

Evaluatie van de peilingen in het Vlaams onderwijs: een terug- en vooruitblik

Departement Onderwijs en Vorming



Vlaanderen
is onderwijs & vorming



Vlaanderen
is onderwijs & vorming

INHOUD

1. Hoofdstuk 1: Context, situering en verantwoording.....	5
1.1. <i>Peilingen binnen een systeem met centraal geformuleerde doelstellingen en een schooleigen evaluatiebeleid.....</i>	<i>5</i>
1.2. <i>Het peilingsonderzoek als antwoord op de nood aan een betere kennis over het behalen van de eindtermen en ontwikkelingsdoelen.....</i>	<i>6</i>
1.3. <i>Algemene schets van het peilingsonderzoek in Vlaanderen.....</i>	<i>7</i>
1.4. <i>Het peilingsonderzoek als instrument in een kennisgedreven beleidsvoorbereiding.....</i>	<i>8</i>
1.5. <i>Aanleiding, opzet en doelstelling van de beleidsevaluatie naar het peilingsonderzoek in Vlaanderen.....</i>	<i>10</i>
1.5.1. <i>Beleidskader.....</i>	<i>10</i>
1.5.2. <i>Ex post beleidsevaluatie van het steunpunt Toetsontwikkeling en Peilingen</i>	<i>12</i>
1.5.3. <i>Ex ante beleidsevaluatie naar toekomstige modellen voor peilingsonderzoek in Vlaanderen.....</i>	<i>12</i>
1.5.4. <i>Auteurs van dit rapport.....</i>	<i>14</i>
2. Hoofdstuk 2: Opzet van het peilingsonderzoek.....	15
2.1. <i>Inleiding.....</i>	<i>15</i>
2.2. <i>Het proces van een peiling.....</i>	<i>15</i>
2.2.1. <i>Toetsontwikkeling.....</i>	<i>16</i>
2.2.2. <i>Peilingsonderzoek.....</i>	<i>19</i>
2.3. <i>Structuur, werking en opvolging van het steunpunt.....</i>	<i>23</i>
2.3.1. <i>Structuur en werking.....</i>	<i>24</i>
2.3.2. <i>Opvolging.....</i>	<i>24</i>
3. Hoofdstuk 3: Evaluatie van de lopende Beheersovereenkomst (2013-2017).....	26
3.1. <i>Situering.....</i>	<i>26</i>
3.2. <i>Beheersovereenkomst, werking, sturing & samenwerking.....</i>	<i>26</i>
3.2.1. <i>De beheersovereenkomst.....</i>	<i>26</i>
3.2.2. <i>De interne werking, sturing & samenwerking.....</i>	<i>27</i>
3.3. <i>Opdrachten uit de beheersovereenkomst.....</i>	<i>29</i>
3.3.1. <i>Toetsontwikkeling.....</i>	<i>29</i>
3.3.2. <i>Afname van peilingen</i>	<i>30</i>
3.3.3. <i>Dataverwerking en –analyse.....</i>	<i>31</i>
3.3.4. <i>Rapportering en feedback</i>	<i>32</i>
3.3.5. <i>Vorming en valorisatie</i>	<i>32</i>
3.3.6. <i>Wetenschappelijke dienstverlening.....</i>	<i>34</i>

3.3.7 Ontwikkeling van specifieke expertise.....	34
4. Hoofdstuk 4: Ex ante evaluatie toekomst van de peilingen	36
4.1. <i>Inleiding: situering en methode.....</i>	<i>36</i>
4.2. <i>Aan welke doelstellingen kunnen de peilingen bijdragen?.....</i>	<i>37</i>
4.2.1. Kwaliteit van het Vlaams onderwijssysteem bewaken en versterken.....	38
4.2.2. De interne kwaliteitszorg van scholen versterken	39
4.2.3. De kwaliteit van scholen bewaken (inspectie).....	41
4.2.4. Versterken van de geïnformeerde schoolkeuze	42
4.2.5. Verhogen van de gelijkwaardigheid van diploma's en getuigschriften	43
4.2.6. Versterken en ondersteunen van de lerende	44
4.3. <i>Hoe kan de peiling als instrument er in de toekomst uitzien?.....</i>	<i>45</i>
4.3.1. Hoe verhouden de peilingen zich tot andere toetsen?	45
4.3.2. Wie neemt deel aan de peilingen?.....	46
4.3.3. Wat moeten peilingen meten?	49
4.3.4. Op welke momenten moet er gepeild worden en met welke frequentie?	51
4.3.5. Naar digitale toetsen?	52
4.3.6. Welke feedback krijgt men na deelname aan de peiling?	52
4.4. <i>Wat zijn noodzakelijke randvoorwaarden om de doelstellingen te bereiken?.....</i>	<i>54</i>
Hoofdstuk 5: Algemene conclusies	57
5.1 <i>Situering.....</i>	<i>57</i>
5.2 <i>Doelstellingen.....</i>	<i>57</i>
5.3 <i>Gebruik en openbaarheid van schoolfeedbackrapporten</i>	<i>60</i>
5.4 <i>Meer en bredere communicatie</i>	<i>60</i>
5.5 <i>Inhoud, tijdstip en frequentie van het peilingsonderzoek.....</i>	<i>61</i>
5.6 <i>Beheer en aansturing.....</i>	<i>61</i>
Bijlage 1: Peilingenkalender 2002-2019.....	63
Bijlage 2: Methode van de ex ante evaluatie.....	65

1. HOOFDSTUK 1: CONTEXT, SITUERING EN VERANTWOORDING

1.1. Peilingen binnen een systeem met centraal geformuleerde doelstellingen en een schooleigen evaluatiebeleid

Sinds midden jaren 1990 worden de leerinhouden van het Vlaamse onderwijs bepaald door eindtermen en ontwikkelingsdoelen. De eindtermen zijn de minimumdoelen die leerlingen in een bepaalde fase van hun schoolloopbaan moeten bereiken. Zo zijn er eindtermen voor het einde van het lager onderwijs en per graad in het secundair onderwijs (uitgezonderd de eerste graad B-stroom). Een school moet er naar streven met om met een meerderheid van haar leerlingen de door het Vlaams Parlement vastgelegde eindtermen te behalen. Ontwikkelingsdoelen zijn minimumdoelen die leerlingen niet noodzakelijk moeten bereiken, maar die de school wel moet nastreven. Ze worden gebruikt in het kleuter- en buitengewoon onderwijs en in de eerste graad van de B-stroom in het secundair onderwijs. Sinds 2008 kan de Vlaamse overheid ook beroeps- en onderwijskwalificaties bepalen die bestaan uit een afgerond en ingeschaald geheel van competenties. Deze kwalificaties maken deel uit van de Vlaamse kwalificatiestructuur die momenteel nog in volle ontwikkeling is.

Met het centraal formuleren van eindtermen, ontwikkelingsdoelen en kwalificaties wil de Vlaamse overheid de gelijkwaardigheid van diploma's en studiebewijzen en ook de onderlinge gelijkwaardigheid tussen onderwijsinstellingen garanderen. Daarnaast vormen deze onderwijsdoelen ook een uiting van wat de samenleving van het onderwijs verwacht dat ze minimaal met haar leerlingen moet nastreven. In deze zin vormen de eindtermen, ontwikkelingsdoelen en kwalificaties de basis voor het kwaliteitszorgbeleid in onderwijs en vormen ze voor de onderwijsinspectie een belangrijk aspect voor de externe kwaliteitscontrole van scholen.

Anders dan vele andere Europese landen koos Vlaanderen er niet voor om het al dan niet behalen van deze centraal geformuleerde onderwijsdoelen door scholen na te gaan via centrale toetsen (Van Hoof et al., 2016¹, Standaert, 2014², OECD, 2013³). Vlaamse scholen hebben een grote autonomie om een eigen visie op kwaliteitsvol onderwijs en een specifiek pedagogisch project uit te tekenen. Ook het evaluatiebeleid behoort tot die Grondwettelijk verankerde pedagogische vrijheid die Vlaamse scholen genieten. Dit

¹ Van Hoof, J. (2016). Scenario's voor leer(winst)monitoring in Vlaanderen: Een ontwerponderzoek naar haalbaarheid en wenselijkheid. Eindrapport OBPWO 13.03. (zie <http://data-onderwijs.vlaanderen.be/onderwijsonderzoek/default.aspx?nr=189>)

² Standaert, R. (2014). De becijferde school. Meetcultus en meetcultuur. Leuven: Acco.

³ OECD (2013). Synergies for better learning. An international perspective on evaluation and assessment. An International Perspective on Evaluation and Assessment. OECD Publishing.

impliceert dat scholen en leerkrachten zelf nagaan in welke mate de leerlingen de eindtermen of ontwikkelingsdoelen al dan niet bereikt hebben. De onderwijsinspectie zal n.a.v. een doorlichting nagaan in welke mate de scholen deze maatschappelijke verwachting effectief realiseren, waarbij scholen die niet of in sterk onvoldoende mate het bereiken van de eindtermen bij hun leerlingen nastreven, de erkenning door de Vlaamse overheid kunnen verliezen.

1.2. Het peilingsonderzoek als antwoord op de nood aan een betere kennis over het behalen van de eindtermen en ontwikkelingsdoelen

Vrij snel na de invoering van de eindtermen en ontwikkelingsdoelen in het Vlaamse onderwijs ontstond bij de overheid de nood om, naast de informatie die de onderwijsinspectie verzamelde n.a.v. de doorlichtingen van scholen, een meer algemeen beeld te krijgen van de onderwijsprestaties van leerlingen. De keuze om niet met centrale examens te werken en het evaluatiebeleid over te laten aan de pedagogische vrijheid van scholen, maakte het moeilijk om een goed beeld te krijgen van de mate waarin de leerlingen in Vlaanderen nu effectief de vooropgestelde eindtermen bereiken. Het is in die context dat Vlaanderen al eind de jaren 1990 mee participeerde aan grootschalig internationaal comparatieve prestatiemetingen zoals het 'Programme for International Student Assessment' (PISA) van de Organisatie voor Economische Samenwerking en Ontwikkeling (OESO) of het 'Trends in International Mathematics and Science Study' (TIMSS) van de International Association for the Evaluation of Educational Achievement (IEA). Deze onderzoeken geven wel een beeld van hoe het Vlaamse onderwijs in vergelijking met andere onderwijssystemen presteert op een aantal domeinen zoals bijvoorbeeld leesvaardigheid, wiskundige en wetenschappelijke geletterdheid, maar zegt niets over de mate waarin de leerlingen ook de eigen eindtermen behalen.

Om toch een beter zicht te krijgen op het bereiken van de eindtermen door leerlingen startte de Vlaamse overheid in 2002 met het peilingsonderzoek. Dat gebeurde voor het eerst in het kader van het Onderwijskundig Beleids- en Praktijkgericht Wetenschappelijk Onderzoek (OBPWO). Het peilingsonderzoek bleef tot 2006 gekaderd binnen het OBPWO-programma, maar kreeg steeds meer een systematisch en periodiek karakter. Daarom werd het peilingsonderzoek vanaf 2006 via een meerjarige kaderovereenkomst gegund aan de KU Leuven. Om de continuïteit van de peilingen verder te garanderen, besliste de Vlaamse Regering in 2012 om een universitair steunpunt voor Toetsontwikkeling en Peilingen op te richten. Dit steunpunt ging van start op 1 januari 2013 en werd gegund aan de KU Leuven, vakgroep onderwijseffectiviteit en -evaluatie, met prof. Dr. Rianne Janssens als promotor-coördinator. De huidige overeenkomst met het steunpunt voor Toetsontwikkeling en Peilingen loopt af op 31 december 2017. De werking van het huidige steunpunt wordt meer in detail toegelicht in hoofdstuk 2.

1.3. Algemene schets van het peilingsonderzoek in Vlaanderen

Een peiling is een grootschalige afname van toetsen. Anders dan bij centrale toetsen waar alle scholen deelnemen, worden de peilingen afgenomen bij een representatieve steekproef van scholen en leerlingen. De peiling bestaat hoofdzakelijk uit de afname van schriftelijke toetsen, maar er wordt ook gebruik gemaakt van observatie en praktische proeven voor de eindtermen die moeilijk schriftelijk te toetsen zijn. De keuze om niet alle scholen te testen, maar te werken met een representatieve steekproef, maakt het net mogelijk om ook eindtermen te testen via observatie en praktische proeven, wat ons een rijker beeld oplevert.

Naast de toetsen, observaties en praktische proeven, worden ook vragenlijsten afgenomen bij leerlingen, ouders en leerkrachten. Hiermee wordt informatie verzameld over o.a. attitudes en de motivatie van leerlingen, de gehanteerde klaspraktijken of didactische aanpak en de mate dat de eindtermen effectief behandeld werden in de les. De koppeling van deze informatie met de prestaties van de leerlingen op de testen, laat toe na te gaan of verschillen tussen scholen in de mate hun leerlingen de eindtermen halen, samenhangen met bepaalde school- of leerlingenkenmerken. Het laat de onderzoekers ook toe om mogelijke verklaringen te vinden voor de resultaten en aan de overheid en scholen geeft het informatie over welke factoren een impact hebben op het al dan niet bereiken van de minimumdoelen door leerlingen.

Er worden elk jaar twee peilingen afgenomen, telkens voor één leergebied in het basisonderwijs en voor één vak in het secundair onderwijs. De peilingen in het secundair onderwijs worden in hoofdzaak afgenomen in ofwel het einde van de eerste graad ofwel het einde van de derde graad. In de derde graad kan dit beperkt blijven tot één onderwijsvorm, dan wel alle onderwijsvormen omvatten.

Voor sommige leergebieden of vakken werden er al meerdere peilingen georganiseerd. Deze zogenaamde herhalingspeilingen laten de overheid en scholen toe om peilingsresultaten doorheen de tijd met elkaar te vergelijken en mogelijke trends op het vlak van het behalen van deze eindtermen en de onderwijskwaliteit in kaart te brengen. Zoals al aangegeven is het peilingsonderzoek periodiek van aard. Om er voor te zorgen dat alle relevante leergebieden en vakken aan bod komen doorheen de tijd, om de samenhang of mogelijke overlap met het internationaal vergelijkend onderzoek op te vangen en om op systematische wijze bepaalde peilingen te herhalen en zo trends te kunnen detecteren, wordt er gewerkt met een peilingenkalender over meerdere jaren heen. De huidige peilingenkalender werd door de minister van Onderwijs bepaald tot en met 2019 (zie Bijlage 1).

De peilingen leveren resultaten op systeemniveau op die de Vlaamse overheid een

algemeen beeld geven van de kwaliteit van het onderwijs, de mate dat de leerlingen de eindtermen al dan niet behalen en de samenhang van deze resultaten met bepaalde school- en leerlingenkenmerken. Daarnaast ontvangen de deelnemende scholen een schoolfeedbackrapport op basis van de eigen resultaten waarmee ze aan de slag kunnen in het kader van de eigen interne kwaliteitszorg. Enkel de scholen hebben toegang tot hun eigen schoolfeedbackrapport. De overheid of de onderwijsinspectie kunnen deze niet inkijken of gebruiken. Daarmee wordt uitgesloten dat scholen onterecht met elkaar vergeleken zouden worden of dat er vormen van ranking van scholen zouden ontstaan wat vaak wel het geval is bij centrale toetsing. De peilingen laten ook niet toe om de prestaties van individuele leerlingen te meten en dus heeft het resultaat van een peilingstoeets geen enkele impact op de schoolloopbaan van de deelnemende leerling. Evenmin kunnen op basis van een peiling uitspraken gedaan worden over het functioneren van een individuele leerkracht.

De paralleltoetsen vormen een afgeleid product van de peilingstoeetsen. Ze worden eveneens ontwikkeld door het steunpunt Toetsontwikkeling en Peilingen. Het zijn toetsen die hetzelfde meten als de peilingstoeetsen maar met andere, gelijkaardige items en waar scholen vrijwillig en kosteloos gebruik van kunnen maken. Deelnemende scholen krijgen eveneens een schoolfeedbackrapport waarmee ze nadien aan de slag kunnen gaan om het eigen schoolbeleid te versterken, al dan niet samen met de pedagogische begeleidingsdienst. In die zin passen zowel de peilingstoeetsen als de paralleltoetsen in het creëren van een informatierijke omgeving voor scholen in het kader van de interne kwaliteitszorg (zie infra). Vanaf 2014 staat het steunpunt Toetsontwikkeling en Peilingen niet alleen in voor de ontwikkeling van de paralleltoetsen, maar heeft het ook als opdracht om deze ter beschikking te stellen aan scholen, de resultaten te verwerken, de schoolfeedbackrapporten op te maken en in te staan voor alle administratief-organisatorische vragen.

In hoofdstuk 2 wordt een meer technische toelichting gegeven over de opmaak, afname, verwerking en analyse van het peilingsonderzoek en wordt ook meer toelichting gegeven bij de gehanteerde begrippen verder in dit rapport.

1.4. Het peilingsonderzoek als instrument in een kennisgedreven beleidsvoorbereiding

Het beleidsdomein Onderwijs en Vorming heeft een lange traditie opgebouwd om beleidskeuzes en de beleidsvoorbereiding mee te voeden met kennis verkregen via wetenschappelijk onderzoek. Hiermee wordt een internationale trend gevolgd naar een meer door onderzoek geïnformeerd en onderbouwd beleid. Het peilingsonderzoek past binnen dit kennisbeleid en vormt een belangrijk instrument om over een aantal aspecten van het onderwijsbeleid interessante kennis te vergaren.

Het peilingsonderzoek geeft kennis over volgende aspecten:

- In welke mate behalen de leerlingen in het Vlaamse onderwijs effectief de eindtermen of ontwikkelingsdoelen?
- Zijn er belangrijke verschillen tussen scholen in het al dan niet behalen van de eindtermen door hun leerlingen en blijven die verschillen bestaan na controle op bepaalde school- en leerlingenkenmerken?
- In welke mate zijn bepaalde leerling-, klas- of schoolkenmerken bepalend voor de mate dat leerlingen de eindtermen al dan niet bereiken?
- Zijn er doorheen de tijd trends of evoluties te onderscheiden in de mate dat leerlingen de eindtermen bereiken?

Het peilingsonderzoek levert ons een bijzonder rijk en gevarieerd pakket aan kennis op dat nadien kan gevaloriseerd worden om zo tot een sterker en meer kwaliteitsvol onderwijs te komen. Dit zowel op het niveau van de centrale overheid, de intermediaire organisaties zoals bijvoorbeeld de pedagogische begeleidingsdiensten, als voor de scholen zelf.

Het peilingsonderzoek kan beschouwd worden als een vorm van ex post beleidsevaluatie van de eindtermen en ontwikkelingsdoelen. Op basis van de resultaten en de inzichten die dit onderzoek oplevert, kan beslist worden om de eindtermen en ontwikkelingsdoelen te behouden, dan wel te actualiseren of verder te verfijnen. Daarnaast vormen de resultaten van het peilingsonderzoek een belangrijke indicator voor de algemene kwaliteit van het Vlaams onderwijs. Sinds de scholen door het kwaliteitsdecreet van 2009 de eerste verantwoordelijke zijn voor het realiseren van kwaliteitsvol onderwijs en hiertoe een intern kwaliteitszorgbeleid dienen uit te bouwen, levert het peilingsonderzoek op systeemniveau kennis over de mate waarin deze doelstelling effectief gerealiseerd wordt alsook welke factoren een impact hebben op het versterken van die kwaliteit.

Ondersteunende organisaties, zoals de koepels van inrichtende machten of pedagogische begeleidingsdiensten, maar ook bijvoorbeeld lerarenopleidingen, kunnen met de resultaten van het peilingsonderzoek aan de slag om de eigen dienstverlening te versterken, de middelen gericht in te zetten om mogelijke pijnpunten weg te werken en kunnen ook de eigen instrumenten meer ontwikkelen op basis van wetenschappelijk opgebouwde kennis en inzichten.

Scholen die deelnemen aan het peilingsonderzoek, maar ook scholen die gebruik maken van de paralleltoetsen, kunnen de kennis die ze meekrijgen in het eigen schoolfeedbackrapport gebruiken om het eigen schoolbeleid te versterken, krijgen een beter zicht op mogelijke te versterken aspecten van de schoolwerking en kunnen meer kennisgedreven beleidsinterventies doorvoeren.

Omdat het peilingsonderzoek voor zoveel onderwijsactoren belangrijke kennis oplevert, wordt elk onderzoek gevolgd door een intensief en laagdrempelig valorisatietraject. Dat start bij de vrijgave van de onderzoeksresultaten door de minister met een brede communicatie naar scholen toe via een toegankelijke brochure en een website. De resultaten worden bekend gemaakt op een colloquium waarop in de eerste plaats de deelnemende scholen worden uitgenodigd. Het is een eerste kanaal om de resultaten van de peiling terug te koppelen aan de onderwijspraktijk.

Ook wordt na elke peiling een werkseminarie georganiseerd waarin er met verschillende partners (koepels, GO!, pedagogische begeleiders, onderwijsinspectie en uitgeverijen) wordt nagedacht over de peilingsresultaten en mogelijke acties die kunnen ondernomen worden.

Finaal gaat elke onderwijspartner aan de slag met deze inzichten om binnen de eigen werking en verantwoordelijkheden acties op te zetten die effectief leiden tot een versterking van de kwaliteit van het onderwijs. Voor de overheid kan dit bijvoorbeeld betekenen dat de eindtermen herzien worden. Maar ook van leerplanmakers, begeleiders en nascholers, opleiders, uitgeverijen, de onderwijsinspectie, directies en leraren wordt verwacht dat ze de nieuwe inzichten die het peilingsonderzoek ons oplevert mee vertalen in het eigen werk.

1.5. Aanleiding, opzet en doelstelling van de beleidsevaluatie naar het peilingsonderzoek in Vlaanderen

1.5.1. Beleidskader

In het regeerakkoord 2014-2019 van de huidige Vlaamse Regering lezen we het volgende over het peilingsonderzoek: *“Met de onderwijsverstreckers wordt overeengekomen om de huidige OVSG-toetsen en de interdiocesane toetsen [sic] te valideren in samenspraak met het Steunpunt Toetsontwikkeling [sic], dat de peilingsproeven en paralleltoetsen ontwikkelt. Op die manier kan een toets toolkit ontstaan. Met de realisatie van dezelfde doelstelling voor ogen voeren we gesprekken met de onderwijsverstreckers van het secundair onderwijs en bekijken we hoe de bestaande peilingsproeven daartoe een basis kunnen vormen”*.

In de beleidsnota Onderwijs 2014-2019 wordt dit verder geconcretiseerd als volgt: *“Zoals voorzien in het masterplan secundair onderwijs worden alle leerlingen op het einde van het basisonderwijs getoetst. Met de onderwijsverstreckers ga ik na hoe we de huidige toetsen van het Onderwijssecretariaat van de Steden en Gemeenten van de Vlaamse*

Gemeenschap (OVSG) en de interdiocesane proeven kunnen laten valideren door het Steunpunt Toetsontwikkeling en Peilingen. Vertrekkend van deze toetsen en de paralleltoetsen die voortkomen uit de peilingsproeven, maak ik werk van een 'toolkit met gevalideerde toetsen' waar scholen naar eigen keuze gebruik van kunnen maken in het kader van interne kwaliteitszorg en heeft ook tot doel gegevens te verzamelen op systeemniveau. Met de onderwijsverstrekkers van het secundair onderwijs bekijk ik of deze werkwijze ook voor hen kan worden opgezet. Wat betreft toetsen die uitspraak doen over prestaties van individuele leerlingen, vertrekken we van het principe dat het evaluatiebeleid blijft behoren tot de volledige autonomie van de school. "

In het kader van de operationele doelstelling om de kwaliteit te versterken door informatierijke omgevingen te creëren staat ook volgende passage in de beleidsnota: *"De resultaten van de peilingen en het internationaal vergelijkend onderzoek zijn een belangrijke indicator voor de kwaliteit van ons onderwijs en de resultaten ervan vragen dus om een meer grondige analyse en debat met alle betrokken partijen. Ik zet het huidige systeem van werkseminaries met belanghebbenden verder. Waar nodig zet ik meer diepgaand onderzoek op om bepaalde resultaten te verklaren en beleidsmatig te vertalen. Ik nodig alle betrokken stakeholders uit om naar aanleiding van nieuwe resultaten concrete verbeteracties te ondernemen."*

In de beleidsbrief Onderwijs 2016-2017 staat ten slotte volgende passage: *"De huidige beheersovereenkomst van het steunpunt Toetsontwikkeling en Peilingen loopt af eind 2017. Zoals vermeld in de conceptnota van de modernisering van het secundair onderwijs wil ik blijven inzetten op het systeem van de periodieke peilingen, om op systeemniveau te kunnen nagaan in welke mate de eindtermen en ontwikkelingsdoelen worden bereikt. Daarom zal ik na de evaluatie van de werking en resultaten van het huidige steunpunt een nieuwe oproep voor het steunpunt Toetsontwikkeling en Peilingen lanceren om de continuïteit van het systeem van het systeem te verzekeren."*

Zoals aangegeven in de beleidsbrief Onderwijs 2016-2017, loopt de huidige overeenkomst met het steunpunt Peilingen en Toetsontwikkeling eind 2017 af. Zowel het in het regeerakkoord, de beleidsnota Onderwijs als in de laatste beleidsbrief Onderwijs is voorzien in een continuering van het steunpunt Toetsontwikkeling en Peilingen na 2017.

Dit rapport vormt de voorbereiding voor de opmaak van het bestek voor de gunningsprocedure voor het steunpunt vanaf 1 januari 2018. Daarbij wordt vooreerst vertrokken van een grondige ex post evaluatie van de werking van het lopende steunpunt Toetsontwikkeling en Peilingen, de scope van het huidige peilingsonderzoek, de wijze waarop dat onderzoek vandaag methodologisch is opgezet als de beleidsrelevantie en de valorisatie van de resultaten.

1.5.2. Ex post beleidsevaluatie van het steunpunt Toetsontwikkeling en Peilingen

Het ex post onderzoek is qua methodologie geïnspireerd door de evaluatie van de vorige kaderovereenkomst 2006-2013 en door de evaluatie van het steunpunt Studie- en Schoolloopbanen dat door het departement Economie, Wetenschappen en Innovatie (EWI) werd uitgevoerd i.s.m. Idea Consult. Dit onderzoek omvat uit drie luiken. Het eerste luik betreft een zelfevaluatie door het onderzoeksteam van het steunpunt Toetsontwikkeling en Peilingen. Daarnaast is er een peer evaluatie door prof. Dr. Daniël Muys van de University of Southampton (UK) en prof. Em. Dr. Norman Verhelst, momenteel werkzaam bij het statistisch consultatiebureau Eurometrics, maar voordien verbonden aan de Universiteit van Twente en het Nederlands Centraal Instituut voor Toetsontwikkeling (CITO). Het derde luik is een evaluatie door de opdrachtgever, vertegenwoordigt door het departement Onderwijs en Vorming, het agentschap voor Hoger Onderwijs, Volwassenenonderwijs, Kwalificaties en Studietoelagen (Ahovoks) en de onderwijsinspectie.

Meer achtergrond bij de gehanteerde methodologie, de resultaten van deze ex post evaluaties en de conclusies die hieruit getrokken kunnen worden, kan in hoofdstuk 3 gelezen worden.

1.5.3. Ex ante beleidsevaluatie naar toekomstige modellen voor peilingsonderzoek in Vlaanderen

Er is geopteerd om niet enkel een ex post beleidsevaluatie uit te voeren ter voorbereiding van de oproep voor het nieuwe steunpunt, maar ook om te kijken naar de impact van mogelijke beleidsontwikkelingen op het toekomstige model voor het peilingsonderzoek in Vlaanderen. Er zijn immers een aantal ontwikkelingen die mogelijk de scope en het karakter van het peilingsonderzoek in belangrijke mate kunnen beïnvloeden.

Vooreerst is er de modernisering van het secundair onderwijs en meer specifiek daarin de beslissing van de Vlaamse Regering dat alle basisscholen op het einde van het lager onderwijs al hun leerlingen moeten toetsen in minstens twee verschillende leergebieden vanaf schooljaar 2017-2018 en in drie verschillende leergebieden vanaf het schooljaar 2018-2019. In het kader van deze beleidsmaatregel ging het steunpunt Toetsontwikkeling en Peilingen de inhoudsvaliditeit na van de OVSG-toetsen en de interdiocesane proeven van Katholiek Onderwijs Vlaanderen. De conclusie was dat deze toetsen voldoende inhoudsvalide waren in relatie tot de vooropgestelde doelstelling. Maar daarnaast wordt aan de basisscholen in Vlaanderen ook de mogelijkheid geboden om gebruik te maken van de paralleltoetsen om hun leerlingen op het einde van het lager onderwijs te testen.

Deze nieuwe beleidsmaatregel zal tot gevolg hebben dat veel meer scholen dan vandaag gebruik zullen maken van de paralleltoetsen. Daarnaast zal het ook meer dan vandaag belangrijk zijn dat er paralleltoetsen zijn voor voldoende verschillende leergebieden en dat ze ook op regelmatige basis vernieuwd worden.

Een andere evolutie in het kader van de modernisering van het secundair onderwijs met een directe impact op het peilingsonderzoek, is de invoering van het concept 'basisgeletterdheid', gekoppeld aan de verwachting dat deze eindtermen - anders dan wat we nu kennen - door elke leerling individueel moeten bereikt worden. Hoewel er op het moment dat dit rapport tot stand kwam, nog geen duidelijkheid was over de verdere invulling van dit begrip basisgeletterdheid, spreekt het voor zich dat de implementatie van dit concept een grote impact zal hebben op het peilingsonderzoek.

Een tweede belangrijk beleidsproject met een mogelijk impact op het peilingsonderzoek is de herziening van de eindtermen. Na een uitgebreide maatschappelijke en wetenschappelijke consultatie over de eindtermen van de toekomst, werkt het Vlaams Parlement aan een kader voor een herziening van de eindtermen. Daarbij gaat het zowel over de inhoud van de eindtermen, het statuut van die eindtermen als over de procedure waarop ze tot stand komen. Omdat het peilingsonderzoek de eindtermen als voorwerp heeft, zullen de keuzes die in het kader van het eindtermendebat gemaakt worden, een heel belangrijke impact hebben op hoe het peilingsonderzoek in de toekomst zal verlopen.

Een derde belangrijke evolutie is de totstandkoming van een referentiekader voor onderwijskwaliteit (ROK) door de onderwijsverstreckers en onderwijsinspectie. Het referentiekader beoogt de kwaliteitsverwachtingen duidelijker te expliciteren en wil een grotere eenheid creëren in het denken en spreken over onderwijskwaliteit. Het is ook de bedoeling om een betere aansluiting te realiseren tussen het interne kwaliteitszorgbeleid van scholen en de externe kwaliteitscontrole door de inspectie. Aangezien het peilingsonderzoek vandaag ons heel wat kennis oplevert over de kwaliteit van het onderwijs op systeemniveau en omdat deelnemende scholen via het schoolfeedbackrapport heel wat relevante kennis krijgen over de eigen resultaten om te gebruiken in het kader van de interne kwaliteitszorg, zal ook het nieuwe referentiekader zijn impact hebben op het toekomstige peilingsonderzoek.

Een laatste element in dit kader vormen twee OBPWO-onderzoeken die tot nieuwe inzichten en ideeën leiden over hoe het peilingsonderzoek in de toekomst vorm kan krijgen. Het betreft het onderzoek *"Hoe zijn competenties grootschalig te toetsen? Ontwikkeling van een evaluatiematrix voor toetsprogramma's en een inventarisatie van good practices."* en het onderzoek *"Scenario's voor leer(winst)monitoring in Vlaanderen: een ontwerponderzoek naar haalbaarheid en wenselijkheid."*, beiden uitgevoerd door de onderzoeksgroep EduBRon van de Universiteit Antwerpen.

Op basis van deze elementen is er voor geopteerd om naast de ex post evaluatie van het steunpunt Toetsontwikkeling en Peilingen, ook een ex ante beleidsevaluatie uit te voeren waarbij wordt nagegaan wat mogelijke toekomstige opties zijn voor het peilingsonderzoek, welke doelstellingen we hiermee willen bereiken, wat mogelijke positieve effecten of risico's zijn die verbonden worden aan een bepaalde keuze. Anders dan bij de ex post beleidsevaluatie, vertrekt deze ex ante beleidsevaluatie van de inzichten, opinies en belevingen van mensen uit de onderwijspraktijk, m.n. schoolleiders, leerkrachten, pedagogisch begeleiders, ouders, leerlingen, experts, de onderwijsinspectie en beleidsmakers. Hierbij werd niet enkel gefocust op het leerplichtonderwijs, maar werden ook het volwassenenonderwijs en het hoger onderwijs betrokken.

Meer informatie over de gehanteerde methodologie van deze ex ante beleidsevaluatie, de resultaten en conclusies, kan u in hoofdstuk 4 lezen.

In het laatste hoofdstuk 5 worden de resultaten van zowel de ex post als de ex ante beleidsevaluatie samengebracht. Op basis van deze conclusies zullen beleidsaanbevelingen aan de minister worden overgebracht m.h.o. op de opmaak van een voorstel van oproep voor de gunning van het nieuwe steunpunt Toetsontwikkeling en Peilingen voor de periode 2018-2023.

1.5.4. Auteurs van dit rapport

Dit rapport werd geschreven door de afdeling Strategische Beleidsondersteuning van het departement, met belangrijke inhoudelijke input van de afdelingen Secundair Onderwijs en Leerlingenbegeleiding, Basisonderwijs en Deeltijds Kunstonderwijs, Hoger Onderwijs en Volwassenenonderwijs en Horizontaal Beleid. Dit rapport kwam tot stand in heel nauwe samenwerking met de afdeling Kwalificaties en Curriculum van Ahovoks en met de onderwijsinspectie.

2. HOOFDSTUK 2: OPZET VAN HET PEILINGSONDERZOEK

2.1. Inleiding

De huidige beheersovereenkomst van het steunpunt Toetsontwikkeling en Peilingen die loopt voor de periode van 2013-2017 voorziet in volgende opdrachten voor het steunpunt:

- toetsontwikkeling
- afname van de peilingen
- dataverwerking en –analyse
- rapportering en feedback
- vorming en valorisatie
- wetenschappelijke dienstverlening
- ontwikkeling en opbouw van specifieke expertise

Elke peiling bestaat uit twee fasen: de toetsontwikkeling en de eigenlijke peiling. In dit hoofdstuk willen we een overzicht geven van de door het steunpunt gehanteerde methodologie. Er wordt ook dieper ingegaan op de structuur, werking en aansturing van het steunpunt.

15

Voor dit hoofdstuk hebben we ons hoofdzakelijk gebaseerd op informatie uit de meerjarenplanning, die de bijlage vormt bij de beheersovereenkomst. Daarbij werden sommige delen overgenomen.

2.2. Het proces van een peiling

Tabel 1 geeft een overzicht van de verschillende fasen van een peilingsonderzoek. Daarbij wordt er een onderscheid gemaakt tussen de toetsontwikkeling en het peilingsonderzoek zelf.

Tabel 1: Algemene fasering van een peiling

Jaar	maand	Fase	Deelfase	
201x	september	1	Opmaak toetsmatrijs en itemontwikkeling	
	oktober	2		
	november	3		
	december	4		
201x +1	januari	5	Vooronderzoeken	
	februari	6		
	maart	7		
	april	8	Toetsontwikkeling	
	mei	9		
	juni	10		
	juli	11		
	augustus	12		<i>Analyses kalibratie</i>
	september	13		<i>Aankondiging peiling</i>
	oktober	14		
	november	15		<i>Ontwikkeling vragenlijsten</i>
	december	16		
201x+2	januari	17	Organisatie afname peiling	
	februari	18		
	maart	19		
	april	20	Peilingsonderzoek	
	mei	21		<i>Afname peiling</i>
	juni	22		
	juli	23		Analyses peiling
	augustus	24		
	september	25		
	oktober	26		
	november	27		Cesurbepaling en interpretatie
	december	28		
201x+3	januari	29	Rapportering en schoolfeedback	
	februari	30		
	maart	31		
	april	32		
	mei	33		

Bron: meerjarenplanning

2.2.1. Toetsontwikkeling

De toetsontwikkeling duurt 16 maanden, start steeds in september van het jaar 201x en loopt af aan het einde van het jaar 201x+1 (zie Tabel 1). Binnen de toetsontwikkeling zijn er verschillende deelfases.

Opmaak toetsmatrijs en itemontwikkeling

De opmaak van de toetsmatrijs en itemontwikkeling neemt 4 maanden in beslag en loopt van begin september tot eind december. Eerst wordt er een **toetsmatrijs** opgesteld. De toetsmatrijs vormt de blauwdruk van de te ontwikkelen toets. Hierin worden de te toetsen eindtermen geclassificeerd volgens een inhoudelijke dimensie en het vereiste verwerkingsniveau (indien gespecificeerd in de eindtermen).

De toetsmatrijs geeft aan voor welke eindtermen items en opdrachten ontwikkeld zullen worden.

Bij het opstellen van de toetsmatrijs worden beslissingen genomen over:

- a) de inhoudelijke specificaties van het te ontwikkelen peilingsinstrument (bijvoorbeeld over het al dan niet opnemen van bepaalde eindtermen);
- b) over welke (groepen van) eindtermen apart gerapporteerd zal worden of welke meetschalen er zullen ontwikkeld worden.

De beslissing over de gehanteerde toetsmatrijs en over de te ontwikkelen meetschalen wordt door de toetsontwikkelaars voorgelegd aan de interne stuurgroep (zie later).

Na de opmaak van de toetsmatrijs wordt gestart met de **itemontwikkeling**. Inspiratiebronnen hiervoor vormen de leerplannen, de handboeken en bestaande toetsen. De toetsvragen mogen daarbij geen enkel handboek of didactische methode bevoordelen. De ontwikkelde toetsvragen worden uitgebreid beoordeeld en besproken door de interne stuurgroep.

Naast schriftelijke proeven worden er in sommige peilingsprojecten ook praktische proeven ontwikkeld. Praktische proeven worden bij een deelsteekproef van leerlingen afgenomen.

17

Vooronderzoeken

Om de ontwikkeling van de toetsitems verder te optimaliseren, worden er vooronderzoeken georganiseerd bij leerlingen en onderwijsdeskundigen.

Het doel van het **vooronderzoek (of pilootafname) bij leerlingen** is om een inschatting te krijgen van de moeilijkheidsgraad van de items, de afnametijd van de toets en om mogelijke dubbelzinnigheden op te sporen in de vraagstelling. Er wordt bij een beperkte steekproef van 100 à 150 leerlingen een klassikale toetsafname georganiseerd. Wanneer ook een praktische proef wordt ontwikkeld, wordt deze eveneens in een vooronderzoek uitgetest.

In het **vooronderzoek bij onderwijsdeskundigen** worden de items voorgelegd aan een vijftiental experts (leerkrachten, lerarenopleiders, pedagogische begeleiders, leden van de inspectie en beleidsmedewerkers). Het doel is om te oordelen over de relevantie van de items en de geschiktheid voor de beoogde leerlingengroep.

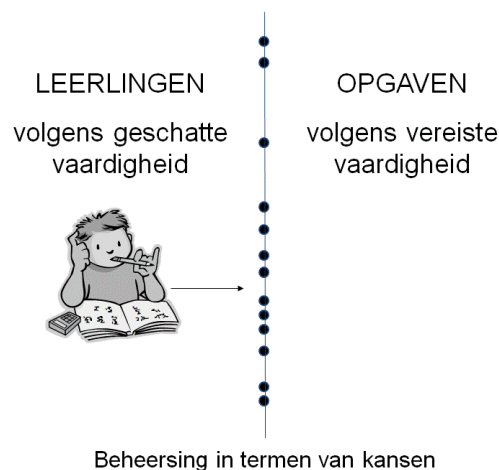
Op basis van de vooronderzoeken worden de ontwikkelde items herwerkt.

Kalibratieonderzoek

Het doel van het kalibratieonderzoek is om de ontwikkelde items per toets op een meetschaal te plaatsen. De meetschaal vormt de basis voor de samenstelling van de peilingstoets en haar parallelversie. Daarbij wordt er voor gezorgd dat de paralleltoets vergelijkbaar is wat moeilijkheidsgraad en de te meten eindtermen betreft. De paralleltoets meet dus hetzelfde als de peilingstoets, maar bevatten andere items.

In het kalibratieonderzoek worden de ontwikkelde toetsen afgenomen bij een representatieve steekproef van leerlingen aan het einde van het schooljaar. Er wordt bij de steekproeftrekking rekening gehouden met het onderwijsnet, de provincie, de schoolgrootte, het schooltype en de verstedelijkingsgraad. Voor het secundair onderwijs wordt er ook rekening gehouden met de basisopties, beroepenvelden, onderwijsvormen of studierichtingen. De antwoorden van de leerlingen op de schriftelijke toetsen worden via scanbare antwoordformulieren verwerkt. De korte-antwoordvragen en open antwoordvragen worden aan de hand van scoringswijzers manueel ingevoerd.

De leerlingresultaten worden vervolgens geanalyseerd aan de hand van de **itemresponsstheorie (IRT)** – modellen. Deze techniek laat toe om leerlingen en items of opgaven op een zelfde meetschaal te plaatsen. Figuur 1 geeft het principe weer van een meetschaal.



Figuur 1. Het principe van een meetschaal in een IRT-model

Bron: Ameer, E., Van Nijlen, D., De Meyst, M., Willem, L., Dekelver, S., Spillier, A., Crynen, M., & Janssen, R. (2015). De constructie van een peilingsinstrument wereldoriëntatie (domeinen natuur en techniek) in het basisonderwijs – Eindrapport. Leuven: KU Leuven, Steunpunt voor Toetsontwikkeling en Peilingen.

De bolletjes op de lijn in figuur 1 stellen de verschillende items voor. Een meetschaal is te vergelijken met een ladder. De sporten van de ladder verwijzen naar de toetsitems. Hoe hoger de items op de ladder staan, hoe moeilijker ze zijn. De vaardigheid van de leerling wordt aangegeven via het pijltje. De items onder het pijltje heeft de leerling onder de knie terwijl de items boven het pijltje te moeilijk zijn voor deze leerling.

Deze techniek biedt verschillende voordelen. Zo is het niet noodzakelijk dat alle leerlingen dezelfde items beantwoorden om op de meetschaal gesitueerd te kunnen worden. Dit laat toe om een grotere set van items op te nemen binnen hetzelfde tijdsbestek. Daarnaast kunnen de prestaties van de leerlingen in opeenvolgende peilingen vergeleken worden zonder dat ze steeds opnieuw moeten worden afgenomen. Een laatste voordeel is dat tijdens de cesuurbepaling (zie later) de onderwijsexperten gebruik kunnen maken van de meetschaal om na te gaan welke items leerlingen moeten beheersen om de eindtermen of ontwikkelingsdoelen te behalen.

Ten slotte krijgen de deelnemende scholen aan het kalibratieonderzoek feedback over de gemiddelde prestatie van hun school afgezet tegenover de gemiddelde prestatie van de volledige groep van leerlingen.

2.2.2. Peilingsonderzoek

De volledige fase van een peilingsonderzoek omvat 17 maanden. Deze start in januari van het jaar 201x+2 en loopt af aan het einde van het volgende jaar 201x+3. Ook binnen het peilingsonderzoek kunnen er verschillende deelfases onderscheiden worden.

Aankondiging en selectie van de scholen

Aan het begin van het schooljaar waarin de peiling plaatsvindt (september 201x+1) wordt de peiling aangekondigd via school- en lerarendirect. Daarna worden de scholen die deel uitmaken van de steekproef aangeschreven met de vraag tot deelname.

Daarbij wordt ervoor gezorgd dat de steekproef representatief is naar het onderwijsnet, een indicator voor de geografische locatie (verstedelingsgraad of provincie) en een indicator voor schooltype (bv. op basis van de schoolgrootte of op basis van de studierichtingen binnen een secundaire school). Om voldoende onderbouwde uitspraken te kunnen doen over het behalen van de eindtermen wordt er naar gestreefd elke meetschaal aan te bieden aan minimaal 1.500 leerlingen. De praktische proeven worden afgenomen bij een deelsteekproef van zes leerlingen per

vestigingsplaats van de deelnemende scholen. De selectie van deze deelsteekproef gebeurt op toeval door het onderzoeksteam.

Ontwikkeling van achtergrondvragenlijsten

Tijdens de fase van de toetsontwikkeling wordt reeds met de opmaak van de achtergrondvragenlijsten gestart maar de verdere afwerking loopt door tijdens de fase van het peilingsonderzoek zelf.

Bij elke peiling worden er achtergrondvragenlijsten opgesteld voor leerlingen, ouders en leerkrachten. Deze bijkomende informatie laat toe om de leerlingenprestaties in verband te brengen met achtergrondvariabelen op leerling-, gezins- en schoolniveau.

Het onderzoeksteam bereidt de inhoud van deze achtergrondvragenlijsten voor in overleg met een focusgroep van vakinhoudelijke experts.

Er worden steeds een aantal vaste variabelen opgenomen in de achtergrondvragenlijsten om sommige kenmerken van leerlingen (geslacht, leeftijd, thuistaal, informatie over de socio-economische status, het cultureel kapitaal van het gezin, ...) en scholen (provincie, % GOK leerlingen op school, ...) in kaart te brengen. Daarnaast worden vragen opgenomen die meer eigen aan een bepaalde peiling zijn.

20

Na proefafnames worden de achtergrondvragenlijsten bijgewerkt en gefinaliseerd in overleg met de interne stuurgroep.

Aan ouders en leerkrachten wordt de mogelijkheid geboden om de achtergrondvragenlijst op papier, dan wel digitaal in te vullen.

Afname

Er worden externe toetsassistenten ingezet die de afname binnen een school coördineren en toezien op het correcte verloop. Wanneer bij de peiling ook een praktische proef is voorzien, staan de toetsassistenten in voor deze afname. Om de toetsassistenten hierop goed voor te bereiden, wordt een opleidingssessie voorzien.

Analyses

Net zoals bij het kalibratieonderzoek worden de leerlingenresultaten op de schriftelijke toetsen verwerkt via scanbare antwoordformulieren en worden de open vragen gescoord aan de hand van scoringswijzers en geanalyseerd aan de hand van IRT-

modellen. De gegevens uit het kalibratieonderzoek worden meegenomen in deze analyses zodat de items uit de paralleltoets op dezelfde meetschaal gesitueerd kunnen worden als de items uit het peilingsinstrument.

Voor de praktische proeven worden er geen meetschalen opgemaakt. De resultaten ervan worden op een descriptieve manier geanalyseerd (bv. % leerlingen dat een bepaalde deelopdracht juist uitvoert). Er wordt daarbij aandacht gegeven aan het identificeren van bepaalde patronen. In sommige peilingsprojecten worden de resultaten van de praktische proeven beoordeeld aan de hand van doelstellingen die gelinkt zijn aan de eindtermen. Op die manier kan gedeeltelijk een uitspraak worden gedaan over het al dan niet behalen van de betreffende eindtermen.

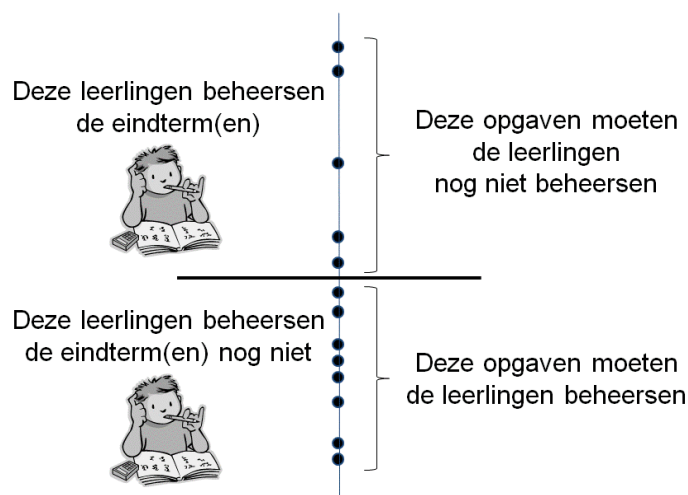
Ten slotte voert het steunpunt multilevelanalyses uit om:

- a) na te gaan in welke mate de verschillen in prestaties tussen leerlingen, klassen en scholen gesitueerd zijn;
- b) te onderzoeken welke factoren op leerling-, gezins-, klas- en schoolniveau samenhangen met de prestaties.

Cesurbepaling en interpretatie

Het doel van de cesurbepaling is om het minimumniveau te bepalen dat leerlingen moeten behalen op de meetschaal om van een voldoende beheersing van de eindtermen te spreken. De cesurbepaling vindt meestal in november plaats.

Het bepalen van een minimumniveau wordt schematisch voorgesteld in figuur 2. Gelijkaardig aan figuur 1 stellen de bolletjes de items voor die zoals een ladder in opgaande richting geordend zijn volgens moeilijkheidsgraad. Het leggen van de cesuur komt neer op het beantwoorden van de vraag welke moeilijkheidsniveau leerlingen moeten kunnen halen opdat ze het vereiste minimumniveau van de eindtermen of ontwikkelingsdoelen bereiken. Dat minimumniveau deelt de items van een meetschaal op in twee groepen, met name de items die beheerst moeten worden om de eindtermen of ontwikkelingsdoelen te bereiken en de items die niet hoeven beheerst te worden door leerlingen die de eindtermen of ontwikkelingsdoelen net behalen.



Figuur 2. Voorstelling van een meetschaal en bepaling van het minimum vaardigheidsniveau

Bron: Ameal, E., Van Nijlen, D., Willem, L., Crynen, M., & Janssen, R. (2016). *Peiling wereldoriëntatie natuur en techniek in het basisonderwijs – Eindrapport. Leuven: KU Leuven, Steunpunt Toetsontwikkeling en Peilingen.*

De gehanteerde methode voor de **cesuurbepaling** bestaat uit een licht aangepaste versie van de Bookmarkmethode (letterlijk: bladwijzermethode) van Mitzel, Lewis, Patz en Green (2001)⁴.

Aan een groep van 25 onderwijsdeskundigen (waaronder beleidsmedewerkers van het ministerie van Onderwijs en Vorming) worden de toetsitems voorgelegd, geordend in stijgende graad van moeilijkheid. De taak van de beoordelaars bestaat erin om een figuurlijke bladwijzer te plaatsen tussen de twee items die de overgang vormen tussen items die wel beheerst moeten worden en items die niet beheerst moeten worden om de getoetste eindterm(en) te halen. Het verschil met de Bookmarkmethode bestaat eruit dat de beoordelaars naast de concrete items gerangschikt volgens moeilijkheidsgraad, ook een visuele weergave van de meetschaal krijgen met daarop de posities van de items, opnieuw gerangschikt volgens hun moeilijkheidsgraad. De meetschaal dient om de beoordelaars een visueel overzicht te geven van de verschillen in moeilijkheidsgraad.

De beoordeling verloopt in drie ronden. In een eerste ronde brengt elke beoordelaar individueel een oordeel op de meetschaal aan. In een tweede ronde is er mogelijkheid tot overleg in kleine groepen over het uitgebrachte oordeel en de argumentatie erachter.

⁴ Mitzel, H. C., Lewis D. M., Patz R. J. & Green D. R. (2001). The bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods and perspectives* (pp. 249-281). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

De beoordelaars geven na deze tweede ronde een nieuw oordeel. Ten slotte volgt er een derde ronde waar met de hele groep aan deelneemt. In deze ronde wordt gestart met het geven van feedback over de mediaan van de oordelen die tijdens de eerste twee rondes werden gegeven. Daarna wordt informatie gegeven over het percentage leerlingen dat volgens de cesuur aangegeven in de tweede ronde, de eindtermen of ontwikkelingsdoelen behaalt. De derde ronde wordt afgesloten met het uitbrengen van een derde en laatste oordeel door elk van de beoordelaars afzonderlijk. De definitieve cesuur wordt vastgelegd op basis van de mediaan van deze finale derde ronde.

Tot slot wordt er nog een vierde ronde georganiseerd om te komen tot een **inhoudelijke interpretatie** van de meetschalen. De onderwijsdeskundigen kunnen daarbij aangeven of er aspecten zijn waarop items boven en onder de cesuur zich van elkaar onderscheiden. Deze resultaten worden eveneens besproken in en teruggekoppeld naar de interne stuurgroep.

Rapportering van de resultaten, valorisatie en schoolfeedback

Het proces van de toetsontwikkeling en de resultaten van een peiling worden gerapporteerd in twee afzonderlijke **wetenschappelijke rapporten**.

Daarnaast wordt ook een **brochure** opgemaakt voor een breed publiek waarin de belangrijkste resultaten worden voorgesteld. De resultaten worden aan een breed publiek bekend gemaakt op het colloquium. Dit vindt meestal plaats aan het eind van het schooljaar volgend op de afname van de peiling.

De scholen ontvangen hun **feedbackrapport** ontvangen op het colloquium. In dit feedbackrapport krijgt de school informatie over de gemiddelde score op de eindtermen voor de eigen leerlingen en over het aantal leerlingen van de school dat de eindtermen behaalt. Daarbij krijgt men ook informatie over het Vlaams gemiddelde, eventueel uitgesplitst per basisoptie of studierichting. Scholen kunnen zich op die manier ook positioneren ten opzichte van andere scholen met een gelijkaardig leerlingenpubliek. Scholen krijgen echter het resultaat van andere scholen niet te zien.

Voor praktische proeven wordt geen schoolfeedback voorzien aangezien het leerlingenaantal te laag is om betrouwbare uitspraken te doen op schoolniveau.

2.3. Structuur, werking en opvolging van het steunpunt

2.3.1. Structuur en werking

Het steunpunt als geheel wordt ondersteund door een groep van **promotoren** met expertise op het gebied van peilingsonderzoek, IRT-modellen, multiniveau-analyses en schooleffectiviteitsonderzoek. Ze zijn verantwoordelijk voor de continuïteit over peilingprojecten heen en zorgen voor elk project voor methodologische ondersteuning.

Daarnaast wordt er voor elke project samengewerkt met **inhoudelijke co-promotoren** die instaan voor de ondersteuning op de inhoudelijke aspecten van elke peiling en mee de afstemming van de peilingsinstrumenten op de gepeilde eindtermen en het vakgebied garanderen.

Het onderzoeksteam van het steunpunt bestaat zelf uit drie onderzoekscellen: de cel organisatie, de cel toetsontwikkeling en de cel data-analyse. Voor elk van deze cellen zijn er vaste, gespecialiseerde medewerkers.

De **cel organisatie** staat in voor de praktische organisatie van de peilingen. Hun taken gaan van het contacteren van scholen, het opmaken en verzenden van toetsmateriaal, het opzetten en organiseren van pilootstudies tot het verwerken van het ingevulde materiaal. Zij kunnen daarbij terugvallen op uitgewerkte draaiboeken. Binnen de **cel toetsontwikkeling** worden per deelproject de toetsen ontwikkeld. Ook medewerking aan de cesuurbepaling en het interpreteren van de resultaten van de peilingen behoren tot het takenpakket. Voor de toetsontwikkeling wordt er per peilingsproject, naast een vaste medewerker, gewerkt met een projectspecifiek (tijdelijk) team van toetsontwikkelaars. Dit om bij de ontwikkeling van de toetsinstrumenten de inhoudelijke deskundigheid voor elk project te garanderen. Het zwaartepunt van de **cel data-analyse** ligt bij de analyses van de gegevens uit de kalibratiestudies en peilingen. Sinds 2014 is er een vierde onderzoekscel die instaat voor de afname van de paralleltoetsen, de bijhorende schoolfeedbackrapporten en het onderhoud van het nieuwe platform paralleltoetsen.be.

De communicatie en afstemming tussen de verschillende cellen gebeurt aan de hand van een wekelijks overleg tussen de projectcoördinator en de coördinatoren van de cellen organisatie, toetsontwikkeling en data-analyse en een maandelijks teamoverleg met alle medewerkers. De cel die werkt rond de paralleltoetsen werkt nauw samen met de projectcoördinator en de informaticus (en occasioneel ook met de andere teamleden van het steunpunt). De **projectcoördinator** bespreekt wekelijks de stand van zaken met de promotor-coördinator en staat in voor de afstemming met de opdrachtgever.

2.3.2. Opvolging

Vandaag worden de werkzaamheden van het steunpunt opgevolgd door drie organen:

- 1) Een stuurgroep
- 2) Een kerngroep
- 3) Een interne stuurgroep

Stuurgroep

De stuurgroep bestaat uit vertegenwoordigers van het departement Onderwijs en Vorming, onderwijsinspectie, Ahovoks en het steunpunt. Zij staan in voor de algehele opvolging en aansturing van het peilingsonderzoek. De frequentie van samenkomst is ongeveer twee keer per jaar.

Kerngroep

De kerngroep staat in voor de dagelijkse werking en opvolging van het steunpunt. Naast vertegenwoordigers van het steunpunt zelf bestaat die uit vertegenwoordigers van de afdeling Strategische Beleidsondersteuning van het departement Onderwijs en Vorming en Ahovoks. Ze komt tweemaandelijks samen en nemen projectspecifieke thema's op. Daarnaast worden ook de stuurgroepvergaderingen voorbereid door de kerngroep.

Interne stuurgroep

De interne stuurgroep wordt voorgezeten door de projectcoördinator en bestaat uit de medewerkers van de cel toetsontwikkeling, vertegenwoordigers van de andere cellen, inhoudelijke partners en de inhoudelijke experts van Ahovoks. De inhoudelijke experts van Ahovoks hebben een observerende rol. Gedurende de fase van de toetsontwikkeling komt de interne stuurgroep vijf keer samen. Het doel van de interne stuurgroep in de fase van de toetsontwikkeling is om de toetsontwikkelaars feedback te geven bij het opstellen van toetsmatrijzen en het ontwikkelen van de items. Tijdens het peilingsonderzoek speelt de interne stuurgroep een rol in het uitwerken van de achtergrondvragenlijsten en bij de interpretatie van de meetschalen.

3. HOOFDSTUK 3: EVALUATIE VAN DE LOPENDE BEHEERSOVEREENKOMST (2013-2017)

3.1. Situering

In de laatste maanden van 2016⁵ evalueerden drie verschillende partijen de lopende beheersovereenkomst van het steunpunt Toetsontwikkeling en Peilingen. Het steunpunt maakte een zelfevaluatie, de opdrachtgever vertegenwoordigd door de leden van de stuurgroep (Ahovoks, departement Onderwijs en Vorming en inspectie) maakte een evaluatie en er was een peerevaluatie door 2 experts, Daniël Muys en Norman Verhelst (ieder afzonderlijk).

In dit hoofdstuk worden de resultaten van deze drie evaluaties gebundeld weergegeven. Daarbij worden verschillende aspecten van de werking van het steunpunt belicht zoals de beheersovereenkomst, de werking, de aansturing en de samenwerking en de verschillende opdrachten.

3.2 Beheersovereenkomst, werking, sturing & samenwerking

3.2.1 De beheersovereenkomst

In de beheersovereenkomst worden verschillende aspecten van het steunpunt geregeld zoals de aansturing, de inhoud van de opdracht, het budget waarover het steunpunt kan beschikken, de intellectuele eigendom, de openbaarheid en vertrouwelijkheid van het ontwikkelde instrumentarium en de wijze waarop omgegaan wordt met de resultaten.

In deze evaluatie wordt nagegaan in welke mate de beheersovereenkomst als instrument in voldoende mate beantwoordt aan deze doelstellingen.

Sturing en opvolging

De wijze van aansturing is duidelijk omschreven in de beheersovereenkomst. Alle betrokkenen krijgen voldoende ruimte om te overleggen en hun inbreng te doen. Naar de toekomst toe wordt wel gevraagd om alle elementen die deel uitmaken van de aansturing op te nemen in de beheersovereenkomst. Zo is bijvoorbeeld de werking en positie van de kerngroep nu niet beschreven in de beheersovereenkomst.

Inhoud en omvang van de opdracht

⁵ Deze evaluaties vonden reeds in de tweede helft van 2016 plaats om de resultaten ervan te kunnen meenemen in het nadenken over en vormgeven van een nieuw Steunpunt.

Een beheersovereenkomst over meerdere jaren heen, zorgt voor de nodige continuïteit op het vlak van financiering en personeelsbeleid en laat de nodige expertise-opbouw toe.

Wel zouden de inhoud en de omvang van de opdracht, de verschillende verantwoordelijkheden en de beslissingsbevoegdheid van de verschillende betrokken partijen nauwkeuriger omschreven kunnen worden.

In de initiële beheersovereenkomst werd een peilingenkalender opgenomen die later verschillende keren gewijzigd werd omwille van aanpassingen aan de eindtermen. Een meer systematische peilingenkalender zonder wijzigingen zou beter zijn. Het uitvoeren van twee peilingen per jaar is haalbaar gebleken, maar het is wel belangrijk dat deze twee peilingen verdeeld worden over basisonderwijs en secundair onderwijs. Vooral het uitvoeren van twee peilingen in het secundair onderwijs zou moeilijk zijn omwille van het te beperkte aantal secundaire scholen.

Budget

In de beheersovereenkomst wordt een jaarlijkse subsidie bepaald, die in het jaarplan nadien nog licht kan bijgestuurd worden. Er is ook een beperkte reserveopbouw mogelijk. Er werd vastgesteld dat de beschikbare financiële middelen, niet maximaal werden aangewend. Dat betekent dat er binnen het huidige budget ruimte is om de bestaande activiteiten te versterken of om innovatieve componenten toe te voegen.

27

Intellectuele eigendom, vertrouwelijkheid en openbaarheid

De regeling die geldt voor het vrijgeven van items en het bekendmaken van de peilingsresultaten wordt als positief geëvalueerd.

3.2.2 De interne werking, sturing & samenwerking

In dit gedeelte worden de aansturing van het steunpunt en de samenwerking tussen de verschillende betrokken partijen geëvalueerd.

De interne organisatiestructuur van het steunpunt

De interne structuur van het steunpunt, bestaande uit de verschillende cellen (zie 2.3.1.), wordt door alle partijen als effectief en efficiënt ervaren. Alle cellen zijn noodzakelijk om de volledige opdracht tot een goed einde te kunnen brengen. Ook de onderlinge samenwerking wordt positief geëvalueerd. De rol van de projectcoördinator is cruciaal. Zo geeft de opdrachtgever aan dat het aanstellen van een coördinator positief is en dat deze rol naar de toekomst toe zelfs verder versterkt kan worden. Door het steunpunt wordt aangegeven dat de coördinatierol veeleisend is. Men pleit ervoor om in de toekomst dezelfde structuur te behouden.

Er wordt door het steunpunt samengewerkt met heel wat externe partners zoals de inhoudelijke promotoren, de toetsontwikkelaars, de onderwijsinspectie en de pedagogische begeleidingsdiensten. De samenwerking met deze partijen verloopt doorgaans vlot en is effectief. Als er knelpunten waren, dan waren deze eerder van administratieve aard. Het tot stand brengen van deze samenwerkingsverbanden is arbeidsintensief, maar de moeite waard en nodig voor een degelijke afstemming met de onderwijspraktijk.

De interne stuurgroep

De interne stuurgroep komt enkel in de zelfevaluatie aan bod en wordt door het steunpunt zelf ervaren als een belangrijk forum waarin opdrachtnemer en opdrachtgever kunnen afstemmen in de fase van de toetsontwikkeling. Voldoende betrokkenheid van alle betrokkenen is een vereiste. Naar de toekomst toe is het nodig duidelijker te bepalen welke beslissingen kunnen genomen worden in de interne stuurgroep en hoe deze stuurgroep zich verhoudt tot de andere overlegorganen.

De kerngroep

De kerngroep wordt door alle evaluatoren die vertrouwd zijn met de werking ervan, positief beoordeeld. In de kerngroep heerst een constructieve en eerder informele sfeer. Hierdoor wordt vragen en bekommernissen vlot op tafel gelegd en kunnen projectspecifieke vragen op korte termijn behandeld worden. In de kerngroep worden ook geregeld projectoverstijgende topics aangebracht, vaak in voorbereiding op de stuurgroep. Vanuit het steunpunt wordt meegegeven dat men in de kerngroep een hoge betrokkenheid van de opdrachtgever ervaart.

Aandachtspunten, ook naar de toekomst toe, zijn de informatiedoorstroom tussen de kerngroep en de interne stuurgroep en een voldoende betrokkenheid van inhoudelijke experts bij de kerngroep.

De stuurgroep

Alle evaluatoren die een voldoende kijk hebben op de stuurgroep concluderen dat de samenstelling goed is, maar dat er vaak te weinig opkomt is waardoor de toegevoegde waarde van de stuurgroep ten opzichte van de kerngroep onvoldoende groot is. Er gaat daardoor te weinig sturing op strategisch niveau uit vanuit de stuurgroep, wat dan wordt overgenomen door de kerngroep.

Naar de toekomst moeten de samenstelling en de taakstelling van de stuurgroep herbekeken worden zodat zij meer strategische inbreng kan hebben op de verschillende peilingsprojecten en de lange termijnplanning. Er wordt ook gesuggereerd linken te leggen met de stuurgroep van het steunpunt SONO en de stuurgroep van het internationaal vergelijkend onderzoek. Wat betreft de

samenstelling dient de opname van de koepels en het GO! overwogen te worden.

3.3 Opdrachten uit de beheersovereenkomst

3.3.1 Toetsontwikkeling

Opdrachtomschrijving

Onder de opdracht toetsontwikkeling valt de ontwikkeling van zowel schriftelijke toetsen, praktische proeven en de achtergrondvragenlijsten als de ontwikkeling van de paralleltoetsen. Ook de kalibratie en de cesuurbepaling behoren tot de toetsontwikkeling.

Voor een volledig overzicht van het ontwikkelingsproces van een peiling, zie hoofdstuk 2.

Sterke punten

De verschillende evaluaties geven volgende positieve punten aan:

- een doorgedreven inhoudelijke analyse van de eindtermen en ontwikkelingsdoelen als startpunt van de toetsontwikkeling;
- de itemontwikkeling die door meerdere personen wordt opgenomen, waardoor een brede blik en onderlinge feedback voor de nodige diversiteit aan items zorgen;
- recente ontwikkelingen in de onderwijspraktijk worden meegenomen in het ontwikkelingsproces dankzij de samenwerking met externe partners;
- de duidelijke procedures die worden gehanteerd en de strikte timing doorheen het hele proces;
- de uitgebreide pilootafnames en itembeoordelingen door diverse experts, wat de kwaliteit ten goede komt;
- de introductie van een toetsenbank.

Mogelijke verbeterpunten

Daarnaast formuleerden de verschillende evaluatoren enkele mogelijke verbeter- of aandachtspunten:

- het beter bewaken van het evenwicht tussen de verwachtingen door de opdrachtgever en de mogelijkheden van grootschalige toetsafnames, zeker bij eindtermen die moeilijker peilbaar zijn;
- het creëren van meer duidelijkheid over welke beslissingen door welke partij kunnen worden genomen en wat het onderlinge gewicht is van al die inbreng;
- het duidelijker formuleren van de opdrachten inzake toetsontwikkeling en de

- ontwikkeling van praktische proeven;
- bij het kiezen voor een praktische proef beter nadenken over de meerwaarde en de haalbaarheid ervan;
 - de rol van de toetsassistenten bij de praktische proeven reeds bij de ontwikkeling nauwkeurig omschrijven;
 - beter beargumenteren waarom de feedback op toetsitems al dan niet meegenomen wordt;
 - meer aandacht voor de achtergrondvragenlijsten in de opdrachtomschrijving van het steunpunt;
 - het duidelijker formuleren van de onderzoeksvragen door de opdrachtgever om zo de ontwikkeling van achtergrondvragenlijsten beter te faciliteren;
 - vanuit de opdrachtgever een sterkere rol van de inhoudelijke co-promotoren bij het opmaken van beleidsrelevante achtergrondvragenlijsten.

De kalibratie is noodzakelijk maar vergt wel een grote inspanning van scholen. Samen met de beperktere schoolfeedback t.o.v. een deelname aan de eigenlijke peiling, zorgt dit voor een mogelijke lagere respons.

Ook wordt in de evaluatie gevraagd om na te denken over mogelijke alternatieven voor de cesuurbepaling waarbij o.a. wordt gewezen op de mogelijke introductie van beheersingsniveaus en het verder professionaliseren van de deelnemers. Vanuit het steunpunt wordt er rond de cesuurbepaling op gewezen dat zij zich bewust zijn van de beperkingen, maar inspanningen leveren om deze zoveel mogelijk te ondervangen, alsook dat de gehanteerde methode ook internationaal veel steun krijgt.

3.3.2 Afname van peilingen

Opdrachtomschrijving

Onder de afname worden de steekproeftrekking, het afnamedesign, de afnameduur, de antwoordmodaliteiten, de toetsassistenten, de praktische organisatie en de responsgraden gevat. Voor een overzicht van het volledige afnameproces: zie hoofdstuk 2.

Sterke punten

Volgende sterke punten worden aangehaald in de verschillende evaluaties:

- de afname wordt goed voorbereid door het gedegen overwegen en het nauwkeurig vaststellen van steekproef, design en duurtijd;
- er is een draaiboek dat regelmatig wordt bijgewerkt waar nodig;
- de eigenlijke afname van de peilingen verloopt steeds vlot en efficiënt;

- de behaalde respons is hoog⁶ en daardoor zijn er voldoende analysemogelijkheden (zie infra);
- er wordt gewerkt met toetsassistenten en dat verhoogt de betrouwbaarheid van de resultaten.

Mogelijke verbeterpunten

Volgende mogelijke verbeterpunten komen uit de verschillende evaluaties:

- er kan meer gebruik gemaakt worden van digitaal toetsen waarbij steeds moet afgewogen worden wat de mogelijke meerwaarde ervan kan zijn. Aandachtspunt hierbij is wel dat dit grote repercussies zal hebben op de toetsontwikkeling;
- er dient een strategie te worden ontwikkeld over de mogelijkheid om redelijke aanpassingen toe te laten;
- er dient voldoende aandacht te zijn voor een mogelijke toekomstige daling van de respons en dan vooral tijdens de kalibratie, waarbij men moet waken over een voldoende draagvlak bij scholen;
- het invoeren van een non-response-biasanalyse wordt aangeraden.

3.3.3 Dataverwerking en –analyse

Opdrachtomschrijving

Onder dataverwerking en –analyse vallen het werken met beoordelaars, het invoeren van antwoorden, de statistische verwerking en de inhoudelijk analyse van de resultaten.

Sterke punten

De statistische verwerking van de resultaten wordt door alle evaluatoren gezien als van een hoog niveau.

Mogelijke verbeterpunten

De opdrachtgever geeft aan dat de verwerking van de resultaten een te lange doorlooptijd kent. Daartegenover staat dat het steunpunt aangeeft dat er telkens extra inspanningen nodig zijn om de timing te kunnen halen. Als mogelijke oplossing om hieraan tegemoet te komen wordt verwezen naar de voordelen die het digitaal toetsen biedt bij de verwerking van de resultaten.

Daarnaast wordt aangegeven dat de inhoudelijke analyse van de resultaten versterkt

⁶ Eén van de wetenschappelijke peers beoordeelt de uitval van scholen als hoog.

kan worden, iets wat al wordt opgenomen door de aanstelling van een coördinator. Er dient daarbij meer aandacht te worden besteed aan de beleidsrelevantie van analyses. Zo kan men werken met hypothesen en kan de expertise van inhoudelijke copromotor meer worden ingezet. Ook stuurgroep en kerngroep moeten hun rol spelen bij het bepalen van beleidsrelevante analyses. Vanuit de inspectie wordt gevraagd naar het meer doorgedreven inzetten van vrijgegeven items bij de inhoudelijke analyse.

Vanuit een bredere perspectief stelt zich de vraag of de verzamelde data vandaag niet onderbenut worden en of, naast het versterken van de inhoudelijke analyses binnen het steunpunt, het beschikbaar stellen van de datasets aan andere onderzoekers hieraan een oplossing kan bieden.

3.3.4 Rapportering en feedback

Opdrachtomschrijving

Onder rapportering en feedback worden de opmaak van de onderzoeksrapporten (zowel toetsontwikkelingsrapporten als eindrapporten) voor de opdrachtgever als de schoolfeedbackrapporten bedoeld.

Sterke punten

De rapporten worden steeds zorgvuldig en met veel oog voor detail opgesteld. Het daarbij gevolgde stramien is duidelijk en wordt consistent gehanteerd. De rapporten worden steeds tijdig opgeleverd.

Mogelijke verbeterpunten

In lijn met de opdracht op het vlak van data-analyse is ook qua rapportering de inhoudelijke analyse een aandachtspunt (zie ook 3.3).

De schoolfeedbackrapporten kunnen meer op maat van de scholen worden opgemaakt want ze zijn momenteel te technisch opgevat. Scholen kunnen ook meer worden ondersteund bij de interpretatie en het gebruik van de schoolfeedbackrapporten.

Vanuit het steunpunt stelt men voor de toetsontwikkelings- en eindrapporten actief te verspreiden via de website omdat deze rapporten heel wat nuttige info bevatten voor het begrijpen van het onderzoeksopzet en het inschatten van de kwaliteit ervan.

3.3.5 Vorming en valorisatie

Opdrachtomschrijving

Onder deze opdracht wordt de vorming aan scholen, koepels, inspectie, ... en de valorisatie en communicatie via persberichten, persartikels, brochures⁷, colloquia, werkseminaries, paralleltoetsen en andere mogelijke initiatieven gevat.

Sterke punten

De evaluatoren vinden het samen op poten zetten van werkseminaries en colloquia een positieve evolutie. Deze worden consequent bij elke peiling georganiseerd, net zoals er ook telkens een brochure wordt gepubliceerd. De brochure is bruikbaar om de resultaten voor te stellen aan een breed publiek. Het informatieve karakter van deze initiatieven moet zeker behouden blijven.

Het platform voor de paralleltoetsen werkt goed. Er werd ook een brochure uitgewerkt rond de paralleltoetsen.

Mogelijke verbeterpunten

Een betere afstemming van de valorisatie en communicatie op het beoogde publiek qua taal, stijl en inhoud is noodzakelijk.

Er kan ook een bredere communicatie opgezet worden rond nieuwe peilingen of over peilingresultaten, bv. via Klasse of nieuwsbrieven. Ook kan de opbouw van de curriculumwebsite nog verbeterd worden. Zo is het bv. moeilijk om de meest recente brochures te vinden.

De valorisatie van de resultaten van het peilingsonderzoek is te beperkt. Er moet daarom nagedacht worden over een mogelijke versterking van het valorisatietraject en de consequenties daarvan voor alle fases van een peilingsonderzoek. Daarbij kunnen eventueel elementen die vroeger bestonden (vb. consultatiefase, de open conferenties voor directies en leraren, de bredere verspreiding van brochures) terug opgenomen worden. De valorisatie kan ook uitgebreid worden door relevante linken te leggen met andere onderzoeksresultaten.

Voor de paralleltoetsen dient er nagedacht te worden over pistes om de werklast voor scholen te verlichten zodat meer scholen deelnemen.

Het steunpunt krijgt weinig vragen voor vormingen vanuit het onderwijsveld. Wel zijn er initiatieven door inhoudelijke co-promotoren, maar dit is wisselend van project tot project.

Het steunpunt en de opdrachtgever zitten niet op één lijn over de verwachtingen m.b.t. de resultaten. Het steunpunt verwijst naar de beheersovereenkomst waarin wordt

⁷ Vanuit het Steunpunt werden de brochures ook opgenomen onder het punt Rapportering.

bepaald welke de op te leveren resultaten zijn en waaraan ook voldaan wordt. De beleidsmatige valorisatie en de mogelijke implicaties voor de onderwijspraktijk ziet men als een taak voor de opdrachtgever. Het steunpunt is zich wel bewust van de vraag naar een meer informatief verhaal of het meer aanreiken van verklaringen voor bepaalde resultaten en ziet wel het gebrek daaraan. Het steunpunt wijst wel op het feit dat er bij het opstellen van de achtergrondvragenlijsten weinig vraagstelling wordt aangereikt vanuit de opdrachtgever (zie ook 3.3.3 en 3.3.4 rond inhoudelijke analyses).

3.3.6 Wetenschappelijke dienstverlening

Opdrachtomschrijving

Wetenschappelijke dienstverlening bestaat uit het ingaan op ad hoc onderzoeksvragen rond toetsontwikkeling en prestatiemetingen.

Sterke punten

De onderzoekers zijn steeds bereid om op diverse fora de resultaten van de peilingen toe te lichten.

Mogelijke verbeterpunten

In de beheersovereenkomst wordt best concreter omschreven wat wordt verstaan onder wetenschappelijke dienstverlening. Uit de evaluaties blijkt immers dat er geen éénduidige interpretatie wordt gegeven aan de invulling van deze opdracht.

De onderzoekers kunnen meer eigen initiatieven nemen inzake wetenschappelijke dienstverlening want nu lijkt deze opdracht vooral vraaggestuurd te worden ingevuld. Zo kunnen de onderzoekers meer naar buiten treden met de resultaten met het oog op een bijdrage aan het maatschappelijk debat over bv. toetsing en evaluatie in het onderwijs.

3.3.7 Ontwikkeling van specifieke expertise

Opdrachtomschrijving

Het verder uitbouwen van Vlaamse expertise op het vlak van toetsontwikkeling en peilingsonderzoek vormt de kern van deze opdracht. Meer specifiek zijn dit verdere digitalisering, de verbreding van de methodologie van de praktische proeven en het uitvoeren van wetenschappelijk onderzoek.

Sterke punten

Het team van het steunpunt is stabiel, waardoor de specialistische expertise daar gegarandeerd is. Zowel binnen de cel toetsontwikkeling als binnen de cel data-analyse werd doorheen de jaren voldoende specifieke kennis en

ervaring opgebouwd, o.a. via deelname aan wetenschappelijke congressen en workshops. Deze expertise werd ook ingezet bij extra opdracht m.b.t. de validering van de IDP's en de OVSG-toetsen.

Er is ook veel expertise opgebouwd rond het kwaliteitsvol afnemen van peilingen (met inbegrip van praktische proeven).

Er wordt meer geïnvesteerd in doctoraatsonderzoek en er zijn ook interessante masterproeven gemaakt, hoewel het niet altijd eenvoudig is geïnteresseerde studenten aan te trekken.

Mogelijke verbeterpunten

Het moet mogelijk zijn om het onderzoek meer weerklank te laten vinden op het internationale wetenschappelijke forum.

De vakinhoudelijke expertise die door de toetsontwikkelaars werd opgebouwd, kan meer gevaloriseerd worden, bv. door deze ook in te zetten bij de ontwikkeling van achtergrondvragenlijsten en bij de inhoudelijke analyses van de resultaten.

Het peilingsonderzoek verloopt via een vast proces en dat heeft als nadeel dat het soms moeilijk is om in te gaan op nieuwe vragen die vereisen dat men deze processen gedeeltelijk loslaat. Dit hangt ook samen met innovatiekracht en de ruimte om te experimenteren.

4. HOOFDSTUK 4: EX ANTE EVALUATIE TOEKOMST VAN DE PEILINGEN

4.1. Inleiding: situering en methode

Om een zicht te krijgen op de meningen, standpunten en perspectieven van verschillende stakeholders werd geopteerd om te werken met enerzijds focusgroepen waarin vragen naar mogelijke doelstellingen, wenselijke inhoud en gebruik van peilingen en anderzijds een terugkomdag met alle deelnemers van de focusgroepen waarin aan de hand van diverse scenario's nagedacht werd over kenmerken van peilingen en randvoorwaarden voor het gebruik ervan.

Dit hoofdstuk bevat een samenvatting van de standpunten en aanbevelingen die daar aan bod kwamen.

In de periode oktober – november 2016 vonden er acht focusgroepen plaats. Deze focusgroepen bestonden uit de volgende doelgroepen:

- experts kwaliteitszorg en inspectie;
- pedagogische begeleiders;
- beleidsmedewerkers van het departement Onderwijs en Vorming en van AHOVOKS;
- hoger onderwijs;
- volwassenenonderwijs (directies en leerkrachten uit Centra voor volwassenenonderwijs en Centra voor basiseducatie);
- directeurs basis- en secundair onderwijs met inbegrip van het buitengewoon onderwijs;
- leerkrachten basis- en secundair onderwijs met inbegrip van het buitengewoon onderwijs;
- ouders en leerlingen.

36

In elk van deze focusgroepgesprekken werd er onder de deelnemers discussie gevoerd over:

- a) de mate waarin peilingen er vandaag in slagen om hun doelstellingen te bereiken;
- b) welke mogelijke alternatieve doelstellingen geformuleerd kunnen worden met betrekking tot de peilingen;
- c) de wenselijke inhoud en gebruik van peilingen.

De resultaten van de focusgroepgesprekken werden verwerkt tot vier mogelijke scenario's die op een terugkomdag (9 december 2016) ter discussie werden voorgelegd aan alle leden van de focusgroepen. Zo probeerden we een antwoord te krijgen op de volgende onderzoeksvragen:

- Welke doelstellingen kunnen de peilingen hebben?

- Door wie en op welke manier kunnen de peilingen gebruikt worden?
- Tot welk soort van informatie kunnen de peilingen leiden?
- Aan welke randvoorwaarden dienen peilingen te voldoen voor een optimaal gebruik?

In [bijlage 2](#) is een gedetailleerde beschrijving van de gevolgde methode opgenomen.

In wat volgt rapporteren we over de belangrijkste resultaten van deze ex ante beleidsevaluatie. We starten met de verschillende doelstellingen van de peilingen. Hoe staat men tegenover de verschillende mogelijke doelstellingen van de peilingen en welke verwachtingen impliceert dit dan naar het gebruik van de peilingen en de paralleltoetsen (4.2)? Vervolgens gaan we dieper in op het instrument van de peiling. Hoe ziet men de peiling van de toekomst (4.3)? Hierbij gaan we dieper in op de gewenste afstemming met andere toetsen, doelgroepen, inhoud van de peiling, meetmomenten en de frequentie, afname en feedback na een peiling. Tot slot besteden we aandacht aan de randvoorwaarden die de betrokkenen noodzakelijk achten om de doelstellingen van de peilingen te kunnen realiseren (4.4).

4.2. Aan welke doelstellingen kunnen de peilingen bijdragen?

Zoals blijkt uit eerdere hoofdstukken hebben de peilingen en paralleltoetsen vandaag twee algemene doelstellingen:

- de kwaliteit van het Vlaamse onderwijssysteem bewaken en versterken;
- de interne kwaliteitszorg van scholen versterken.

Naast de vraag aan de deelnemers in welke mate de peilingen en paralleltoetsen de huidige doelstellingen realiseren, werd er in deze ex-ante evaluatie verkend of er nog andere mogelijke doelstellingen te realiseren zijn met peilingen en paralleltoetsen. Verder werd in deze ex-ante beleidsevaluatie verkend welke de randvoorwaarden zijn om deze doelstellingen te realiseren.

In de ex ante evaluatie kwamen de volgende doelstellingen aan bod:

- de kwaliteit van het Vlaams onderwijssysteem bewaken en versterken;
- de interne kwaliteitszorg van scholen versterken;
- de kwaliteit van scholen bewaken;
- het versterken van de geïnformeerde schoolkeuze;
- het verhogen van de gelijkwaardigheid van diploma's en getuigschriften;
- het versterken en ondersteunen van de lerende.

Voorafgaand aan de bespreking van deze doelstellingen, stellen we in de ex-ante beleidsevaluatie vast dat er nood is aan het expliciteren ervan. Er heerst immers

vandaag onduidelijkheid over wat er precies beoogd wordt met de peilingen.

In wat volgt bespreken we voor elk doelstelling hoe men er tegenover staat en in welke mate men vindt dat de peilingen en paralleltoetsen aan deze doelstelling kunnen bijdragen. Zo ja, welke verwachtingen impliceert dit dan naar het gebruik van de peilingen en de paralleltoetsen?

Hoewel we de doelstellingen één voor één bespreken, willen we er toch voor pleiten om de **samenhang en mogelijke interferenties tussen de doelstellingen** niet uit het oog te verliezen. Dit is ook een belangrijke bezorgdheid die geuit werd tijdens de focusgroepen. Algemeen beschouwd waarschuwt men ervoor om al te veel doelstellingen te willen bereiken aan de hand van één instrument. Men wijst onder meer op moeilijkheden wanneer men een ontwikkelingsgericht en verantwoordingsgericht probeert te combineren in één instrument.

4.2.1. Kwaliteit van het Vlaams onderwijssysteem bewaken en versterken

In deze ex-ante beleidsevaluatie was er een grote consensus dat het bewaken en versterken van de kwaliteit van het Vlaams onderwijssysteem een belangrijke doelstelling van de peilingen is. Ook naar de toekomst wil men deze doelstelling behouden.

Er werd aangegeven dat de peilingen een **belangrijke informatiebron** vormen om op Vlaams niveau een beeld te krijgen van de mate waarin leerlingen de verschillende eindtermen halen. Het gebruik van de peilingen op Vlaams niveau mag zich echter niet beperken tot informeren. De deelnemers gaven aan dat de peilingen op systeem – meer dan nu het geval is - gebruikt zouden kunnen worden om de kwaliteit van het onderwijs te **verbeteren en gerichte acties daartoe op te zetten**. Het is momenteel niet altijd duidelijk voor de deelnemers welke acties er op Vlaams niveau vanuit de peilingsresultaten genomen worden om de onderwijskwaliteit verder te versterken.

Om tegemoet te komen aan de nood om de peilingsresultaten sterker te benutten, werden in de ex-ante beleidsevaluatie de volgende voorstellen tot **versterkte analyse van de resultaten** aangehaald:

- het uitvoeren van secundaire analyses op de peilingsdata;
- het vergelijken van de peilingsresultaten met resultaten van koepelgebonden proeven en andere vergelijkbare toetsen;
- het vergelijken van de peilingsresultaten met resultaten van internationale onderzoeken;
- het vergelijken van de peilingsresultaten met de bevindingen van de onderwijsinspectie;
- het zoeken naar mogelijke verklaringen voor de resultaten bijvoorbeeld aan de hand van diepgaand verklarend onderzoek. Hierbij dient men

verder te kijken dan verklaringen op leerlingniveau, maar bijvoorbeeld ook het leerkracht- en schoolniveau mee te nemen;

- het identificeren van goede praktijkvoorbeelden waardoor scholen van mekaars praktijken kunnen leren.

Ook de **communicatie** over de peilingsresultaten kan volgens een aantal geconsulteerde belanghebbenden nog beter. Deze nood aan optimalisering betreft niet enkel communicatie ten aanzien van de direct betrokkenen, maar ook naar de ruimere maatschappij. Aandachtspunten die ze hierbij formuleren zijn de volgende:

- duidelijk communiceren over wat een peiling al dan niet is en wat men ervan kan verwachten en heldere situering van de doelgroep van de peilingen en de kenmerken van de doelgroep;
- in de communicatie niet de indruk wekken dat zaken die minder meetbaar zijn ook minder belangrijk zijn;
- communicatie tot op de klasvloer en eventueel zelfs tot bij ouders;
- communicatie op maat van de doelgroepen;
- communicatie over de verbeteracties die men onderneemt naar aanleiding van de resultaten (zie ook verder).

Naast een betere analyse en communicatie van de peilingsresultaten, werd er ook gevraagd om de **valorisatie en het gebruik van de peilingsresultaten op systeemniveau** verder te versterken. Nadat de peilingsresultaten bekend gemaakt zijn, zouden de verschillende betrokken partijen samen verbeteracties kunnen formuleren in een **actieplan op Vlaams niveau**. Mogelijke betrokken partijen zijn de Vlaamse overheid, de koepels en het GO!, de pedagogische begeleidingsdiensten, lerarenopleiders, scholen, schoolleiders, leerkrachten, Op deze wijze wil men de vertaling van de resultaten naar concrete acties op Vlaams niveau – met een explicitering van de rol van alle betrokkenen - versterken.

39

Hoewel het bewaken van de kwaliteit van het onderwijs door de meesten aanzien als een belangrijke doelstelling wordt de bezorgdheid geuit dat men de kwaliteit van onderwijs niet mag verengen tot enkel dat wat een peiling meet (zie ook verder).

4.2.2. De interne kwaliteitszorg van scholen versterken

Eén van de huidige doelstellingen van de peilingen en paralleltoetsen is het versterken van de interne kwaliteitszorg van scholen, vooral door het ter beschikking stellen van uitgebreide feedbackrapporten waarmee scholen (op schoolniveau) een gedetailleerd zicht krijgen op de leerlingenprestaties.

In deze ex-ante evaluatie vinden respondenten het belangrijk dat scholen over **goede instrumenten** beschikken om de eigen kwaliteit te bewaken en te versterken. De meeste

respondenten zijn van mening dat de peilingen en paralleltoetsen een mogelijk instrument voor interne kwaliteitszorg kunnen zijn naast vele andere instrumenten. Ze kunnen de school één mogelijke spiegel voorhouden m.b.t. de eigen kwaliteit, scholen uitdagen om hiervoor verklaringen te zoeken en acties op te zetten

Deze mening wordt evenwel niet door iedereen gedeeld. Sommigen pleiten voor voorzichtigheid in het gebruik van de huidige peilingen en paralleltoetsen als instrument voor interne kwaliteitszorg. Zo vindt men o. m. dat de peilingen en paralleltoetsen in hun huidige vorm te beperkt zijn om de volledige kwaliteit van een school in kaart te brengen. Men ervaart dat er maar een deel van de kwaliteit van onderwijs getest wordt en men beschouwt het als een momentopname van overwegend cognitieve kennis en vaardigheden. Sommigen geven ook aan dat de peilingen en paralleltoetsen in de huidige vorm niet volledig in lijn liggen met de opvatting over brede evaluatie. In de peilingen en paralleltoetsen wordt op een klassieke manier een summatieve evaluatie opgezet. Ook wordt er gewezen op het bestaan van andere, meer bruikbare instrumenten die eveneens de interne kwaliteitszorg kunnen ondersteunen, zoals o.m. de interdiocesane proeven- en OVSG-toetsen. Deze instrumenten zouden immers meer dan de peilingen en paralleltoetsen handvaten en ondersteuning bieden aan de scholen om met de resultaten aan de slag te gaan.

Hoewel de meeste respondenten het potentieel van de paralleltoetsen en peilingen als een mogelijke spiegel voor de school onderkennen, geeft iedereen aan dat scholen deze instrumenten momenteel (te) **weinig aanwenden** in het kader van hun interne kwaliteitszorg. Naast de hoger aangegeven redenen, ervaart men het feedbackrapport dat scholen ontvangen na een peiling of paralleltoets momenteel als te weinig toegankelijk om daadwerkelijk in te zetten voor interne kwaliteitszorg. Naast de complexiteit van de formulering, mist men ook vaak nuances, o.m. door meer rekening te houden met de specifieke context. Daarnaast blijkt uit de reacties dat scholen zelf niet steeds aandacht geven in hun schoolbeleid (en meer concreet het evaluatiebeleid) aan deze feedbackrapporten en er dus de mogelijkheden ervan niet ervaren. Bovendien blijkt dat om dergelijk feedbackrapport adequaat te gebruiken, ook nood is aan ondersteuning, zowel van binnen als buiten de school. Dergelijk gebruik moet immers gefaciliteerd worden, o.m. door het voorzien van tijd om dit door te nemen en te vertalen. Een andere reden voor het beperkte gebruik van de peilingen en paralleltoetsen is dat deze – in tegenstelling tot de koepelgebonden toetsen – niet op de leerplannen, maar op de eindtermen geënt zijn.

Om het ondersteunende karakter van de peilingen en paralleltoetsen te versterken, acht men het zinvol om n.a.v. de resultaten op deze toetsen op schoolniveau een **schoolspecifiek actieplan** op te stellen. In dit actieplan vormen de resultaten op de peilingen en paralleltoetsen één van de spiegels die scholen zich kunnen voorhouden en gebruiken om gerichte acties te formuleren. Deze actieplannen

zouden deel kunnen uitmaken van het strategisch beleid van een school en worden ingezet in de interne kwaliteitszorg.

Hierbij wordt het belang van betrokkenheid van alle actoren onderstreept, zodat er bij deze actoren een groot eigenaarschap kan ontstaan over dit actieplan. Volgens de deelnemers biedt dergelijk actieplan kansen om meer gerichte input en ondersteuning te krijgen van pedagogische begeleidingsdiensten en/of andere actoren.

Het mogelijk gebruik van de peilingsresultaten voor de **evaluatie van leerkrachten** werd algemeen als eerder onwenselijk geacht. Hooguit ziet men het als een aanleiding tot een gesprek in teamverband. Men acht het wel zinvol om de resultaten met het schoolteam te bespreken om zo te komen tot betere inzichten in leerprocessen en mogelijke verklaringen ervoor. Immers, niet één enkele leerkracht kan verantwoordelijk gesteld worden voor leerlingenresultaten want de leerlingen hebben reeds voorafgaand een leerproces doorlopen. Zo wordt in het basisonderwijs verwacht dat de eindtermen gerealiseerd worden op het einde van het zesde leerjaar. Hieraan kan doorheen de hele lagere schooltijd gewerkt worden. Men kan dan de leraar van het zesde leerjaar niet alleen verantwoordelijk achten voor de bekomen peilingsresultaten. Tot slot geeft men aan dat de huidige peilingen vooral momentopnames zijn.

4.2.3. De kwaliteit van scholen bewaken (inspectie)

Vandaag de dag heeft de inspectie geen inzage in de resultaten van scholen op peilingen of paralleltoetsen, tenzij de school er zelf voor kiest deze tijdens een doorlichting te delen. Het bewaken van de kwaliteit van individuele scholen is vandaag dus geen doelstelling van de peilingen.

In deze ex-ante beleidsevaluatie werden **zowel argumenten voor als tegen** de mogelijkheid om de inspectie toegang te geven tot de resultaten van individuele scholen geformuleerd.

De voorstanders benadrukken dat de inspectie dankzij deze specifieke toetsresultaten beschikt over bijkomende objectieve informatie om een beeld te kunnen vormen van de kwaliteit van een school. Een belangrijke voorwaarde hierbij is wel dat de inspectie doorheen het hele proces van de doorlichting deze toetsresultaten op schoolniveau gebruikt als één van de vele bronnen van informatie. Voorstanders geven ook aan dat op die manier de inspectie – eventueel samen met de pedagogische begeleidingsdiensten – de scholen beter zouden kunnen begeleiden. Ze beklemtonen dat de interne kwaliteitszorg niet mag losgekoppeld worden van de externe kwaliteitszorg (inspectie). Het is volgens hen immers wenselijk dat de acties die scholen nemen om zich te verantwoorden ten aanzien van de overheid (inspectie)

samenhangen met de acties die worden ontwikkeld om zichzelf te verbeteren. Scholen zouden op die manier meer inzicht krijgen in hun eigen kwaliteit en in mogelijke verklaringen voor hun kwaliteit. Er zouden volgens een aantal respondenten zo duidelijke en afgestemde indicatoren kunnen geformuleerd worden en concrete en meer gerichte aanknopingspunten tot verbetering aangegeven kunnen worden. Dit vereist volgens deze groep meer samenwerking tussen de inspectie en pedagogische begeleidingsdiensten.

Er waren echter ook tegenstanders van het gebruik van de toetsresultaten door de inspectie. Een eerste bezorgdheid die hierbij wordt geuit is dat men de kwaliteit van scholen niet mag verengen tot wat een peiling of paralleltoets meet. De tegenstanders wijzen tevens op de spanning tussen het ontwikkelings- en verantwoordingsperspectief en de mogelijke gevaren die het gebruik van de toetsresultaten door de inspectie met zich kan meebrengen. Indien de inspectie de toetsresultaten zou gebruiken als één element in de verantwoording van de kwaliteit van scholen, zou dit de ontwikkelingskansen van scholen kunnen hypothekeren. Bovendien leeft de vrees dat dit zou kunnen leiden tot teaching to the test met de mogelijke negatieve neveneffecten als gevolg (zoals strategisch gedrag, onbetrouwbaarheid van de toets, verenging van het gerealiseerde curriculum) (zie ook verder). Bij het gebruik van de resultaten van de peilingen of de paralleltoetsen tijdens de doorlichtingen wijzen de deelnemers op het risico dat scholen strategische keuzes zouden kunnen maken en bijvoorbeeld ervoor kiezen om enkel de toetsen af te nemen waarop ze verwachten hoog te scoren. Men concludeert dat de inhoud van de doorlichtingen niet mag bepaald worden door de keuze van de toetsen en omgekeerd.

Hoewel de voorstanders aangeven dat het gebruik van de resultaten op peilingen en paralleltoetsen door de inspectie zou kunnen leiden tot een betere afstemming tussen interne en externe kwaliteitszorg, wijzen de tegenstanders op mogelijke gevaren hiervan. Ze geven aan dat de vrijheid van scholen om aan interne kwaliteitszorg te doen mogelijks kan ingeperkt worden als de externe kwaliteitszorg hierop geënt wordt.

Volgens de tegenstanders zou de rol van de inspectie er eerder in moeten bestaan om bij de doorlichting na te gaan of de school een instrument (zoals de peilingen en paralleltoetsen) gebruikt en op welke manier dit wordt ingezet in het verbeteren van de eigen kwaliteit. Het onderwerp van de focus van een doorlichting zou dus moeten liggen op de verandering of verbetering van scholen.

4.2.4. Versterken van de geïnformeerde schoolkeuze

Vandaag de dag hebben ouders en leerlingen over het algemeen geen inzage in de peilingsresultaten van de school of van de leerlingen.

In deze ex ante evaluatie onderschrijft iedereen het **belang** van voldoende informatie over scholen voor ouders, o.m. om een goede schoolkeuze voor hun kinderen te maken. In deze ex-ante beleidsevaluatie werd verkend of en in welke mate peilingsresultaten hierbij een mogelijke informatiebron kunnen zijn.

De haalbaarheid en wenselijkheid om via peilingsresultaten de gewenste informatie aan ouders te bieden wordt betwijfeld. Er kunnen twee aspecten in het **voorbehoud** onderscheiden worden:

- Het gebruik van dergelijke resultaten van peilingen en paralleltoetsen voor schoolkeuze veronderstelt dat deze publiek gemaakt worden. Dergelijke openbaarheid van de resultaten kan uiteindelijk leiden tot een ranking en volgens sommigen een versterkte competitie tussen scholen. In elk geval wijst men op de nood aan voldoende informatie over de context van de school om voldoende genuanceerde interpretaties te kunnen maken (o.m. informatie over de leerlingkenmerken, ...) in geval van openbaarheid van gegevens.
- De peilingsresultaten hebben een beperkte focus. Het gaat telkens over een specifiek aspect dat getoetst wordt op één bepaald moment. Dergelijke resultaten bieden dan ook geen volledig beeld van de kwaliteit van de school. Er zijn namelijk nog andere aspecten die men van belang acht bij een schoolkeuze, zoals o.m. de schoolcultuur en –klimaat, de visie van de school op zorg en ondersteuning, het welbevinden van leerlingen,

Als mogelijk **alternatief** wordt voorgesteld om de informatie over scholen die momenteel reeds beschikbaar is (o.m. uit schooldoorlichtingen, OKI-indicatoren, ...) meer zichtbaar te maken en toegankelijker aan te bieden aan ouders.

4.2.5. Verhogen van de gelijkwaardigheid van diploma's en getuigschriften

Bij het verkennen van mogelijke doelstellingen en gebruikswijzen van peilingsresultaten, kwam ook het verhogen van de gelijkwaardigheid van diploma's en getuigschriften naar boven. Peilingsresultaten (weliswaar op leerlingniveau) kunnen interessante objectieve bijkomende informatie geven aan leerkrachten over het bereiken van de eindtermen en kunnen de beslissing van een klassenraad om een getuigschrift al dan niet uit te reiken ondersteunen. Het zou een objectiverend element kunnen toevoegen aan de evaluatie van de leerlingen. Sommigen geven ook aan dat dergelijke informatie leerkrachten kan ondersteunen in hun evaluatiebeleid en de werking van de klassenraad. Toch waren de respondenten duidelijk: de beslissing over het al dan niet toekennen van getuigschriften blijft steeds de verantwoordelijkheid van de klassenraad. Peilingsresultaten kunnen hier niet in de plaats treden.

Toch blijkt uit deze ex-ante evaluatie dat er weinig voorstanders zijn om met peilingen deze doelstelling te realiseren. Peilingsresultaten in de huidige context betreffen steeds een specifiek aspect op een bepaald moment en kunnen nooit een globale evaluatie van een leerlingen omvatten. Ze kunnen slechts als één element meegenomen worden in een globale evaluatie, maar deze evaluatie mag hiertoe niet beperkt worden. Immers, men vreest zo te kunnen evolueren naar een situatie waar de beslissingsgrond verengd wordt tot enkel het meetbare in een peiling (zie ook verder). Bovendien wordt ervoor gewaarschuwd dat het gebruik van de peilingen bij de evaluatie van leerlingen ongewenste gevolgen met zich kan meebrengen zoals teaching to the test, toetsangst, ...

Anderzijds werd door enkele respondenten aangegeven dat teaching to the test niet per se een probleem hoeft te zijn indien men zou inzetten op een beperkte set van minimumdoelen of basisdoelen waarvan men wil dat iedereen ze haalt en hiervoor een goede toets kan ontwikkeld worden.

Er wordt evenwel een **alternatief** voorgesteld: scholen gebruiken de peilingresultaten als een bron van reflectie over hun evaluatiebeleid en over de normen die scholen gebruiken voor het uitreiken van het getuigschrift en diploma's (en worden hierin bijgestaan door de inspectie: zie lager). Voor deze oefening zouden scholen gebruik moeten kunnen maken van hun feedbackrapporten, o.m. door zich te vergelijken met andere scholen, en zouden ze beroep moeten kunnen doen op de ondersteuning door de pedagogische begeleidingsdiensten.

4.2.6. Versterken en ondersteunen van de lerende

Hoewel de peilingen vandaag de dag geen rapportage op leerlingenniveau geven en dus niet op individueel niveau gebruikt kunnen worden om de lerende te ondersteunen en te versterken, bleek er bij een aantal respondenten wel een nood te zijn aan instrumenten om leerlingen beter op te volgen en dit vanuit een **ontwikkelingsgericht perspectief**.

De respondenten formuleerden de behoefte aan instrumenten om:

- de vooruitgang van leerlingen op te volgen;
- leerlingen, studenten en cursisten beter te oriënteren en te begeleiden, onder meer bij de overgang van basisonderwijs naar secundair onderwijs en bij de overgang naar hoger onderwijs;
- leerlingen meer inzicht te doen krijgen in hun eigen competenties en hun leerproces waardoor ze meer eigenaar worden van hun eigen onderwijstraject.

Een belangrijke voorwaarde is dat er **duidelijke leerlijnen** moeten uitgestippeld worden zodat de ontwikkeling en groei van leerlingen in kaart gebracht kan worden. Deze toetsen zouden ook ingepast moeten worden in het toets- en evaluatiebeleid van

scholen (zie ook verder). De respondenten gaven aan dat wanneer men vooral wil inzetten op het ondersteunen en versterken van de lerende er ruimte is om ook andere evaluatiemethodes (portfolio's, zelfevaluaties, ...) te exploreren en om meer formatief dan summatief te werken. Ze vragen zich ook af hoe deze ontwikkelingsgerichte doelstelling kan sporen met de andere doelstellingen van de peilingen die toch eerder summatief van aard zijn en een verantwoordingsaspect bevatten met het behalen van de eindtermen als norm.

Hoewel er dus wel een duidelijke behoefte geformuleerd werd aan instrumenten om de lerende te ondersteunen en versterken, was het niet duidelijk of de peilingen en paralleltoetsen hiervoor het meest aangewezen instrument zijn. Redenen die hiervoor aangegeven werden zijn mogelijke contradicties met de huidige doelstellingen (zie boven) en het huidige opzet (summatieve karakter, toetsfrequentie, norm....) van de peilingen en paralleltoetsen. De peilingen en paralleltoetsen zouden dus eerder gezien kunnen worden als één onderdeel in een ruimer beleid om de lerende meer te ondersteunen en te versterken. Dit zou wel enkele fundamentele aanpassingen vergen in de manier waarop er gepeild wordt en er feedback wordt gegeven (zie verder).

4.3. Hoe kan de peiling als instrument er in de toekomst uitzien?

4.3.1. Hoe verhouden de peilingen zich tot andere toetsen?

45

In deze ex ante beleidsevaluatie werd het belang van **afstemming tussen (aanbieders van) diverse toetsen**, o.m. van de overheid en de onderwijskoepels, benadrukt. Er is nood aan afstemming inzake de inhoud (verhouding van de peilingen en paralleltoetsen tot de koepelgebonden toetsen), afname (tijdstip van afname), analyse en aard van gebruik. Deze discussie situeert zich voornamelijk in het basisonderwijs, waar er verschillende toetsen bestaan. In het secundair onderwijs zijn er enkel peilingen en geen koepelgebonden toetsen.

Algemeen beschouwd formuleert men als voordeel van een peiling dat er een cesuur bepaald wordt op basis van de eindtermen die netoverschrijdend geldt voor alle scholen, wat niet het geval is bij koepelgebonden toetsen.

Met betrekking tot afstemming over de **inhoud** konden dan ook twee pistes worden geïdentificeerd. Ten eerste werd aangehaald dat de verschillende instrumenten geïntegreerd konden worden tot een algemeen en gestandaardiseerd instrument. Er zou dan sprake kunnen zijn van een uniform en gestandaardiseerd instrument opdat scholen zich kunnen vergelijken met andere scholen en meer inzicht krijgen in de eigen kwaliteit. Idealiter wordt er goed samengewerkt tussen de verschillende actoren om tot dit

geïntegreerd instrument te komen. Hierbij wijst men op het belang van het erkennen van de contextualisering van leerresultaten i.f.v. een schoolspecifieke interne kwaliteitszorg.

Ten tweede werd ook de mogelijkheid geopperd om de verschillende testen complementair in te zetten. Men biedt scholen als het ware een waaier van kwaliteitsvolle toetsen aan in plaats van één gestandaardiseerde test.

Naast peilingen wordt vaak ook deelgenomen aan andere onderzoeken. Uit deze ex-ante beleidsevaluatie blijkt dat een geïntegreerde analyse van de peilingsresultaten, samen met andere testresultaten, wenselijk geacht wordt. Tegelijk wordt hierbij de nood aan ondersteuning en verdere professionalisering (i.c. inzake datageletterdheid) geformuleerd.

In elk geval blijkt uit deze ex-ante beleidsevaluatie dat de verhouding tussen de (doelstellingen van de) diverse toetsen uitklaring vereist. Bovendien is duidelijk dat men de opportuniteit onderschrijft om met peilingsresultaten het gesprek aan te gaan met andere scholen (zelfs niveauoverstijgend), lerarenopleiders en begeleiders over de resultaten.

4.3.2. Wie neemt deel aan de peilingen?

46

Welke scholen? Welke leerlingen?

In het huidige systeem wordt een representatieve set van scholen geselecteerd om deel te nemen aan de peilingen. Scholen zijn evenwel niet verplicht om hieraan deel te nemen. De deelname aan de paralleltoetsen is volledig vrijwillig voor alle scholen (zie 1.3).

De gewenste deelname van scholen en leerlingen aan de peilingen en paralleltoetsen kan verschillen naargelang de doelstellingen die men wil bereiken. Sommige doelstellingen lijken een verplichte deelname van scholen te impliceren, bijvoorbeeld het gebruik van toetsresultaten door inspectie of bij certificering. In andere doelstellingen lijkt dan weer het belang van vrijwillige deelname van scholen en leerlingen primordiaal, bijvoorbeeld bij het gebruik in kader van interne kwaliteitszorg of bij het versterken en ondersteunen van de lerende.

In de ex ante beleidsevaluatie was er geen eenduidigheid over al dan niet vrijblijvende deelname.

Een aantal respondenten zagen een aantal voordelen in een verplichte deelname van alle

scholen: Zij gaven hiervoor onder meer de volgende redenen aan:

- het is belangrijk dat de kwaliteit van alle scholen versterkt wordt en alle scholen dus de kans krijgen om deel te nemen aan de peilingen en paralleltoetsen;
- elke school dient gelijk behandeld te worden (gelijkheidsprincipe);
- dit komt de leerlingen ten goede en alle leerlingen hebben hier recht op;
- de verplichte deelname kan een krachtig middel zijn om het beleid bij te sturen;
- de verplichte deelname biedt het voordeel dat de eindtermen centraler komen te staan en dat dit de kennis van de eindtermen bij leerkrachten (en scholen in het algemeen) zal bevorderen.

Er waren echter ook een heel aantal tegenstanders van een verplichte deelname van alle scholen. Zij vrezen dat een verplichte afname van de peilingen bij alle scholen het risico inhoudt dat scholen hun resultaten openbaar moeten maken (cf. openbaarheid van bestuur) met een ranking van scholen tot gevolg. Ze wijzen er ook op dat dit bijkomende stress kan veroorzaken bij leerlingen, ouders en scholen. Een aantal deelnemers stellen zich ook vragen bij het verplicht maken van de deelname aan één toets. Deze deelnemers kunnen zich eerder vinden in het feit dat scholen zouden moeten verplicht worden om een toets af te nemen maar waarbij ze zelf kunnen kiezen uit een toetsenbank waarbij dan de kwaliteit van alle aangeboden toetsen gegarandeerd wordt.

Hoewel er dus zeker geen consensus was over een al dan niet verplichte deelname van alle scholen, waren er wel deelnemers die ook vragen hadden bij de effecten van de huidige vrijblijvende deelname aan de peilingen op de representativiteit van de steekproef. Zij wijzen op een mogelijke vertekening van de resultaten en waren er voorstander van om scholen te verplichten tot deelname indien ze geselecteerd worden via de steekproef.

Uit de ex-ante beleidsevaluatie blijkt dus dat het noodzakelijk is om – rekening houdend met de (ev. combinatie van) geformuleerde doelstellingen – expliciete keuzes te maken inzake al dan niet verplichte deelname van scholen aan peilingsproeven. Ook is het belangrijk om deze keuze duidelijk te communiceren en de scholen te blijven wijzen op hun verantwoordelijkheid hierin.

De ex ante evaluatie focuste vooral op de vraag of scholen vrijblijvend kunnen deelnemen (zie boven). De concrete implicaties van de verschillende mogelijke doelstellingen van de peilingen voor de deelname van leerlingen (verplicht/vrijblijvend, aantal leerlingen, ...) werd in deze ex ante evaluatie niet systematisch besproken. Dit dient nog verder uitgeklaard te worden.

Wat alvast werd aangegeven is het belang om gegeven het M-decreet bij de afname van de toetsen in de toekomst meer rekening te houden met leerlingen met specifieke onderwijsbehoeften. Dit zou bijvoorbeeld kunnen door redelijke aanpassingen toe te staan bij het afleggen van de toets.

Kunnen peilingen ook nuttig zijn buiten het gewoon leerplichtonderwijs?

Vandaag de dag zijn er enkel peilingen in het gewoon leerplichtonderwijs. In deze evaluatie werd ook nagegaan of er een behoefte is aan peilingen buiten het gewoon leerplichtonderwijs.

Vooreerst vroegen we aan de respondenten of het wenselijk zou zijn om peilingen in het **buitengewoon onderwijs** te organiseren. De respondenten meenden dat het wel zinvol zou zijn om peilingen af te nemen in het basisaanbod, type 9, OV3 en OV4. In het basisaanbod en type 9 wordt vandaag niet gewerkt a.d.h.v. eindtermen, hoewel het vaak wel de bedoeling is dat deze leerlingen doorstromen naar het gewoon onderwijs. Voor deze groep leerlingen zou men de peilingen wel kunnen gebruiken. Anderzijds zijn er mogelijks bepaalde ontwikkelingsdoelen die peilbaar zijn. Men zou dus kunnen nagaan hoe ver een leerling uit het buitengewoon onderwijs staat op vlak van het bereiken van eindtermen en ontwikkelingsdoelen. Belangrijke voorwaarde hierbij is dat leerlingen hulpmiddelen mogen gebruiken. Het dient dus zeker verder onderzocht te worden hoe de peilingen of paralleltoetsen zouden kunnen gebruikt worden voor relevante doelgroepen in het buitengewoon onderwijs.

Daarnaast werd de vraag gesteld of het zinvol zou zijn om peilingen in te richten in het **hoger onderwijs**. Over het algemeen waren hier geen voorstanders voor en wel om de volgende redenen:

- in het hoger onderwijs is een andere soort verantwoording nodig;
- instellingen voor hoger onderwijs kunnen zelf hun eigen kwaliteitsbewaking organiseren;
- het hoger onderwijs werkt niet met eindtermen wel met meer generieke leeruitkomsten

We moeten deze resultaten wel met enige voorzichtigheid interpreteren aangezien dit niet de hoofdfocus was van deze evaluatie en dit onderwerp dus slechts beperkt aanbod is kunnen komen.

De peilingen in het leerplichtonderwijs vinden de respondenten uit hoger onderwijs wel zinvol omdat deze informeren over de kwaliteit van het Vlaams onderwijs en over de manier waarop eindtermen geïnterpreteerd kunnen worden. Verschillende deelnemers uit de lerarenopleidingen gaven aan dat ze vandaag de peilingen gebruiken in hun eigen werking (als didactisch (toets)materiaal, om lessen aan te passen zeker in het geval van minder goede resultaten op de peilingen). Men is vanuit het hoger onderwijs ook vragende partij om meer samen te werken met het secundair onderwijs. Zo zouden de startcompetenties in het hoger onderwijs beter kunnen afgestemd worden met de eindtermen van de derde graad secundair onderwijs.

Ten slotte zijn we nagegaan of peilingen zinvol zouden kunnen zijn in het **volwassenenonderwijs** (centra voor volwassenenonderwijs en centra voor basiseducatie). Hoewel dit niet de hoofdfocus was van dit onderzoek en deze vraag dus slechts beperkt aan bod is gekomen geven we toch kort mee wat de respondenten hierover dachten. Men is van mening dat peilingen voor sommige opleidingen relevant kunnen zijn, bijvoorbeeld voor meer beroepsgerichte opleidingen, het tweedekansonderwijs en NT2-opleidingen. Het tweedekansonderwijs vindt men het meest voor de hand liggend gezien de parallellen met het voltijds secundair onderwijs. Voor andere opleidingen lijkt het volgens de deelnemers moeilijker haalbaar gezien de grote diversiteit tussen cursisten en centra en door de modulaire werking.

Men ziet in het volwassenenonderwijs wel voordelen in het systeem van de peilingen en dan voornamelijk in het kader van het gebruik ervan in de interne kwaliteitszorg. Eén van de belangrijke randvoorwaarden is dat men bij het onderling vergelijken tussen de centra er voldoende rekening wordt gehouden met de context (in doelgroepen en regio's). Een andere voorwaarde is dat het leercontinuüm mee in kaart wordt gebracht. Peilingen zouden zowel een start- als een eindmoment van het traject in kaart kunnen brengen en gebruikt worden bij de verdere oriëntering van cursisten.

Hoewel dit geen uitdrukkelijk thema was dat met alle stakeholders in de ex-ante beleidsevaluatie besproken werd, suggereerden enkele leerlingen tot slot om na te gaan in welke mate peilingen bij leraren naar hun kennis en vaardigheden wenselijk en haalbaar zouden zijn.

4.3.3. Wat moeten peilingen meten?

Welke kennis, vaardigheden en attitudes?

Peilingen zijn opgesteld om na te gaan in welke mate leerlingen de eindtermen en ontwikkelingsdoelen beheersen (zie 1.3)

Algemeen geeft men aan dat het meten van vakinhoudelijke competenties belangrijk blijft. Verschillende deelnemers gaven hierbij aan dat ze het belangrijk vinden dat men streeft naar het peilen van een **brede diversiteit** aan vakken in verschillende onderwijsvormen en studierichtingen. De deelnemers waarschuwen voor het feit dat er verkeerdelijk de indruk kan gewekt worden dat bepaalde vakken belangrijker zijn dan andere indien men niet voldoende breed peilt. Anderzijds zijn er ook wel deelnemers die ervoor pleiten om aan bepaalde 'basisvakken' meer aandacht te besteden.

Hoewel **vakinhoudelijke competenties** belangrijk worden geacht, vroegen verschillende respondenten om daarnaast ook **vakoverschrijdende competenties** te meten. Hierbij werd verwezen naar verschillende mogelijke kaders om dit verder vorm te geven waaronder de 21^{ste}-eeuwse competenties, de Europese sleutelcompetenties, basisgeletterdheid, ... Voorbeelden van competenties die gegeven werden zijn onder meer informatieverwerving en –verwerking, kritisch nadenken, leren leren en onderzoekscompetenties, metacognitieve vaardigheden, samenwerken en zelfredzaamheid.

Er was een grote consensus onder de respondenten dat de peilingen niet alleen aandacht mogen hebben voor kennis en vaardigheden maar ook voldoende aandacht moeten besteden aan **socio-emotionele vaardigheden en attitudes, motivatie en welbevinden**. Er leefde echter wel duidelijk het besef dat dit technisch niet eenvoudig is om te meten in een grootschalig onderzoek. Er werd aangegeven dat hierbij ook dient nagedacht te worden over alternatieve evaluatiemethodes zoals observaties of dagelijkse evaluatie van leerkrachten.

Bij het meten van attitudes stelden sommige deelnemers zich wel de vraag hoe men hiermee in de peilingen kan omgaan aangezien het hier gaat over een inspanningsverplichting en deze doelen dus niet te beheersen zijn. Deze bezorgdheid hangt uiteraard nauw samen met de doelstellingen waarvoor de peiling gebruikt wordt.

We kunnen concluderen dat de meeste respondenten vinden dat de peilingen zowel kennis, vaardigheden als attitudes zouden moeten meten. Daarnaast gaven verschillende deelnemers het belang aan van het geïntegreerd meten van kennis, vaardigheden en attitudes (competentiegericht meten). Men merkt daarbij wel op dat het een uitdaging is om dit te realiseren in een grootschalig onderzoek.

Waar ligt de norm?

In de huidige peilingen wordt de norm van het al dan niet behalen van de eindtermen bepaald in een proces van cesuurbepaling (zie 2.2).

Er werd door een aantal deelnemers vragen gesteld bij de **methode van cesuurbepaling** en er werd voor gepleit om deze methode verder te optimaliseren en de subjectiviteit hierbij te verminderen.

De deelnemers benadrukten ook het belang van **voldoende expertise bij de toetsontwikkeling**, niet enkel methodologisch, maar ook op inhoudelijke expertise op het vlak van de eindtermen en kennis van de onderwijsrealiteit. Er werd gevraagd om het onderwijsveld voldoende te betrekken bij het opstellen van de toetsen.

Een aantal deelnemers gaven aan dat de peilingen en paralleltoetsen nu een uitspraak doen over een minimumniveau maar dat deze informatie beperkt is. Het zou ook interessante beleidsinformatie opleveren als er naast het minimumniveau ook beheersingsniveaus in kaart zouden kunnen gebracht worden.

Welke achtergrondvariabelen moeten we meenemen?

In de huidige peilingen worden leerling-, school en leerkrachtkenmerken in kaart gebracht onder meer aan de hand van achtergrondvragenlijsten (zie 2.2).

Om de peilingsresultaten beter te kunnen contextualiseren en duiden, gaven de respondenten aan dat het wenselijk is om niet enkel leerlingenkenmerken in kaart te brengen, maar daarnaast ook zicht te krijgen op leerkracht- en schoolkenmerken alsook op kenmerken van het onderwijsproces (zoals o.a. lespraktijken en –methodes, evaluatiepraktijken). Dit is vooral belangrijk om de resultaten te kunnen begrijpen en op zoek te gaan naar mogelijke verklaringen. Dit ondersteunt ook het opzetten van gerichte verbeteracties.

4.3.4. Op welke momenten moet er gepeild worden en met welke frequentie?

In de huidige situatie worden er peilingen afgenomen op het einde van het lager onderwijs en aan het einde van elke graad van het secundair onderwijs. De peilingen zijn steeds een momentopname. Er is geen beginmeting. Er zijn wel herhalingspeilingen die het mogelijk maken evoluties in kaart te brengen (zie 1.3).

De keuze van de meetmomenten en de frequentie van de peilingen hangt uiteraard nauw samen met de doelstellingen die men wil bereiken. Zo vindt men het in kaart brengen van de ontwikkeling van leerlingen aan de hand van meerdere meetmomenten belangrijk wanneer men het heeft over het gebruik van peilingen op school- en leerlingniveau. Er bestaat echter minder consensus in welke mate dit ook nodig is voor de kwaliteitsbewaking op Vlaams niveau.

De deelnemers gaven aan dat ze peilingen nodig vinden op belangrijke ijkpunten in de onderwijsloopbaan. In het basisonderwijs vinden de meeste deelnemers dat het einde van het zesde leerjaar een belangrijk moment is om te peilen. Dit is immers een belangrijk eindpunt in de onderwijsloopbaan en markeert de overgang naar het secundair onderwijs.

In het secundair onderwijs is er consensus dat er minstens aan het einde van de eerste en de derde graad dient gepeild te worden aangezien deze momenten in het secundair onderwijs als het belangrijkste aanzien worden. Een aantal

deelnemers vindt dat peilingen zouden moeten meten op alle momenten waar er eindtermen geformuleerd zijn, dus ook aan het einde van de 2^{de} graad.

Verschillende deelnemers kaarten de nadelen aan van het meten op één moment, zoals vandaag gebeurt. Zij zijn voorstander om minstens een begin- en een eindmeting te voorzien zodat de ontwikkeling, evolutie en leerwinst in kaart kunnen gebracht worden. Dit zowel op school- als op leerlingniveau. Op die manier kunnen de resultaten beter ingezet worden in de interne kwaliteitszorg en bij het ondersteunen en versterken van de lerenden.

Op Vlaams niveau zijn de respondenten vragende partij om meer systematische **herhalingspeilingen** te organiseren om zo de evolutie en de ontwikkeling van het Vlaams onderwijs beter in kaart te brengen. Vooral in het secundair onderwijs vindt men dat er momenteel te weinig herhalingspeilingen zijn. Er dient verder onderzocht te worden wat een zinvolle frequentie zou kunnen zijn. Er wordt door een aantal deelnemers gevraagd hierbij een onderscheid te maken tussen peilingen die de basisvakken of -competenties meten (bijvoorbeeld talen, wiskunde,) en andere peilingen.

4.3.5. Naar digitale toetsen?

De huidige peilingen en paralleltoetsen bestaan in hoofdzaak uit pen- en papiertoetsen.

52

Verschillende respondenten hebben gepleit om te evolueren naar computer gebaseerde toetsen. Digitale toetsen zouden niet alleen de afname van de peilingen vergemakkelijken maar zouden ook toelaten om sneller feedback te voorzien. Ook de duur van de afname zou korter kunnen door adaptief te testen. Dit kan testmoeheid voorkomen.

Bij de paralleltoetsen dienen scholen zelf hun gegevens in te voeren, wat tijdrovend is. Dit kan voor extra planlast zorgen en met zich meebrengen dat de paralleltoetsen minder gebruikt worden. Een digitalisering zou hierop een antwoord kunnen bieden.

4.3.6. Welke feedback krijgt men na deelname aan de peiling?

Scholen die deelnemen aan peilingsproeven en paralleltoetsen ontvangen na deelname een feedbackrapport, waarin de resultaten op schoolniveau gepresenteerd worden (zie 1.3).

Zoals reeds aangehaald, maken scholen momenteel (te) weinig gebruik van de feedbackrapporten op basis van de peilingen. Hiervoor wordt

vaak verwezen naar de beperkte gebruiksvriendelijkheid van het feedbackrapport: complexiteit van het rapport, moeilijk en eerder wetenschappelijk taalgebruik, weinig concrete handvaten, te weinig afstemming op de school zelf... Een algemene verwachting in deze ex-ante beleidsevaluatie is dat de feedbackrapporten meer op maat van de school opgesteld worden zodat deze sterker ingezet kunnen worden voor de interne kwaliteitszorg. Meermaals wordt er ook gepleit om de datageletterdheid van leraren(teams) te versterken (zie ook verder).

Momenteel wordt er in de schoolfeedbackrapporten geen feedback gegeven op item- en leerlingenniveau. Dit staat volgens de respondenten een gerichte foutenanalyse in de weg. Dit werd ook aangehaald als een reden waarom de feedbackrapporten minder gebruikt worden. In deze ex-ante evaluatie was er een grote eensgezindheid om naast informatie op schoolniveau ook informatie op leerlingniveau te voorzien. Deze informatie is nodig om tegemoet te komen aan diverse doelstellingen zoals o.m. het versterken en ondersteunen van de lerende en het verhogen van de gelijkwaardigheid van diploma's en getuigschriften. Er kan op basis van de leerlingenspecifieke resultaten niet enkel gerichte ondersteuning aan het leerproces en studiekeuzeproces gegeven worden, maar leerlingen krijgen door deze feedback ook zicht op hun eigen competenties. Ze kunnen op deze manier het eigenaarschap over hun leerproces verhogen. Men gaat ervan uit dat door deze rechtstreekse feedback de motivatie van de leerlingen kan verhogen om zich voor de toets in te zetten.

53

Zoals reeds aangehaald, veronderstelt men dat dergelijke feedback aan de leerlingen op leerlingniveau de motivatie tot deelname van de leerlingen aan dergelijke toetsen kan verhogen en dus ook de kwaliteit van de antwoorden ten goede kan komen. Zo kan informatie op leerlingniveau dan ook – zij het onrechtstreeks - van belang zijn voor de realisatie van de doelstelling “de kwaliteit van het Vlaamse onderwijs versterken”.

Er wordt gepleit dat niet enkel leerlingen kunnen beschikken over hun eigen resultaten, maar dat deze informatie op leerlingniveau ook toegankelijk is voor leraren(teams). Dit is niet enkel belangrijk om zo specifieke handvaten te krijgen om het leerproces én de studiekeuze van de leerlingen te ondersteunen, maar dergelijke terugkoppeling op leerlingniveau is eveneens zinvol in het kader van het versterken van interne kwaliteitszorg; er kan dan immers zo gericht aan de slag gegaan worden op schoolniveau (zie ook eerder, het opstellen van een schoolspecifiek actieplan).

Een cruciale voorwaarde hiertoe is dat de resultaten op leerlingniveau voldoende snel (bij voorkeur binnen hetzelfde schooljaar) beschikbaar zijn zodat deze nog gericht kunnen aangewend worden om het leerproces te ondersteunen en leerlingen snel een terugkoppeling krijgen over hun eigen resultaten. Ook een uitgebreide kadering van de resultaten is onontbeerlijk.

Momenteel worden de feedbackrapporten aan de scholen bezorgd. Vanuit hun rol als ondersteuners van schoolontwikkeling zijn pedagogische begeleidingsdiensten vragende partij om ook over deze feedbackrapporten te beschikken. Er was echter geen eensgezindheid over de mogelijkheid dat pedagogische begeleidingsdiensten vanzelfsprekend alle resultaten van de scholen die ze begeleiden, ontvangen. Hoewel pedagogische begeleidingsdiensten dit zinvol achten om hun ondersteuningsaanbod nog beter te kunnen afstemmen op de vastgestelde noden, wijzen anderen op het belang van eigenaarschap door de school zelf.

Hoewel deze vraag niet frequent aan bod kwam, blijkt uit deze ex-ante evaluatie dat een uitklaring van de gebruikers van de rapporten noodzakelijk is.

4.4. Wat zijn noodzakelijke randvoorwaarden om de doelstellingen te bereiken?

Om met de peilingen de vooropgestelde doelstellingen te realiseren, zijn er verschillende randvoorwaarden nodig. Hieronder worden de belangrijkste randvoorwaarden die uit de ex-ante beleidsevaluatie naar voren kwamen, (willekeurig) opgesomd.

In deze ex-ante beleidsevaluatie is men het erover eens dat peilingen een geschikt en legitiem middel zijn om uitspraken te doen over het al dan niet beheersen van de gemeten eindtermen, als één dimensie om zicht te krijgen op de kwaliteit van het Vlaamse onderwijs. Doorheen de hele ex-ante beleidsevaluatie werd echter de zorg geformuleerd dat onderwijskwaliteit niet louter verengd mag/kan worden tot peilingsresultaten alleen. Er is een grote bezorgdheid dat aspecten die minder of nauwelijks meetbaar zijn minder belangrijk geacht zouden worden.

Er wordt dan ook gevraagd om de mogelijkheden en beperktheden van de peilingen als instrument om de kwaliteit van het Vlaamse onderwijs in kaart te brengen, duidelijk en blijvend te communiceren.

Het belang van **communicatie** werd voortdurend beklemtoond. Niet alleen de doelen, maar ook een duidelijke contextualisering van de resultaten is noodzakelijk. Men wijst op het belang van een heldere situering van de doelgroep van de bevraging (welk onderwijsniveau, onderwijsvorm, studierichting,) zodat dit voldoende gekaderd kan worden en geen onterecht generaliserende uitspraken gedaan worden. Tot slot wijst men op de noodzaak van communicatie over mogelijke vervolgacties naar aanleiding van de resultaten.

Een ander generieke randvoorwaarde is vertrouwen in scholen en in de professionaliteit van leraren. Het gaat hierbij niet enkel om professionaliteit in het dagdagelijks handelen

in de (klas-) praktijk, maar uit de ex-ante evaluatie blijkt dat verkend kan worden op welke wijze de expertise van leraren doorheen het hele peilingsonderzoek – ruimer dan binnen de eigen klas- en schoolcontext - kan benut worden, bijvoorbeeld bij de toetsontwikkeling en bij de valorisatie van de resultaten op Vlaams niveau. Zoals reeds aangegeven, wijst deze ex-ante beleidsevaluatie ook op het belang van een sterk beleidsvoerend vermogen van scholen om met de peilingsresultaten aan de slag te gaan. De school heeft immers de regierol in het omgaan met de feedbackrapporten. Dit vraagt ook een goede datageletterdheid om het feedbackrapport goed te interpreteren. Ondersteuning hierbij is dan ook van cruciaal belang (zie ook verder).

Zoals reeds aangegeven, is men te vinden voor het gebruik van objectieve instrumenten (zoals de peilingen en paralleltoetsen) om het oordeel van de klassenraad te ondersteunen en mee te nemen als één element in de beoordeling van het uitreiken van diploma's en getuigschriften. Dit wil evenwel niet zeggen dat men wil evolueren naar centrale toetsen als middel tot certificering. De beslissing dient steeds bij de klassenraad te blijven liggen (zie hoger) en de professionaliteit van de leraren moet erkend blijven. Immers, elke toets (en dus ook een peiling) kan slechts een beperkt beeld geven van de competenties van een leerling.

In deze ex ante beleidsevaluatie was bovendien nagenoeg iedereen het erover eens dat een evolutie naar een centrale toetsing niet wenselijk is. Men vreest immers negatieve gevolgen van een ranking n.a.v. centrale toetsing indien de resultaten van de scholen openbaar gemaakt worden en/of alle scholen verplicht zijn om deel te nemen aan de peilingen (zie ook eerder).

Een andere mogelijk neveneffect van centrale examens waarvoor de deelnemers waarschuwen, is teaching to the test. Indien de peilingen zouden gebruikt worden bij certificering van leerlingen dan bestaat het risico dat leerkrachten het curriculum gaan reduceren tot wat de peiling meet. Een peiling zoals ze nu geconcipieerd wordt, kan evenwel niet alles meten.

Doorheen de hele ex-ante beleidsevaluatie kwam de bezorgdheid naar voren dat we een evolutie naar een toetscultuur moeten vermijden. Men waarschuwt voor de negatieve gevolgen van een situatie waarin leerlingen te frequent getoetst zouden worden. Hierbij verwijst men naar o.m. overtesting, stress bij de leerlingen, het gebrek aan coherentie en taakbelasting, praktische haalbaarheid en mogelijke planlast bij leraren.

Zoals reeds eerder aangehaald, wezen respondenten dan ook het belang van het inpassen van peilingen in het evaluatiebeleid van de school. Algemeen beschouwd, wordt in deze ex-ante beleidsevaluatie het belang van breed evalueren sterk benadrukt, met o.m. aandacht voor zowel formatieve als summatieve evaluatie, meerdere evaluatiemomenten, meerdere en diverse instrumenten, ... Men

pleit ervoor om blijvend te zoeken naar andere (relevante) instrumenten zodat men een breed zicht krijgt op de competenties van leerlingen. Het exploreren van voorbeelden in andere landen wordt hierbij aangemoedigd.

Opdat de peilingen en paralleltoetsen van waarde kunnen zijn voor de leerlingen en de schoolontwikkeling (interne kwaliteitszorg), is eigenaarschap door de scholen een vereiste. In deze ex-ante beleidsevaluatie werd gewezen op de ervaring dat de peilingen(kalender) top down bij scholen aankomen (scholen kiezen niet zelf aan welke peiling ze deelnemen – de paralleltoetsen kunnen ze wel zelf kiezen), zonder dat dit afgestemd is op de prioriteiten of noden van de school (incl. leerlingen). Daarom pleit men ervoor om de autonomie van de scholen in het gebruik van de toetsen maximaal mogelijk te maken zodat deze een grotere waarde hebben voor o.m. de interne kwaliteitszorg in scholen. Zo werd aangehaald dat een school autonoom zou moeten kunnen beslissen over de frequentie van testen (en het moment van afname). Op die manier kan de school zelf kiezen waarop ze focust en welke competenties ze wil evalueren in het kader van haar interne kwaliteitszorg. Dit bevordert het eigenaarschap van de school.

Tot slot is een belangrijke randvoorwaarde voor het optimaal gebruik van de resultaten door scholen dat er voldoende ondersteuning voorzien wordt. Dit vereist niet alleen het inzetten op datageletterdheid van scholen maar ook werken aan de capaciteit van scholen om effectief met de resultaten aan de slag te gaan en in te zetten op hun interne kwaliteitszorg.

Samenhangend met die ondersteuning én de nood aan eigenaarschap, is het cruciaal dat schoolfeedbackrapporten meer op maat van de school geschreven worden (zie ook eerder). Daarmee wordt o.a. bedoeld dat de rapporten meer gebruiksvriendelijk kunnen gemaakt worden en aangepast worden aan de noden van scholen. Verdere afspraken over deze feedbackrapporten met het oog op maximale aanwending door de scholen, moeten dan ook gemaakt worden.

HOOFDSTUK 5: ALGEMENE CONCLUSIES

5.1 Situering

We besluiten dit rapport met een overzicht van de conclusies die overkoepelend naar voor komen op basis van de beide evaluaties. Aangezien het opzet van de beide evaluaties fundamenteel verschillend is, is het niet steeds mogelijk om deze conclusies consequent op beide evaluaties te baseren. Sommige, vooral toekomstgerichte, vragen kwamen uitgebreid aan bod in de ex ante evaluatie (hoofdstuk 4); het bekijken van het functioneren van het lopende steunpunt was dan weer de focus van de evaluatie van de beheersovereenkomst (hoofdstuk 3).

Wanneer een aspect duidelijk naar voor komt, hetzij in één hetzij in beide evaluaties, heeft het een plaats gekregen in dit afsluitende hoofdstuk. Deze conclusies houden geen rekening met de werkelijke haalbaarheid van de eraan gekoppelde toekomstgerichte veranderingen en zijn dus geen beleidsaanbevelingen vanwege de administratie.

5.2 Doelstellingen

Het vastleggen van de te bereiken doelstellingen is stap één bij het bepalen van de contouren van de toekomstige peilingen en vormt dan ook het eerste deel van de conclusies. De keuze van de doelstellingen heeft vanzelfsprekend implicaties op het mogelijk gebruik en de opzet van de peilingen; waar mogelijk worden deze gekoppeld aan de doelstelling.

Uit de evaluaties blijkt vooreerst dat het combineren van al te veel doelstellingen binnen het instrument peilingen geen goed idee is. Het is allicht in de praktijk niet haalbaar om een instrumentarium op te bouwen en te implementeren dat voor alle doelstellingen voldoende geschikt is. We brengen in deze conclusie een duidelijke prioritering aan in de doelstellingen.

5.2.1 Peilingen inzetten als beleidsinstrument

De primaire doelstelling van de peilingen is het leveren van informatie op systeemniveau. De ideeën die daarover bestaan, wijzen allemaal in dezelfde richting. Wat de leerlingenprestaties op het vlak van het bereik van de eindtermen en ontwikkelingsdoelen betreft, vormen de peilingen zelfs een unieke bron. De beschikbare informatie is daarbij ook van hoge kwaliteit en dit moet in de toekomst absoluut behouden blijven.

De resultaten van het peilingsonderzoek kunnen nog meer gebruikt worden in de voorbereiding van het onderwijsbeleid, waardoor de aanzienlijke investering in het peilingsonderzoek een grotere return on investment zou krijgen.

Om dit mogelijk te maken moet er meer afstemming komen tussen de verwachtingen van de beleidsmakers en het gehanteerde instrumentarium.

Toekomstgericht zou er voor kunnen geopteerd worden om bij de toetsontwikkeling te werken met meerdere beheersingsniveaus in plaats van het plaatsen van één cesuur tussen het wel of niet behalen van de eindtermen.

Ook dient te worden gedacht aan het inzetten van adaptieve toetsen die er voor kunnen zorgen dat er meer nuance komt in de testresultaten.

Om meer gebruik te maken van de informatie die voortkomt uit de achtergrondvragenlijsten, dient de opdrachtgever meer duidelijke beleidsvragen te formuleren die daarna methodologisch vertaald worden door de opdrachthouder (zie ook 5.5 rond de inhoud van het peilingsonderzoek).

Nadat de resultaten beschikbaar zijn, kan werk gemaakt worden van meer verklarende analyses. Bij deze analyses kunnen, waar relevant ook andere onderzoeksresultaten (vb. uit internationaal vergelijkend prestatieonderzoek) worden gebruikt. Op basis van deze combinatie van onderzoeksresultaten kan een door alle belanghebbenden gedragen actieplan worden opgesteld.

5.2.2 Peilingen inzetten voor interne kwaliteitszorg

Het bijdragen aan de interne kwaliteitszorg van scholen is een secundaire doelstelling van het peilingsonderzoek, al zijn er duidelijk verschillende standpunten over hoe dit best kan gebeuren.

Momenteel gebeurt dit via het schoolfeedbackrapport (zie ook 5.3 over toegang tot het rapport) dat scholen ontvangen na deelname aan een peiling of het afnemen van een paralleltoets. Het is voor scholen niet altijd eenvoudig om de inhoud van dit rapport te gebruiken om het eigen schoolbeleid mee vorm te geven. De rapporten zijn eerder technisch-wetenschappelijk opgesteld en in scholen ontbreken vaak de nodige competenties op het vlak van datageletterdheid om er goed mee aan de slag te kunnen gaan.

Er wordt aangegeven dat de vorm en inhoud van het rapport beter kunnen worden afgestemd op wat nuttig is voor scholen (bv. via een gebruikersonderzoek). Ook kan er meer ondersteuning worden geboden aan scholen om de schoolfeedbackrapporten te kunnen interpreteren. Deze ondersteuning kan deel uitmaken

van de opdracht van het steunpunt, maar moet goed worden afgestemd met andere onderwijspartners, in het bijzonder de pedagogische begeleidingsdiensten.

Aansluitend wordt gesuggereerd om via de peilingen te trachten de interne kwaliteitszorg in scholen op systeemniveau te monitoren en de resultaten ervan mee te nemen in een gemeenschappelijk actieplan dat wordt opgesteld in het kader van de valorisatie van de peilingsresultaten (zie infra). Op deze manier kunnen de primaire en secundaire doelstelling van het peilingsonderzoek op elkaar aansluiten.

5.2.3 Overige doelstellingen

Bij meerdere onderwijspartners leeft de vraag naar feedback op leerlingenniveau. Dergelijke feedback kan zowel dienen voor de interne kwaliteitszorg van een school als extra informatie opleveren voor de school in het kader van de evaluaties van individuele leerlingen.

In het buitengewoon onderwijs leeft de vraag om paralleltoetsen te kunnen afnemen. De resultaten kunnen gebruikt worden bij de afweging of een leerling klaar is voor een overstap naar het gewoon onderwijs. Het betreft hier dus een gebruik op individueel leerlingenniveau waarbij de school beslist voor welke leerlingen dit relevant is.

Ook het informeren van ouders en leerlingen over de leerloopbaan en op het vlak van schoolkeuze kwam aan bod tijdens de evaluaties. Daarbij ziet men de peilingen hooguit als een mogelijke informatiebron, die steeds moet gelezen worden in combinatie met andere bronnen.

Het feit dat een peiling vandaag een momentopname is waarbij enkel een snapshot wordt genomen, zien velen als een zwak punt. Door zich te beperken tot één meetmoment op het einde van een fase in de leerloopbaan, wordt de startpositie van de lerende niet bekeken waardoor met dit instrument geen leerwinst kan gemeten worden. Hierdoor zijn de mogelijkheden om de resultaten van een peiling op formatieve wijze aan te wenden ernstig beperkt. De haalbaarheid van het meten van leerwinst via peilingsonderzoek wordt betwijfeld, wat niet wegneemt dat in het onderwijsveld vragen leven m.bt. het meten van leerwinst.

Het instrumentarium van de peilingen is ook bruikbaar voor andere onderwijsniveaus of andere onderwijssectoren. In het hoger onderwijs werd de nood om peilingsonderzoek daar in te voeren niet gezien, maar in het volwassenenonderwijs ziet men wel mogelijke toepassingen binnen NT2 en tweedekansonderwijs.

5.3 Gebruik en openbaarheid van schoolfeedbackrapporten

Scholen die deelnemen aan een peiling of een paralleltoets, krijgen een schoolfeedbackrapport met daarin de geanonimiseerde resultaten van de deelnemende leerlingen. Over de vraag of deze rapporten ook ter beschikking moeten worden gesteld aan bijvoorbeeld de inspectie in het kader van een schooldoorlichting of aan de pedagogische begeleidingsdienst om zo de begeleiding van de school te versterken, bestaat geen eensgezindheid.

Men is het wel eens dat het eigenaarschap over de schoolfeedbackrapporten in de eerste plaats bij de scholen zelf moet blijven liggen. Er is daarvoor een voldoende vertrouwen in het beleidsvoerend vermogen van de school en in de professionaliteit van de leraren waardoor gesteld wordt dat een school vrij moet kunnen kiezen welke instrumenten toetsen, vragenlijsten het inzet met het oog op haar interne kwaliteitszorg. Hier wordt gewezen op mogelijke ongewenste effecten wanneer éénzelfde instrument zou worden ingezet voor zowel verantwoording als ontwikkeling.

Bovendien zou dit vereisen dat alle scholen deelnemen aan de peilingen of paralleltoetsen en betekent dit dat in de externe evaluatie door de inspectie structureel een element wordt ingebouwd dat bestaat uit een momentopname waarbij geen rekening is gehouden met de startpositie van de leerlingen.

Er wordt gevreesd voor een mogelijke ranking van scholen of zelfs het invoeren van een vorm van centrale toetsing mochten de resultaten van scholen en de schoolfeedbackrapporten breder gedeeld worden in de toekomst, zeker wanneer dit zou gecombineerd worden met een verplichte deelname van scholen aan het peilingsonderzoek. Daaruit volgt dan dat er een risico bestaat van een verenging van het curriculum tot die aspecten die van belang zijn voor de toetsing. Men vreest ook voor een verhoogde stress die zal ontstaan bij leerlingen, leraren en scholen wanneer er een sterke toetscultuur geïnstalleerd wordt. Als mogelijk voordeel van gestandaardiseerde toetsing wordt wel gewezen op de inbreng van een objectief evaluatief element naast de evaluaties door een individuele leerkracht.

5.4 Meer en bredere communicatie

De peilingresultaten en ook de peilingen als instrument zijn vandaag weinig bekend. Dat geldt vandaag zeker voor het brede publiek, maar ook voor sommige directe belanghebbenden en onderwijspartners. Dit heeft een weerslag op deelname aan peilingen en paralleltoetsen en op het gebruik van de resultaten achteraf.

In beide evaluaties wordt dan ook gepleit voor het voeren van meer en bredere communicatie rond de peilingen. Het opzetten van een **communicatieplan** kan niet enkel

meer algemene bekendheid opleveren, maar ook het beter aanwenden van de peilingsresultaten bevorderen (zie ook 5.2.1 en 5.2.2).

5.5 Inhoud, tijdstip en frequentie van het peilingsonderzoek

Het peilingsinstrumentarium dient te worden ontworpen om op een brede wijze het bereik van eindtermen of ontwikkelingsdoelen door leerlingen te meten. Inhoudelijk wordt gepleit om een ruime waaier aan vakinhoudelijke en vakoverschrijdende competenties te peilen. Daarbij wordt een meer systematische peilingenkalender aangeraden.

Naast de leerinhouden wordt ook gevraagd om in de peilingen systematisch ruimte te maken voor het bevragen van welbevinden en motivatie bij leerlingen.

Peilingen moeten plaatsvinden op cruciale punten in de onderwijsloopbaan. Gezien de band met eindtermen is het van belang dat er op die punten eindtermen geformuleerd zijn. Concreet wordt het einde van het lager onderwijs aangeduid als een scharniermoment in de schoolcarrière van elke leerling, evenals het einde van de eerste en de derde graad secundair onderwijs.

5.6 Beheer en aansturing

Het peilingsonderzoek vereist een goede aansturing en voldoende gemandateerde fora voor overleg, adviesverstrekking en beslissingsbevoegdheid.

Doorheen het volledige proces moet het duidelijk zijn hoe de agendasetting en de adviesverstrekking gebeurt en wie, waar en wanneer de nodige beslissingen neemt.

In functie van een voldoende draagvlak en eigenaarschap en een breed gebruik van de peilingsresultaten, zijn bij voorkeur alle stakeholders betrokken bij de aansturing van het peilingsonderzoek.

BIJLAGE 1: PEILINGENKALENDER 2002-2019

Tabel 2 geeft een overzicht van de peilingen die sinds 2002 zijn afgenomen. Hoewel de beheersovereenkomst van het steunpunt loopt tot eind 2017, zijn er reeds peilingen gepland tot 2019. De toetsontwikkeling van een peiling start immers reeds in september van het jaar x-2 voor de afname van de peiling in jaar x.

Tabel 2. Peilingenkalendar

Jaar afname	basisonderwijs	secundair onderwijs	
			graad
2002	wiskunde en begrijpend lezen		
2003			
2004		informatieverwerving en -verwerking	1 ^{ste} graad A-stroom
2005	wereldoriëntatie: natuur met praktische proef onderzoeksvaardigheid en		
2006		natuurwetenschappen - biologie met praktische proef onderzoeksvaardigheden	1 ^{ste} graad A-stroom
2007	Nederlands: lezen (bis) en luisteren	Frans: lezen, luisteren en schrijven	1 ^{ste} graad A-stroom
2008	Frans: lezen, luisteren en schrijven praktische proef spreken	wiskunde	1 ^{ste} graad B-stroom
2009	wiskunde (bis)	wiskunde	1 ^{ste} graad A-stroom
2010	wereldoriëntatie: tijd, ruimte en maatschappij met praktische proef brongebruik	Nederlands: lezen en luisteren met praktische proef spreken	3 ^{de} graad aso, tso, kso
2011		informatieverwerving en -verwerking (bis)	1 ^{ste} graad A-stroom
		wiskunde	2 ^{de} graad aso
2012	informatieverwerving en -verwerking met praktische proef ICT	Frans: luisteren met praktische proef spreken	3 ^{de} graad aso, tso, kso
2013	Nederlands: lezen	PAV	3 ^{de} graad bso

	(tris) en luisteren (bis)		
2014		wiskunde wiskunde	3 ^{de} graad aso 3 ^{de} graad tso, kso
2015	wereldoriëntatie: natuur (bis) + techniek	Natuurwetenschappen (bis)	1 ^{ste} graad A-stroom
2016	wiskunde (tris)	Burgerschap	3 ^{de} graad aso, tso, kso, bso
2017	Frans (bis)	Techniek	1 ^{ste} graad A-stroom
2018	Nederlands: lezen (4) en luisteren (tris) praktische proef: schrijven	wiskunde (bis)	1 ^{ste} graad A-stroom
2019	wereldoriëntatie: mens en maatschappij (bis)	wiskunde (bis)	1 ^{ste} graad B-stroom

BIJLAGE 2: METHODE VAN DE EX ANTE EVALUATIE

Om een antwoord te krijgen op de onderzoeksvragen in de ex ante evaluatie werd er geopteerd om de meningen, standpunten en perspectieven van verschillende stakeholders werd er geopteerd om:

- in een eerste fase de mogelijke doelstellingen in kaart brengen van de peilingen bij verschillende stakeholders- of gebruikersgroepen alsook de wenselijke inhoud en aanpak. Dit gebeurde aan de hand van doelgroepgebonden focusgroepen (zie lager).
- in een tweede fase feedback te krijgen over concrete scenario's waarin er rekening is gehouden met de al dan niet noodzakelijke verdere uitbouw van het huidige systeem die de geïdentificeerde doelen realiseren. De feedback werd verkregen via het organiseren van een gezamenlijke terugkomdag voor alle deelnemers uit de focusgroepen.

In dit hoofdstuk bespreken we eerst hoe we de deelnemers geselecteerd en gecontacteerd hebben. Vervolgens bespreken we het verloop van de focusgroepen en de verwerking van de resultaten. In een derde deel gaan we in op de methode die gebruikt werd bij de ontwikkeling van de scenario's. Tenslotte bespreken we het verloop van de terugkomdag en de verwerking van de resultaten.

1. Deelnemers

1.1. Selectie van de deelnemers

Er werden acht focusgroepen georganiseerd:

- (1) leerkrachten,
- (2) schoolleiders,
- (3) medewerkers uit het hoger onderwijs;
- (4) medewerkers van centra voor volwassenenonderwijs en centra voor basiseducatie;
- (5) pedagogische begeleiders;
- (6) ouders en leerlingen;
- (7) experts kwaliteitszorg en inspectie en
- (8) beleidsmedewerkers van het Ministerie van Onderwijs en Vorming.

We streefden naar een gemiddelde grootte van 15 à 20 deelnemers per focusgroep. Het doel was niet zozeer om een representatieve vertegenwoordiging van de doelgroep binnen elke focusgroep te realiseren maar om een zo divers mogelijke samenstelling (deelnemers met en zonder voorkennis van de peilingen) te verkrijgen. Om dit te bereiken, gebruikten we verschillende strategieën afhankelijk van de doelgroep.

Voor de focusgroep '**pedagogische begeleiders**' schreven we alle onderwijskoepels (Katholiek Onderwijs Vlaanderen, GO!, OVSG, POV en OKO) en pedagogische begeleidingsdiensten (PBDKO, PB GO!, PB OVSG, PB Provinciaal, FOPEM, VONAC vzw, Steinerscholen - OKO PBD en IPCO) aan met de vraag om - indien van toepassing - om één vertegenwoordiger uit het basisonderwijs en één vanuit het secundair onderwijs af te vaardigen. Ook de koepels van de CLB's werden aangesproken (Vrije CLB Koepel, GO! Koepel, OVSG-dienst CLB, CLB - POC OGO).

Voor de focusgroep '**ouders en leerlingen**' werd er via de Koning Boudewijnstichting een oproep gedaan om een vijftal deelnemers van het ouderpanel uit te nodigen voor de focusgroep. In kader van het eindtermendebat kregen deze ouders al een toelichting over eindtermen en peilingen en hebben ook een aanbeveling geformuleerd rond het toetsen van leerlingen. Zij hebben dus al een goede startpositie om mee na te denken over de toekomst van de peilingen. Daarnaast spraken we ook de ouderverenigingen (KOOGO, VCOV en GO! Ouders vzw) aan met de vraag om telkens twee ouders af te vaardigen, één voor basisonderwijs en één voor secundair onderwijs of één ouder met een kind in het buitengewoon onderwijs. Voor de scholieren vroegen we aan de Vlaamse Scholierenkoepel om een tiental scholieren af te vaardigen waarbij er gevraagd werd naar een mix tussen jongens en meisjes, verschillende onderwijsvormen en leeftijd. Er werd daarbij ook aangegeven dat het interessant zou zijn om leerlingen die deelnamen aan het eindtermendebat te betrekken.

66

Voor de focusgroep '**hoger onderwijs**' werden de VLIR en de VLHORA aangeschreven. Aan de VLIR werd er gevraagd om vijf vertegenwoordigers vanuit de universiteiten en aan de VLHORA tien vertegenwoordigers vanuit de hogescholen af te vaardigen. Met betrekking tot het profiel stelden we voor om vertegenwoordigers vanuit de verschillende lerarenopleiding (GLO en SLO) zowel uit het basisonderwijs als het secundair onderwijs af te vaardigen en daarnaast vertegenwoordigers die voornamelijk bezig zijn met de startcompetenties van studenten zoals bv. studietrajectbegeleiders. Deze vijftien vertegenwoordigers werden aangevuld met een vijftal deelnemers die we selecteerden uit een lijst van personen uit het hoger onderwijs die recent deelnamen aan de conferenties en colloquia van de peilingen en/of meewerkten aan de peilingen (via bv cesuurbepalingen). Deze selectie gebeurt op basis van de deelname aan de meest recente conferenties, colloquia en werkseminaries van de peilingen en/of medewerking aan de meest recente peilingen. Op die manier ontstond er een mix van deelnemers uit het hoger onderwijs met en zonder voorkennis van de peilingen.

Voor de focusgroep '**experten en inspectie**' schreven we experts aan op het gebied van kwaliteitszorg. Daarnaast werd er aan de inspectie gevraagd om vertegenwoordigers af te vaardigen.

Wat betreft de focusgroep **'beleidsmedewerkers'** vroegen we naar vertegenwoordigers uit de verschillende afdelingen van het departement Onderwijs & Vorming (afdeling SBO, HB, SLB, BOP, BDK, HV), een vertegenwoordiger uit de VLOR, een vertegenwoordiger vanuit AgODi en één uit Ahovoks.

Voor de focusgroep **'volwassenonderwijs (CVO en CBE)'** werden er op toevallige wijze tien centra gekozen uit alle Vlaamse centra voor volwassenenonderwijs en vijf centra uit alle centra voor basiseducatie die een aanbod NT2 hebben. We nodigden daarbij steeds de directie uit (en indien die niet aanwezig kon zijn één van de coördinatoren) en één van de leraren. Om een zo divers mogelijk publiek aan te trekken, werd er bij de steekproeftrekking rekening gehouden met het net van de centra en de grootte. Er werden ook steeds twee reservecentra voorzien om te contacteren in het geval het eerst aangeschreven centrum niet wenste deel te nemen.

Voor de focusgroepen **'leerkrachten'** en **'directies'** schreven we 20 deelnemers aan per focusgroep waarvan ongeveer de helft (zeven) van de deelnemers reeds in contact waren gekomen met de peilingen via deelname aan colloquia, werkseminaries of conferenties of die reeds meewerkten aan het peilingsonderzoek. De andere helft (acht) van de leerkrachten en directies werden geselecteerd op toevallige wijze uit onze databanken. De steekproef van scholen uit de databank gebeurde op basis van de strata net, schoolgrootte en ligging (centrumsteden versus niet-centrumsteden) in het basisonderwijs en net, schoolgrootte, onderwijsvormen en ligging (centrumsteden versus niet-centrumsteden) in het secundair onderwijs. Daarnaast werd er ook een selectie gemaakt van vijf scholen uit het buitengewoon basisonderwijs waarbij we ons beperkten tot scholen met een aanbod type 1, 3 of 8. Voor de selectie van de vijf scholen uit het buitengewoon secundair onderwijs beperkten we ons tot opleidingsvormen 3 en 4. Er werden steeds twee reservescholen voorzien om te contacteren in het geval de eerst aangeschreven school niet wenste deel te nemen. Voor het basisonderwijs werd er gevraagd om indien mogelijk leerkrachten uit het zesde leerjaar af te vaardigen. Indien de directie verhinderd was, werd er gevraagd of de zorgcoördinator in de plaats kon komen. Voor het secundair onderwijs werd gevraagd om leerkrachten uit verschillende graden te vertegenwoordigen en om eventueel een graadcoördinator in de plaats van de directie te laten deelnemen indien de directie verhinderd was.

1.2. Contacteren van de deelnemers

De geselecteerde deelnemers ontvingen in september- oktober een mail met een uitnodiging tot deelname. Na een eerste introductiemail ontvangen te hebben, werden alle geselecteerde scholen, centra voor volwassenenonderwijs en basiseducatie opgebeld door medewerkers van het departement Onderwijs en Vorming met de vraag tot medewerking. Een week na het versturen van de eerste introductiemail ontvingen de

overige geselecteerden een herinneringsmail. Tenslotte kreeg men een laatste herinneringsmaal twee dagen voordat de focusgroep plaatsvond met praktische informatie.

In Tabel 3 wordt er een overzicht gegeven van de data waarop de focusgroepen plaatsgrepen en het aantal deelnemers per focusgroep.

Tabel 3. Overzicht van de data en het aantal deelnemers per focusgroep

Focusgroep	Datum	Aantal deelnemers aanwezig
Experten en inspectie	14/10/2016	10
Pedagogische begeleidingsdiensten	14/10/2016	7
Hoger Onderwijs	21/10/2016	15
Volwassenenonderwijs	21/10/2016	13
Leerkrachten	25/10/2016	16
Directies	25/10/2016	15
Beleidsmedewerkers	26/10/2016	7
Ouders en leerlingen	9/11/2016	9
Totaal		92

Alle deelnemers ontvingen het verslag van de focusgroep waaraan men deelnam en een uitnodiging voor de 'terugkomdag peilingen' op 9 december 2016. Op de terugkomdag waren er 63 deelnemers met een goede vertegenwoordiging uit alle focusgroepen (4 beleidsmedewerkers, 8 volwassenenonderwijs, 8 directies, 11 leerkrachten, 6 experts en inspectie, 12 hoger onderwijs, 8 ouders en leerlingen, 6 pedagogische begeleiders). Enkel de scholieren konden niet deelnemen aan de terugkomdag wegens de start van de examens. Een vertegenwoordiger van de Vlaamse Scholierenkoepel was aanwezig tijdens de terugkomdag om de stem van de studenten te vertegenwoordigen.

2. Focusgroepen

2.1. Verloop van de focusgroepen

De acht focusgroepen werden georganiseerd in de periode oktober – november 2016 te Brussel. In elk van deze sessies werd diepgaand ingegaan op de doelstellingen van de peilingen en de wijze waarop ze gerealiseerd (kunnen) worden.

Er werd gekozen samen te werken met externe procesbegeleiders (Levuur cvba) om vanuit een extern perspectief de gesprekken in goede banen te leiden. In onderlinge afstemming werd een draaiboek voor de gesprekken opgemaakt, vergezeld van een

mindmap waarin de te bespreken items overzichtelijk waren samengevat. In overleg en samenspraak met Levuur cvba werden de volgende vragen voorgelegd aan de deelnemers:

- In welke mate slagen de peilingen er vandaag in om hun doelstellingen te bereiken?
- Zijn er andere doelstellingen van de peilingen wenselijk of nodig?
- Hoe en door wie dienen de peilingen gebruikt te worden op macro-, micro- en mesoniveau
- Wat moeten peilingen meten?
- Wanneer of hoe frequent moeten peilingen gemeten worden?

Om deze vragen te beantwoorden werd in de focusgroepen gewerkt in twee rondes. In eerste ronde werd ingegaan op de mate waarin de huidige doelstellingen van de peilingen gerealiseerd worden en welke mogelijke alternatieve doelstellingen geformuleerd kunnen worden. In de tweede ronde werd ingegaan op de vraag naar de wenselijke inhoud en aanpak van de peilingen.

De methodiek die door de procesbegeleider voor deze groepsgesprekken was uitgewerkt verzekerde dat elke deelnemer aan het gesprek een gelijke stem kreeg en dat alle punten die de deelnemers op voorhand hadden voorbereid om in te brengen in het groepsgesprek aan bod kwamen.

De focusgroepen verliepen op een gelijkaardige manier. Er waren naast de externe procesbegeleider telkens twee verslaggevers aanwezig. Dit leidde tot een gezamenlijk verslag in een vast sjabloon, waarin alles wat aangehaald werd in thematische clusters werd ondergebracht. Er vond in deze fase geen correctie van eventuele feitelijke onjuistheden plaats en er was ook geen selectie op basis van wat wel of niet onder de scope van de peilingen valt. De verslagen werden ter kennisgeving aan de focusgroepsleden bezorgd.

2.2. Verwerking van de resultaten van de focusgroepen

We structureerden de resultaten van de focusgroepgesprekken aan de hand van de vragen die tijdens de gesprekken aan bod kwamen. Bij de beschrijving van de standpunten en meningen zagen we erop toe dat zowel de voor- en tegenstanders aan bod kwamen. Gezien de verschillende perspectieven van de verschillende gebruikersgroepen werd steeds aangeduid uit welke focusgroep een bepaald standpunt werd ingenomen.

We geven de meningen en standpunten van de deelnemers zo accuraat en beschrijvend mogelijk mee zonder daarbij eigen interpretaties toe te voegen.

3. Ontwikkeling van de scenario's

Op basis van de verslagen van de focusgroepen werden er vier scenario's met betrekking tot de peilingen uitgewerkt.

In een eerste stap werden de mogelijke doelstellingen van de peilingen geïdentificeerd die tijdens de focusgroepen aan bod kwamen, namelijk:

- de kwaliteit van het Vlaams onderwijs bewaken en versterken;
- de kwaliteit van Vlaamse scholen bewaken, versterken en ondersteunen;
- de gelijkwaardigheid van getuigschriften en diploma's verhogen;
- het versterken van de geïnformeerde schoolkeuze van ouders;
- de afstemming tussen de interne kwaliteitszorg en de externe kwaliteitszorg;
- de lerende ondersteunen;

In een tweede stap werden er een aantal kenmerken van de peilingen opgesteld waarop de deelnemers verschillende posities innamen tijdens de focusgroepgesprekken. Deze zijn:

- het verantwoordings- versus het ontwikkelingsperspectief;
- verplichte deelname van scholen aan de peilingen versus vrijwillige deelname;
- openbaarheid van de resultaten versus geen openbaarheid;
- een focus van de peilingen op het Vlaams niveau, schoolniveau of leerlingenniveau;
- het meten van resultaten via een momentopname versus het meten van een proces of ontwikkeling;
- het meten van (een beperkt aantal) competenties versus breed evalueren;
- het geven van feedback op individueel niveau versus op een geaggregeerd niveau (Vlaams of schoolniveau);

Een combinatie van deze twee bovenstaande elementen (doelstellingen en kenmerken) leverde het raamwerk op waarbinnen de scenario's ontwikkeld werden. Omdat het in vrijwel alle discussies een centrale plaats in nam, gebruikten we in de verdere uitwerking van de scenario's het spanningsveld tussen verantwoordings- en ontwikkelingsperspectief als leidend principe. Een belangrijk verschil tussen de vier scenario's lag in de mate waarin de nadruk lag op een verantwoordings- dan wel een ontwikkelingsperspectief.

Dit spanningsveld 'verantwoording versus ontwikkeling' werd ook in de literatuurstudie en de Delphi-studie, in het kader van het OBPWO-onderzoek naar leerwinstmonitoring, als één van de belangrijkste bouwstenen geïdentificeerd voor het opzetten en

vormgeven van leer(winst)monitoringssystemen.

Een laatste belangrijk principe dat werd gehanteerd bij het opstellen van de scenario's is om zoveel mogelijk ideeën met betrekking tot mogelijke bijkomende doelstellingen en mogelijke nieuwe manieren van gebruik, die door de deelnemers geuit werden in de focusgroepen, te verwerken in de scenario's.

4. Terugkomdag

4.1. Verloop van de terugkomdag

Na de ontwikkeling van de scenario's (zie hoger) vond een terugkomdag plaats op 9 december 2016 te Brussel. Voor deze terugkomdag werden alle deelnemers aan de eerder genoemde focusgroepen uitgenodigd.

De ontwikkelde scenario's werden ter feedback voorgelegd aan groepen met een gemengde samenstelling om de verschillende standpunten en perspectieven zoveel mogelijk met elkaar te kunnen confronteren. Zij dienden niet gezien te worden als realistische scenario's maar als didactische instrumenten waarin sommige standpunten bewust werden scherp gesteld om de discussie te voeden en aan te wakkeren.

De terugkomdag vond in meerdere rondes plaats. Na de voorstelling van de verschillende scenario's, kozen de deelnemers in een eerste ronde een scenario uit dat ze verder wilden uitwerken. De groepen per scenario bestonden uit actoren uit verschillende focusgroepen. Op deze wijze kon in het gesprek vanuit verschillende invalshoeken/actoren naar het scenario gekeken worden. Aan de hand van een template werden de verschillende kenmerken en randvoorwaarden voor dat betreffende scenario bediscussieerd. In de tweede ronde konden deelnemers feedback geven op andere scenario's. De derde ronde nodigde de deelnemers uit om de analyse van hun eerst gekozen scenario te vervolledigen. De terugkomdag werd afgerond met een vierde ronde, waar de essentiële elementen (kenmerken of randvoorwaarden) m.b.t. de verschillende scenario's plenair teruggekoppeld worden.

4.2. Verwerking van de input van de terugkomdag

De feedback van de deelnemers op elk scenario werd gestructureerd per scenario aan de hand van de thema's die aan bod kwamen tijdens de discussies. Daarbij kan er een onderscheid gemaakt worden tussen de feedback die gegeven werd enerzijds op de doelstellingen van elk scenario en anderzijds op het gebruik van de peilingen. Bij de beschrijving van de standpunten en meningen zagen we erop toe dat zowel de voor- en tegenstanders aan bod kwamen. Waar mogelijk werden steeds

de argumenten opgesomd. Gezien de verschillende perspectieven van de verschillende gebruikersgroepen werd indien mogelijk aangeduid welke 'soort' deelnemer feedback gaf.

Ook hier werd ervoor gekozen om de meningen en standpunten van de deelnemers zo accuraat en beschrijvend mogelijk mee zonder daarbij eigen interpretaties toe te voegen.

Departement Onderwijs en Vorming
Koning Albert II laan 15
1210 Brussel
<http://www.ond.vlaanderen.be/>