in de geest van dit OK-kader verder te blijven werken aan een gemeenschappelijke taal over wat gedeeld wordt in de kijk op onderwijskwaliteit.

[Verfijn het OK-kader voor verder gebruik, ook voor gebruik door de leraar](#)

Deze studie leert dat het voor een behoorlijk deel van de bevroegde schoolleiders en leraren niet ten volle duidelijk is welke de concrete kwaliteitsverwachtingen zijn waaraan zij geacht worden tegemoet te komen. Ondanks de verdienste van het Referentiekader Onderwijskwaliteit (cf. vorige aanbeveling) wordt het kader ook als weinig concreet en functioneel omschreven. Verwachtingen en standaarden blijven daardoor nog te vaak impliciet, waardoor ze dreigen in te boeten op gemeenschappelijkheid,

richtinggevendheid en afdwingbaarheid. Dat weegt op de mate waarin verwachtingen een bron kunnen zijn voor professionele dialoog en een concreet instrument voor vormen van afleggen van verantwoording. Verdere concretisering van het referentiekader tot meer concrete omschrijvingen en heldere indicatoren lijkt ons daarom wenselijk.

Daarnaast creëert het OK-kader weliswaar een gemeenschappelijke taal omtrent onderwijskwaliteit, maar de vaststelling is dat de uitwerking ervan zich sterk heeft gericht op het beleidsniveau in scholen. Een dragend idee bij de ontwikkeling van het OK-kader was echter om ook 'het kernproces leren en onderwijzen' een meer zichtbare, transparante en fundamentele plaats te geven. Inzetten op een Referentiekader voor Onderwijskwaliteit dat ook leerkrachten bereikt, net daarin lag volgens de ontwikkelgroep van het OK-kader de kans tot toegevoegde waarde. Die meerwaarde wordt momenteel nog niet waargemaakt. Daarom zou het zinvol zijn het OK-kader – of een parallelle complementaire versie ervan - ook verder te verfijnen en uit te werken voor gebruik door de leraar.

[Bewaak dat autonomie bij gebruik van het OK-kader niet overhelt naar een ervaren gebrek aan houvast](#)

Het OK-kader wordt gepercipieerd als een integraal kader, dat een eerder holistisch beeld wil schetsen van wat 'onderwijskwaliteit bieden' inhoudt, zo leert evidentie in deze studie. Dat is in lijn met de opzet van het kader, gegeven dat het net aan de gebruiker is om in de eigen specifieke context de beschikbare autonomie aan te wenden om tot een invulling op maat te komen. Die autonomie aanwenden vergt echter beleidskracht én de beschikbaarheid van tijd bij gebruikers, twee voorwaarden die niet als gegeven mogen verondersteld worden. Dat maakt dat rekenen op de autonomie van gebruikers beperkingen met zich meebrengt in behoorlijk wat schoolcontexten. Zij zijn vaak vragende partij voor verdergaande houvast, en voelen zich als het ware deels in de steek gelaten wanneer onderwijsverstrekkers, inspectie en/of overheidsinstanties omwille van principes van vrijheid van onderwijs het verdere initiatief tot invulling en uitwerking te sterk bij de school zelf leggen. Het is dan ook van belang de mate van autonomie bij gebruik van het OK-kader te bewaken, zodat deze niet overhelt naar een ervaren gebrek aan houvast.

Deze aanbeveling richt zich niet alleen op de wijze waarop het OK-kader centraal uitgewerkt wordt maar ook op de wijze waarop scholen ondersteund worden in het gebruik ervan. Het ervaren van passende ondersteuning bij het implementeren en gebruiken van het OK-kader blijkt immers een noemenswaardige factor in het begrijpen van verschillen in gebruik van het OK-kader in groepen van scholen. Ondersteuning kan variëren van zeer open, over inspirerend tot sturend. Ondersteuning die stuurt vanuit de uitgesproken intentie om de gebruikers te voorzien van passende kaders, die meteen voldoende voor gebruik uitgewerkt zijn, wordt niet als sturend maar als ondersteunend ervaren. Uiteraard zonder dat de sturing als opgelegd zonder ruimte voor eigen inbreng ervaren wordt. Een gebrek aan zulke ondersteuning leidt tot ongewenste neveneffecten als demotivatie en onzekerheid. Vandaar een pleidooi voor het durven bieden van houvast aan gebruikers. Dat brengt ons naadloos bij de volgende beleidsaanbeveling.

[Durf in communicatie over het OK-kader sterker normerend zijn omtrent inhouden en gebruik](#)

Uit de resultaten van dit onderzoek blijkt dat diegenen die vertrouwd zijn met het OK-kader, het kader omschrijven als een relevant kader dat een meerwaarde biedt binnen de werking van de school, weze het - in de huidige versie - veelal op schoolbeleidsniveau. Vooral voor startende schoolleiders of schoolleiders die alle aspecten van het schoolbeleid (grondig) willen reorganiseren of heropbouwen, biedt het OK-kader naar eigen zeggen de nodige structuur en richting. Daarnaast leidt het OK-kader in de scholen waar het een weg vond vaak tot instrumentele effecten op vlak van kwaliteitsontwikkeling, met meer aandacht voor de evaluatie van onderwijskwaliteit.

Hoewel het OK-kader beoogt toetsstenen voor onderwijskwaliteit aan te reiken en duidelijkheid te scheppen in de beoordeling van onderwijskwaliteit, wordt het OK-kader in beleidsmiddelen als 'niet-normerend' beschouwd. Dat is ook conform de opzet van het OK-kader. Ondanks het potentieel van het kader en de onderzoeksevidentie die er achter schuilgaat, is er in de communicatie over het kader vaak sprake van vrijblijvendheid, in die zin dat het slechts één kader is naast andere. Geenszins willen we hier de indruk wekken dat dit onderzoek leert dat scholen die intensief werken met het OK-kader per definitie ook kwaliteitsvol onderwijs aanbieden. Toch is ons standpunt dat de kracht waarmee dit kader in de markt gezet wordt sterker kan, en meer normerend mag. Het is immers onmiskenbaar zo dat de onderwijsinspectie tijdens een schooldoorlichting nagaat of de onderwijskwaliteit van de desbetreffende school voldoet aan de kwaliteitsverwachtingen zoals geformuleerd in het OK-kader. Dat maakt het kader wel degelijk normerend, en verklaart ook waarom het kader door gebruikers in scholen vanuit dat perspectief als normerend gepercipieerd wordt. Indien alle partners van mening zijn dat de inhoudelijke focus van het OK-kader wezenlijk is om onderwijskwaliteit te monitoren, en als men van mening is dat gebruik van het kader kwaliteitsbevorderend werkt, dan rijst de vraag waarom men in de communicatie over het OK-kader niet sterker normerend durft te spreken over inhouden en gebruik van het OK-kader.

Zorg voor een meer naadloze afstemming tussen het OK-kader en het TOK

Uit de resultaten van dit onderzoek blijkt dat de bevroegde schoolleiders het toezichtkader (TOK) van de Vlaamse onderwijsinspectie als een meer functioneel kwaliteitszorginstrument ervaren dan het huidige OK-kader als dusdanig. Hoewel het OK-kader en het daarop gebaseerde TOK verschillende kaders zijn die in de bestaande beleidscontext afzonderlijke doelen dienen, pleiten de onderzoeksresultaten voor een meer verregaande afstemming tussen OK en TOK die verder reikt dan enkel de inhoudelijke afstemming. Een vaststelling is immers dat het inschalen van de eigen onderwijskwaliteit louter met het OK-kader in handen een uitdaging is voor gebruikers. Een andere vaststelling is dan weer dat verschillende instrumenten die verder bouwen op het OK-kader en die door de onderwijsverstrekkers ter ondersteuning van scholen uitgewerkt werden telkens de vorm aannemen van ontwikkelingsschalen zoals beschikbaar in het TOK. Aan de andere kant is het zo dat vanuit de onderwijsinspectie geenszins wordt gesteld dat eigen evaluaties in de scholen aan de hand van de ontwikkelingsschalen dienen te gebeuren. Deze vaststellingen overschouwend stelt zich de vraag of in de verdere onderlinge afstemming van het OK-kader en het toezichtkader niet net een concrete strategie schuilt om de vorige beleidsaanbevelingen te realiseren. Waarom niet inzake inhoud en gehanteerde terminologie voor een naadloze koppeling zorgen tussen beide kaders? En waarom niet van de in het TOK beschikbaar zijnde ontwikkelingsschalen een integraal onderdeel maken van het OK-kader?

Aanbevelingen bij Inspectie 2.0

Waarderend, procedureel en interactief rechtvaardig: Koester de sterktes van Inspectie 2.0

Dat Inspectie 2.0 zich kenmerkt door het opzoeken van een waarderend dialoog met de school is volgens de onderzoeksresultaten één van de sterktes van de nieuwe doorlichtingsaanpak. Procedurele rechtvaardigheid, in Inspectie 2.0 onder meer nagestreefd door het gebruik van datatriangulatie (denk aan gesprekken met ouders en leerlingen) om tot een eindoordeel te komen, lijkt in sterke mate gerealiseerd en bovendien van invloed op de algehele beleving van een doorlichting door de respondenten. Ook de rondgang met de leerlingen en reflectiegesprekken met het schoolteam worden over het algemeen goed onthaald. De inspecteurs en hun aanpak krijgen doorheen het doorlichtingsproces een goed rapport van de bevroegde respondenten. De interactionele rechtvaardigheid (bijvoorbeeld de respectvolle en zorgzame omgang met de inspecteurs) en professionaliteit van de betrokken inspecteurs blijken bovendien betekenisvolle voorspellers in het

begrijpen of in scholen na afloop van een doorlichting al dan niet met inspectiefeedback aan de slag wordt gegaan. Het is dan ook aangewezen om de uitgangspunten van Inspectie 2.0 te koesteren en te blijven inzetten op een professioneel inspectiekorps dat het belang van dialoog en de eigenheid van de doorgelichte school blijft erkennen.

[Wees niet tevreden met de huidige impact, scherp standaarden verder aan](#)

Ondanks de waarderende geluiden bij het proces van de doorlichting leert dit onderzoek dat de feitelijke impact van de doorlichting in de grote meerderheid van de scholen wel degelijk bestaat, maar tegelijk eerder beperkt is. Er is geen uitgesproken effect van de doorlichting op de bestaande inzichten (conceptueel effect) en acties in scholen (instrumenteel effect). Dat zou op zich logisch zijn en weinig problematisch moesten alle andere indicatoren omtrent aspecten van onderwijskwaliteit verder op groen staan. Quod non. De waarderende inschatting van respondenten bij Inspectie 2.0 zou dus ook wel eens een keerzijde kunnen hebben. Mogelijks bracht het streven naar het waarderende karakter van de doorlichtingsaanpak met zich mee dat de verwachte standaarden minder ambitieus dan wenselijk ingevuld werden en wordt er op verbeterpunten die er in scholen vastgesteld worden minder sterk doorgedruwd dan een echt stimulerend effect van de doorlichting vereist. Wij adviseren dan ook als onderwijsveld in het algemeen en als onderwijsinspectie in het bijzonder niet tevreden te zijn met de in dit onderzoek vastgestelde impact. Wij adviseren om de ontwikkelingsschalen tegen het licht te houden en na te gaan of deze werkelijk de minimale standaarden omvatten om in elke school kwaliteitsvol onderwijs te kunnen garanderen. De inschatting of deze standaarden zich al dan niet in de zogenaamde zone van de naaste ontwikkeling van het onderwijsveld bevinden is een argument dat vanuit het perspectief van een controlerend instantie niet te sterk mag doorwegen. Wij adviseren daarnaast om vanuit het principe 'hard in de boodschap, zacht in de relatie' verder in te zetten op het onomwonden vaststellen en benoemen van sterktes, uitdagingen én tekortkomingen in de doorgelichte scholen. Dat kan samengaan met een begripvolle reactie op het relationele. Verder is het aan alle onderwijspartners om van elke school in principe te verwachten dat ze tegemoetkomt aan de vooropgestelde standaarden. Hoeveel begrip er ook kan (en mag) zijn voor de particuliere context van de school, de basisboodschap moet zijn dat elke school geacht wordt aan de verwachtingen te voldoen.

[\(Her\)evalueer en respecteer vervolgens de goed afgebakende opdrachten van inspectie en pedagogische begeleidingsdiensten, en versterk binnen dat kader de samenwerking tussen beide verder](#)

De bevindingen van dit onderzoek wijzen uit dat mondelinge, informele inspectiefeedback (die bijvoorbeeld tijdens het synthesegesprek van een doorlichting wordt gegeven ter verduidelijking van het doorlichtingsverslag) door gebruikers sterk gewaardeerd wordt. Deze bevat volgens schoolleiders namelijk ook concrete adviezen en aanbevelingen om met werkpunten aan de slag te gaan. Schoolleiders en leraren in het onderwijsveld zijn vragende partij om dit nog te versterken. Dat zulke feedback en advies verwacht worden van een doorlichtingsteam en 'geloofwaardiger' zou zijn als deze van de Vlaamse Onderwijsinspectie komen, is mede te wijten aan het spanningsveld dat bestaat tussen de controlerende en stimulerende rol van de Vlaamse Onderwijsinspectie enerzijds en de begeleidende rol van de pedagogische begeleidingsdiensten anderzijds. Er heerst in het werkveld een weinig scherpe kijk op wat de decretaal afgebakende opdrachten van de onderwijsinspectie zijn. Daardoor wordt de onderwijsinspectie meegetrokken in een deels ondersteunende rol, een bevoegdheid die in strikte zin de pedagogische begeleidingsdiensten toekomt.

Vlaamse scholen kennen vooralsnog geen sterke cultuur van interne professionele dialoog en van het betrekken van externen bij kwaliteitszorg. Doorgelicht worden via de principes van Inspectie 2.0 doet hen echter de meerwaarde daarvan aan den lijve ervaren. Het zou ons inziens een verkeerde keuze zijn om mee te stappen in de vraag van het onderwijsveld om die ontwikkelings- en ondersteuningsgerichte elementen te versterken tijdens het doorlichtingsproces. Het is zaak scholen via andere instrumenten

de vruchten van zulke kwaliteitszorginitiatieven te doen plukken, zodat de onderwijsinspectie zich net kan terugplooiën op haar kerntaak, met name controle en de stimulerende rol die daarvan uitgaat. Pedagogische begeleidingsdiensten kunnen hier flankeren om ondersteuning op maat van de school te voorzien. Deze ondersteuning heeft betrekking op de voorbereiding, het doorlopen en het opvolgen van de doorlichting en dus ook op het inbedden van de doorlichting in een ruimer schooleigen kwaliteitszorgsysteem.

Hoewel er een structureel overlegorgaan bestaat tussen beide onderwijspartners, beschouwt het werkveld hen vooralsnog nogal eens als twee los van elkaar fungerende organen. In een ideaal plaatje hebben zij elkaar nodig, en versterken zij elkaar. Het is daarom aangewezen de afgebakende opdrachten van beide onderwijspartners goed te (her)evalueren, en vanuit respect voor dat kader de samenwerking tussen beiden verder te versterken. Een doeltreffende communicatie met het werkveld over wat scholen wel en niet mogen verwachten van beiden is dan noodzakelijk om de contouren van dat partnerschap in de feitelijke uitrol in het Vlaamse onderwijsveld niet onder druk gezet te zien worden.

Zorg als inspectie voor een voldoende frequente vorm van tastbare aanwezigheid op de scholen

Er is in het Vlaamse onderwijsveld veel steun voor het werken met een vorm van doorlichten als onderdeel van het kwaliteitszorgsysteem. De vooropgestelde doorlichtingstermijn van zes jaar is ook breed gedragen. Dit zou echter een strikt minimum moeten zijn, voor elke school. Wanneer er invulling gegeven wordt aan vormen van gedifferentieerd doorlichten dan hebben deze best betrekking op variatie omtrent pas na maximaal zes jaar dan wel sneller opnieuw doorlichten, op variatie in de inhoudelijke breedte van de inspectie of op variatie in intensiteit van het doorlichtingsproces. Uitgaan van het principe dat sommige scholen minder frequent dan om de zes jaar zouden doorgelicht worden is daarbij te vermijden. In het geval van een doorlichting die uitmondt in een gunstig advies met werkpunten zijn respondenten in deze studie vragende partij om de onderwijsinspectie een goed gefaseerde rol te laten spelen in de opvolging van hun doorlichting. Zelfs in het geval van een doorlichting die leidt tot een inschaling boven de verwachting zijn respondenten vragende partij om die termijn niet los te laten, gegeven dat het vooruitzicht op een herhaalde doorlichting na zes jaar in die context prikkels aan het schoolteam blijft geven om de vinger aan de pols te houden en de kwaliteitszorgactiviteiten levendig te houden.

Een onderwijsinspectie stimuleert overigens niet alleen tijdens het moment waarop de school doorgelicht wordt. Ook het uitdragen van een helder en functioneel toezichtkader werkt kwaliteitsbevorderend. Helder houdt in dat opgenomen indicatoren en bijhorende standaarden eenduidig zijn. Functioneel betekent dat scholen tussen doorlichtingen door ook aan de slag kunnen met het toezichtkader. Het eerdere advies om voor een meer naadloze afstemming tussen het OK-kader en het toezichtkader te zorgen, is dus eveneens een manier om als onderwijsinspectie meer tastbaar aanwezig te kunnen zijn op scholen. Dit verder uitwerken kan in scholen mogelijkheden creëren om - als het ware in 'afwezigheid' van de onderwijsinspectie en waar nodig ondersteund door hun begeleidingsdienst - ook zelfstandig maar in de geest van het toezichtkader stappen te zetten inzake borging, verbetering en verandering van aspecten van onderwijskwaliteit. In diezelfde geest van afstemming tussen kaders en tussen kwaliteitszorgactoren willen we het Vlaamse onderwijsveld dan ook oproepen om ernstig te onderzoeken of en op welke manier zelfevaluaties van scholen als aangrijpingspunt en tussenpunt bij doorlichtingen vorm kunnen gegeven worden. Of zo men wil, hoe de doorlichting nog meer op een coherente manier een aangrijpingspunt en tussenpunt kan zijn voor de zelfevaluaties in scholen.

Besef als overheid dat het controleren en stimuleren van onderwijskwaliteit de inzet van voldoende middelen vergt

Een onderwijsinspectie en het goed gefaseerd en uitgerold organiseren van doorlichtingen zijn hefboomen voor onderwijskwaliteit. Deze studie toont aan dat het daarbij van belang is dat een doorlichting zich vertaalt in een goed gedocumenteerd, transparant en eerlijk proces dat ook door de medewerkers in scholen gepercipieerd wordt als valide en betrouwbaar. Daarnaast wordt ook het belang erkend van de professionaliteit van de betrokken inspecteurs. Om met impact te kunnen doorlichten is het dan ook een noodzaak om zowel de frequentie, de minimumduur als de minimuminzet aan personeel van een doorlichting goed te bewaken. Onderwijsinspecteurs zijn als een vliegwiel voor onderwijskwaliteit in het onderwijsveld, op voorwaarde dat zij goed geïnformeerd, goed omkaderd en in werkbare doorlichtingsprocedures kunnen functioneren. Als gedifferentieerd doorlichten te sterk gemotiveerd wordt door schaarste van middelen en zich daardoor noodgedwongen vertaalt in onderwijskundig moeilijk te verantwoorden keuzes zoals een doorlichting te sterk in focus of duur beperken of doorlichtingsteams in omvang verkleinen, dan wordt de potentiële impact uitgehouden.

De uitdagingen inzake doelrealisatie in het Vlaamse onderwijs zijn groot. In tijden van lerarentekort lijkt de populistische lokroep om inspecteurs weer voor de klas te brengen aanlokkelijk. Ons advies is de onderwijsinspectie te blijven zien als een cruciale actor en hun slagkracht in lijn met bovenstaande aanbevelingen nog te versterken, hen uit te rusten met de middelen om die rol effectief waar te kunnen maken en hen vervolgens ook aan te spreken op hun realisaties en inzet van middelen.

Aanbevelingen bij de bronnen van externe kwaliteitszorg

Hou de verschillende indicatoren voor onderwijskwaliteit zo rijk en divers mogelijk

Geen enkele bron van externe kwaliteitszorg kan uitspraken doen over alle aspecten van onderwijskwaliteit op school- en systeemniveau. Elke bron belicht immers een verschillend perspectief. Dit onderzoek toont aan dat die diversiteit erg waardevol is, ook wanneer meerdere bronnen leiden tot een divergerend beeld. Bijgevolg is het belangrijk om de verschillende indicatoren die aan de hand van de verschillende bronnen gemeten worden zo divers en rijk mogelijk te houden. Op die manier kunnen zowel het Vlaamse onderwijs als scholen beroep blijven doen op meerdere bronnen die samen de multidimensionaliteit van onderwijskwaliteit in kaart brengen.

Met oog op bovenstaande vaststelling is het waardevol om in te zetten op bronnen die een absoluut criterium (zoals de doorlichtingen en de peilingen) hanteren én bronnen die een relatief criterium gebruiken (zoals de internationaal vergelijkende onderzoeken en de peilingsonderzoeken). Focus bovendien niet alleen op wiskunde en Nederlands, maar heb ook uitdrukkelijke aandacht voor andere domeinen. Het is aangewezen om te zorgen voor een waaier aan indicatoren die de verschillende aspecten van onderwijskwaliteit in kaart brengen op zowel het systeem- als schoolniveau. Koppel tot slot geen gevolgen aan het resultaat van slechts één bron en stimuleer daarentegen scholen om meerdere bronnen te combineren. Kortom, bewaak dat wat met onderwijskwaliteit bedoeld wordt niet (onbedoeld) verengd wordt en koester de diversiteit aan indicatoren voor onderwijskwaliteit.

Wanneer het kwaliteitsmonitoringssysteem in de toekomst uitgebreid zou worden met bijkomende bronnen, is het waardevol om te bekijken welke aspecten van onderwijskwaliteit in het huidige systeem onderbelicht blijven of minder belicht worden. Vervolgens kunnen er nieuwe bronnen voor externe kwaliteitszorg ontwikkeld worden die specifiek als doel hebben om die onderbelichte aspecten sterker op te volgen (bijvoorbeeld de link tussen toetsen en processen in de klas). Op die manier kunnen de

verschillende bronnen van externe kwaliteitszorg beter op elkaar afgestemd worden om samen de multidimensionaliteit van onderwijskwaliteit in kaart te brengen.

[Communiceer op een heldere, transparante en toegankelijke manier over de verschillende bronnen van externe kwaliteitszorg – en in de toekomst over de Vlaamse toetsen](#)

Als scholen of andere actoren niet weten wat elke bron wel en niet meet, kunnen ze onmogelijk de resultaten van verschillende bronnen interpreteren vanuit het juiste perspectief en ermee aan de slag gaan. Daarom is het aan te bevelen om naar scholen en andere actoren duidelijk te communiceren over het onderzoeksopzet van de verschillende bronnen. Vermeld daarbij wat ze wel en niet kunnen leren van elke bron. Bij feedbackrapporten is het van belang om de resultaten op een bevattelijke, toegankelijke en duidelijke manier te presenteren. Geef daarbij op een laagdrempelige manier aan hoe je de resultaten van een bron wel en niet kan lezen. Dit onderzoek toont dat duiding over het onderzoeksopzet nog belangrijker zal worden bij de feedbackrapporten van de Vlaamse toetsen. Het is cruciaal om steeds mee te geven hoe die resultaten wel en niet geïnterpreteerd kunnen worden en wat er wel en niet geleerd kan worden van die resultaten.

[Leg de resultaten van verschillende bronnen samen](#)

Wanneer scholen slechts één bron van externe kwaliteitszorg gebruiken, biedt hen dat weinig inzicht in mogelijke acties of verbeterstrategieën. Eén bron belicht namelijk slechts een perspectief. Uit dit onderzoek blijkt immers dat geen enkele bron uitspraken kan doen over alle aspecten van onderwijskwaliteit. Door meerdere bronnen – en dus verschillende perspectieven – te combineren, kunnen het Vlaamse onderwijs en scholen vanuit een helikopterperspectief wel tot concrete actiepunten ter verbetering komen.

Daarom is het aanbevolen om de beschikbare data van scholen op één plaats te verzamelen en daarbij te communiceren op een heldere, transparante en toegankelijke manier (zie eerdere aanbeveling). Op schoolniveau is het waardevol om scholen te stimuleren om in dialoog te gaan met collega's, pedagogische begeleidingsdiensten, inspecteurs, ouders... over de resultaten van de bronnen van externe kwaliteitszorg. Moedig scholen daarenboven aan om samen met de pedagogische begeleidingsdiensten de resultaten van meerdere bronnen naast elkaar te leggen en – indien nodig – op basis daarvan verbeteringsstrategieën te formuleren. Op systeemniveau is het aangewezen om te blijven inzetten op deelname aan internationaal vergelijkend onderzoek. Durf bovendien scholen verplichten om deel te nemen aan die internationale onderzoeken. Zo zit er geen bias op de steekproef en krijgen alle scholen op regelmatige basis inzicht in resultaten van een extra bron, naast de doorlichtingen en de Vlaamse toetsen.

[Zet in op datageletterdheid bij alle actoren in het onderwijsveld](#)

Het lezen van de informatie in feedbackrapporten of doorlichtingsverslagen is slechts de eerste stap om aan de slag te gaan met de resultaten van verschillende bronnen. Vaak loopt het echter al mis bij die eerste stap, zoals blijkt uit dit onderzoek. Om doelgericht aan de slag te gaan met die resultaten is het noodzakelijk dat leraren en directies beschikken over de nodige kennis en vaardigheden om die resultaten te interpreteren en te vertalen naar hun eigen onderwijspraktijk. Niet alleen op schoolniveau, maar ook op systeemniveau is die vertaalslag nodig. Organiseer daarom niet alleen professionaliseringsactiviteiten voor leraren en directeuren, maar ook voor medewerkers van de pedagogische begeleidingsdiensten, onderwijsinspectie, overheid enzoverder. Meer concreet zouden die activiteiten zich moeten focussen op het aanreiken van concrete handvaten – bijvoorbeeld een gestructureerd stappenplan van een onderzoekscyclus –, het begrijpen van statistische terminologie en het lezen en interpreteren van visueel weergegeven resultaten. Verder is het aangewezen om te investeren in datageletterdheid tijdens de opleiding van toekomstige schoolleiders en leraren.

Blijf achtergrondkenmerken in rekening brengen om een eerlijke vergelijking tussen scholen te maken

Om de resultaten van verschillende scholen (ongeacht hun leerlingenpubliek) op een eerlijke manier te vergelijken, is het noodzakelijk om de achtergrondkenmerken van leerlingen en school in rekening te brengen. Ook bij de Vlaamse toetsen is dat van belang, zo blijkt uit dit onderzoek. Het gaat immers over hoeveel extra leerwinst een school creëert en niet over welke instroom van leerlingen die school heeft. Bijgevolg is het aanbevolen om statistisch te controleren op achtergrondkenmerken van scholen en leerlingen – bijvoorbeeld door toegevoegde waardemodellen te hanteren – of op een andere manier hiermee rekening te houden – bijvoorbeeld door groepen van leerlingen te vergelijken met gelijkaardige leerlingen in de totale populatie. Van belang is wel dat onderzoekers helder en transparant communiceren over hoe ze die achtergrondkenmerken in rekening brengen, welke modellen ze daarvoor gebruiken en voor welke variabelen ze precies controleren. Bij de resultaten dient steeds duidelijk aangegeven te worden hoe de lezers die resultaten kunnen interpreteren.

Afrondende aanbeveling: Onderwijskwaliteit realiseren is zorgen voor leerkrachten én schoolleiders

Tot slot. Vertrouwd geraken met het OK-kader en het kader ingang doen vinden in de eigen school- en klaspraktijk. Gericht en met vertrouwen toeleven naar een doorlichting en aan de slag gaan met een doorlichtingsverslag. Deelnemen aan peilingsonderzoek (of meer toekomstgericht: aan centrale toetsen) of internationaal vergelijkend onderzoek, de schoolfeedback die deze met zich meebrengen ter harte nemen. Er is één voorwaardenscheppende factor die hiertoe doorheen elk onderdeel van deze studie prominent naar voren komt: het kunnen garanderen van voldoende continuïteit in het leiderschap en het onderwijzend corps van de school. Het samenspel van Inspectie 2.0, het OK-kader en andere externe bronnen van onderwijskwaliteit kan maar succesvol zijn als hefboom voor onderwijskwaliteit als met voldoende en daartoe opgeleide leidinggevendenden, leraren en andere medewerkers een duurzame relatie kan opgebouwd worden. De zorg daartoe moet daarom een absolute prioriteit blijven, als voorwaardenscheppend. Echter ook omgekeerd. Verder inzetten op mechanismen en standaarden voor kwaliteitszorg in de ruime zin, in de klas, binnenschools, bovenschools en op systeemniveau zijn een voorwaarde om van onderwijs een aantrekkelijke werkplek te maken om in te functioneren én om leiding aan te geven. Onderwijskwaliteit realiseren is zorgen voor leerkrachten én schoolleiders.

6. Referentielijst

- Allen, R., & Burgess, S. (2012). *How should we treat under-performing schools? A regression discontinuity analysis of school inspections in England* CMPO Working Paper Series. Bristol, UK: CMPO, University of Bristol.
- Audia, P. G., & Locke, E. A. (2003). Benefiting from negative feedback. *Human Resource Management Review*, 13(4), 631–646.
- AHOVOKS. (2021). *Peilingen: Wat en waarom?* <https://einddoelen.be/wat-en-waarom>
- Brimblecombe, N., & Ormston, M. (1995). Teachers' perceptions of school inspection: A stressful experience. *Cambridge Journal of Education*, 25, 53-61.
- Creswell, J. W. (2005). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (2nd Edition). Thousand Oaks, CA: Sage Publications.
- Denzin, N. K., & Lincoln, Y. S. (2000). *Handbook of qualitative research*. Thousand Oaks, CA: Sage Publications.
- De Maeyer, S., Penninckx, M., Vanhoof, J., & Van Petegem, P. (2014). Exploring and explaining the effects of being inspected. *Educational Studies*, 40(4), 456–472.
- Dockx, J., Van Landeghem, G., Aesaert, K., Van Damme, J., & De Fraine, B. (2019). *Begrijpend lezen van het vierde naar het zesde leerjaar. Herhalingsmeting van PIRLS in 2018 vergeleken met PIRLS 2016*. Centrum voor Onderwijseffectiviteit en –evaluatie, KU Leuven.
- Ehren, M. C. M., & Visscher, A. J. (2008). The relationships between school inspections, school characteristics and school improvement. *British Journal of Educational Studies*, 56(2), 205–227. <https://doi.org/10.1111/j.1467-8527.2008.00400.x>.
- Ehren, M. C. M., Perryman, J., & Shackleton, N. (2015). Setting expectations for good education: How Dutch school inspections drive improvement. *School Effectiveness and School Improvement*, 26, 296–327.
- Elmore, R. F., & Fuhrman, S. H. (2001). Holding Schools Accountable: Is it Working? *Phi Delta Kappan*, 83(1), 67–72.
- Gärtner, H., Wurster, S., & Pant, H. A. (2014). *The effect of school inspections on school improvement*. *School Effectiveness and School Improvement*, 25(4), 489-508.
- Hutjes, J. M., & Van Buuren, J. A. (1992). *De gevalsstudie: strategie van kwalitatief onderzoek*. Meppel: Boom.
- Kelchtermans, G. (2007). Macropolitics caught up in micropolitics: the case of the policy on quality control in Flanders (Belgium). *Journal of Education Policy*, 22(4), 471-491.
- King, F. (2011). The role of leadership in developing and sustaining teacher's professional learning. *Management in Education*, 25(4), 149-155.
- Matthews, P., & Sammons, P. (2004). *Improvement through inspection: An evaluation of the impact of Ofsted's work*. London: Ofsted.
- Miles, M., & Huberman, A. (2014). *Qualitative Data Analysis: An Expanded Sourcebook*. Sage Publications.
- Mullis, I., Martin, M., Foy, P., Kelly, D., & Fishbein, B. (2020). *TIMSS 2019 International Results in Mathematics and Science*. TIMSS & PIRLS International Study Center, IEA.

- Mullis, I., Martin, M., Foy, Pierre., & Hooper, M. (2017). *PIRLS 2016 International Results in Reading*. TIMSS & PIRLS International Study Center, IEA.
- OECD. (2019b). *PISA 2018 Results (Volume I): What Students Know and Can Do*. OECD Publishing.
- Onderwijsinspectie. (z.d.). *Het Concept Inspectie 2.0*. Vlaams Ministerie van Onderwijs en Vorming. Geraadpleegd op 2 juli 2021 van <https://www.onderwijsinspectie.be/nl/het-concept-inspectie-20>.
- Penninckx, M. (2015). *Inspecting school inspections*. Niet gepubliceerd doctoraatsproefschrift: Antwerpen, UA.
- Penninckx, M., Vanhoof, J., De Maeyer, S., & Van Petegem, P. (2015). Effects and side effects of Flemish school inspection. *Educational Management, Administration and Leadership*, 44(5), 728–744.
- Quintelier, A. (2020). *School Inspection Feedback: Unravelling Determinants and Consequences of Teachers' Feedback Acceptance*. Niet gepubliceerd doctoraatsproefschrift: Antwerpen, UA.
- Scheerens, J. (2011). *Measuring Educational Quality by Means of Indicators*. In J. Scheerens, H. Luyten, & van Ravens, J. (Eds.), *Perspectives on Educational Quality* (pp. 3-34). Dordrecht: Springer.
- Schildkamp, K., Lai, M., & Earl, L. (2013). *Data-based Decision Making in Education. Challenges and Opportunities*. Springer.
- Sun, H., Creemers, B.P.M., & de Jong, R. (2007). Contextual factors and effective school improvement. *School Effectiveness and School Improvement*, 18(1), 93-122.
- Swanborn, P. (2010). *Casestudy Research: What, Why and How?* Thousand Oaks, CA: Sage Publications.
- Tashakkori, A., & Teddlie, C. (2003). *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks, CA: Sage Publications.
- Treviño, E., & Órdenes, M. (2017). *Exploring Commonalities and Differences in Regional and International Assessments*. Information Paper No. 48, UNESCO.
- Vlaams Ministerie van Onderwijs en Vorming, & Vlaamse Onderwijsinspectie. (2017). *Bronnendocument Referentiekader voor Onderwijskwaliteit*.