



Vlaams
Parlement

ingediend op **251** (2019-2020) – Nr. 1
24 maart 2020 (2019-2020)

Verslag van de hoorzitting

namens de Commissie voor Onderwijs
uitgebracht door Loes Vandromme en Jan Laeremans

over de gevalideerde toetsen
voor het basisonderwijs

Samenstelling van de Commissie voor Onderwijs:

Voorzitter: Karolien Grosemans.

Vaste leden:

Arnout Coel, Koen Daniëls, Karolien Grosemans, Kathleen Krekels, Annabel Tavernier;
Roosmarijn Beckers, Jan Laeremans, Kristof Slagmulder;
Jo Brouns, Loes Vandromme;
Jean-Jacques De Gucht, Sihame El Kaouakibi;
Stijn Bex, Elisabeth Meuleman;
Hannelore Goeman.

Plaatsvervangers:

Marius Meremans, Axel Ronse, Nadia Sminate, Kris Van Dijck, Katja Verheyen;
Leo Pieters, Els Sterckx, Sam Van Rooy;
Katrien Schryvers, Brecht Warnez;
Els Ampe, Steven Coenegrachts;
Johan Danen, Gustaaf Pelckmans;
Steve Vandenberghe.

Toegevoegde leden:

Kim De Witte.

INHOUD

I.	Uiteenzetting door OVSG	4
1.	Waarom een OVSG-toets?.....	4
2.	Wat is de OVSG-toets?	4
3.	Hoe aan de slag met de OVSG-toets?	5
4.	Hoe versterken we de OVSG-toets?.....	6
II.	Uiteenzetting door Katholiek Onderwijs Vlaanderen.....	7
1.	Doel en inhoud IDP	7
1.1.	Situering in de evaluatiebox.....	7
1.2.	Breed aanbod met scherpe focus	8
1.3.	Resultaten	8
1.4.	Evolutie Nederlands – begrijpend lezen	9
2.	Kwaliteitsontwikkeling als organisatie	9
3.	Kwaliteitsontwikkeling op school	10
III.	Uiteenzetting door het Steunpunt voor Toetsontwikkeling en Peilingen	11
1.	Peilingen.....	11
2.	Doelstellingen.....	11
2.1.	Systeemniveau.....	11
2.2.	Schoolniveau.....	11
3.	Onderzoeklogica	12
3.1.	Leerprestaties leerlingen.....	12
3.2.	Schoolprestaties	14
4.	Schoolfeedback bij paralleltoetsen.....	15
5.	Gebruik	15
6.	Besluit	15
IV.	Bespreking.....	16
1.	Tussenkomen van de leden.....	16
2.	Antwoorden.....	17
3.	Bijkomende vragen en antwoorden	22
	Gebruikte afkortingen	26
	Bijlagen: zie de dossierpagina van dit document op www.vlaamsparlement.be	

Op 20 februari 2020 organiseerde de Commissie voor Onderwijs een hoorzitting over de gevalideerde toetsen voor het basisonderwijs met:

Onderwijsvereniging van Steden en Gemeenten:

- Patriek Delbaere, algemeen directeur;
- Steven De Laet, pedagogisch adviseur, coördinator Basisonderwijs regio Oost en projectleider OVSG-toets;
- Elias Hemelsoet, coördinator Studie- en Ontwikkelingsdienst;

Katholiek Onderwijs Vlaanderen:

- Lieven Boeve, directeur-generaal;
- Hilde Hendrickx, schoolbegeleider;
- Anne Verhoeven, teamverantwoordelijke Dienst Curriculum en Vorming;

Steunpunt Toetsontwikkeling en Peilingen:

- prof. dr. Rianne Janssen, hoofd Eenheid Onderwijseffectiviteit en -evaluatie, Faculteit Psychologie en Pedagogische Wetenschappen, KU Leuven;
- dr. Koen Aesaert, projectleider Basisonderwijs en coördinator Data-analyse;
- dr. Mieke Heyvaert, teamleider.

De gebruikte presentaties kunnen geraadpleegd worden op de [dossierpagina](#) van dit document op www.vlaamsparlement.be.

I. Uiteenzetting door OVSG

Patriek Delbaere leidt de presentatie van de OVSG-toets in. De vereniging heeft begrepen uit het regeerakkoord dat toetsen niet alleen een prominente rol zullen spelen in het beleid maar dat er ook een evolutie is. Het instrument zal immers meer gericht worden op leerwinst.

De onderwijsvereniging wijdde samen met haar Nederlandse zusterorganisatie reeds een expertdag aan het onderwerp, tijdens dewelke beide praktijken werden vergeleken. Dat werk wordt op 13 maart voortgezet op een congres met de schoolbesturen over onderwijskwaliteit en de plaats daarin van de OVSG-toets, die inmiddels 25 jaar bestaat.

1. Waarom een OVSG-toets?

Elias Hemelsoet vervolgt met een schets van de historiek. De OVSG-toets is voortgekomen uit de drang om de onderwijskwaliteit te versterken. Die beoogt het verhogen van de leerresultaten, met andere woorden de leerwinst. Op dit moment doet de toets weliswaar nog niet aan systematische leerwinstmonitoring, erkent hij, maar de verhoging ervan vormt wel een impliciete doelstelling.

Om de complexiteit van het toetsen te illustreren, gebruikt hij het beeld van Rubiks kubus. Het is niet zo eenvoudig dat de invoering van (centrale) toetsen vanzelf leidt tot meer onderwijskwaliteit. Zij kunnen ertoe bijdragen maar dan vanuit een breed perspectief. Er moet immers op verschillende acties tegelijk ingezet worden. Toetsen vormen een waardevolle hefboom indien tegelijk werk wordt gemaakt van de professionalisering van leerkrachten, de begeleiding, de ondersteuning enzovoort. Hij vat samen dat toetsen een belangrijk instrument voor interne kwaliteitszorg vormen naast vele andere.

2. Wat is de OVSG-toets?

Steven De Laet licht de concrete werking toe. De OVSG-toets is gestandaardiseerd en wetenschappelijk gevalideerd. Hij is opgenomen in de toolkit 'breed evalueren'

op het einde van het zesde leerjaar. Tegelijk is hij een instrument voor de zelfevaluatie van scholen. Hij is ook relevant voor de praktijk, want vanaf het begin was de vrijwillige deelname hoog en zij werd steeds groter. De recente regelgeving heeft daar nog toe bijgedragen. Verder is hij nog een bron van materiaal voor leerlingenbegeleiding en voor de evaluatie op leerlingenniveau, naast de schooleigen gegevens. Men mag er echter geen high-stakes tests van maken die zouden oriënteren voor het secundair onderwijs, waarschuwt hij.

De spreker geeft een voorbeeld van een vraag uit de toets, die moet opgelost worden met behulp van een bronnenboekje. In de handleiding voor de leraren staat de vermelding van de bijhorende leerplandoelen, de juiste oplossing, de puntenverdeling en een verwijzing naar de getoetste eindtermen.

In een toets over alle leergebieden en domeinen is het niet mogelijk om alle eindtermen te toetsen, al krijgt de auteur steeds de opdracht daarin zo breed mogelijk te gaan, schetst de spreker. Voor Frans bijvoorbeeld komen, gespreid over de jaren, alle doelen aan bod, uitgezonderd twee of drie eindtermen die moeilijk in het toetsenformat te vatten zijn.

3. Hoe aan de slag met de OVSG-toets?

De toets verloopt snel. De leerlingen maken hem op maandag, dinsdag en woensdag in zes blokken. De leraren hebben donderdag en vrijdag om de resultaten in te geven. De daaropvolgende maandag wordt de interactieve website met de resultaten vrijgegeven. De tabel bevat in de rechterkolom de resultaten van de school, met vermelding van het aantal leerlingen en de SES-categorie. In de tabel staan ook de resultaten van de voorbije jaren en de cijfers van het Vlaamse gemiddelde, van de scholen uit de gemeente en van de SES-categorie. Scholen beschikken dus al na drie dagen over kwantitatieve feedback.

Van de vele andere mogelijkheden die de interactieve omgeving nog biedt, geeft de spreker als voorbeeld de mogelijkheid om in een grafiek de evolutie van het resultaat van de school, die van de SES-groep en die van het Vlaamse gemiddelde over de jaren heen weer te geven.

Daarnaast wordt even snel gezorgd voor kwalitatieve analyse-instrumenten. Per vraag krijgt men de juiste antwoorden en het gemiddelde op Vlaams niveau. De leerplandoelen en de klassen die daar volgens het leerplan aan horen te werken, worden ook weergegeven. Doel is in de scholen een gesprek op gang te brengen, omdat kwaliteit niet alleen een verantwoordelijkheid is van de leerkracht van het zesde leerjaar. Door dergelijke visualisering verhoogt men de betrokkenheid, stelt de spreker. Op verzoek worden scholen ook verder begeleid. In de zomervakantie zet OVSG haar werk verder en bereidt zij het volgende schooljaar voor. De vereniging bekijkt het mogelijke nascholingsaanbod, zet netwerken op zoals voor data-letterdheid en inhoudelijk-didactische verdieping, en begeleidt scholen op maat, naast het bieden van ikz-instrumenten en didactische readers.

Ook een groot deel van het GO! neemt deel aan de OVSG-toetsen. Als motief geeft men aan dat het een brede alletalententoets is, die de leraren inspireert en elk jaar volledig nieuw is. Het deelnamepercentage bij OVSG- en GO!-scholen bedraagt respectievelijk 99 en 94 procent. De pedagogische begeleidingsdienst van het GO! krijgt ook de kans om feedback te geven op de proefversies, en er wordt voor gezorgd dat zij de concordantie met hun leerplan kunnen maken in het licht van de ondersteuning van de scholen.

4. Hoe versterken we de OVSG-toets?

Wat de kwaliteitszorg voor de toets zelf betreft, vermeldt de spreker het valideeronderzoek van onder meer professor Janssen. De gouden tips daarin werden ter harte genomen bij de verdere verbetering van de technische kwaliteit en van de toepassing als instrument voor kwaliteitszorg in de school. Noodzakelijk daartoe was de digitalisering. Naast de genoemde feedback en de reflecties van leraren en leerlingen, worden ook statistische principes ingezet om de interne consistentie en de discriminerende kracht te bekijken en op een hoger niveau te brengen.

Verder in dit verband wijst Steven De Laet op de schoolfeedbackrapporten. De interactieve omgeving biedt veel analysemogelijkheden, maar er blijft handelingsverlegenheid op het vlak van datageletterdheid bestaan. Om die te overwinnen, werden op basis van de digitale gegevens voor het eerst geïndividualiseerde rapporten gemaakt, waarmee de pedagogische adviseurs thans naar de scholen stappen.

Nog een recente verandering is de wetenschappelijk onderbouwde inzet van 'rubrics' bij de beoordeling van competenties. Voordeel is dat ze nu allemaal op gelijkaardige manier door de scholen in beeld worden gebracht. De doelstellingen worden weergegeven, uitgedrukt op verschillende beheersingsniveaus. Daarnaast wordt het op grond van de eindtermen verwachte niveau beschreven, maar de spreker benadrukt dat dit geen glazen plafond is en dat ook excelleren aan bod komt. Deze nieuwe aanpak heeft geen noemenswaardige vragen opgeroepen, stelt hij vast.

Steven De Laet legt uit dat 'boven het verwachte niveau' dezelfde score krijgt als 'het verwachte niveau' (twee en niet bijvoorbeeld drie punten), omdat de eindterm in beide gevallen wordt beheerst. Het excelleren wordt alleen verbaal uitgewerkt.

In het kader van een sterker aanklappend beleid wordt extra ingezet op de rol van de schoolbesturen, die voortaan desgewenst rechtstreeks toegang krijgen tot de toetsresultaten en geprofessionaliseerd worden (zoals op de genoemde congresdag van 13 maart). De spreker attendeert op het onderzoek van het Rekenhof uit 2019, dat uitwijst hoezeer het actieve partnerschap van schoolbesturen een effect heeft op de werking. Verder werd een proeftuin opgezet waarin ook de pedagogische begeleiders toegang hebben tot de resultaten, met algemene toestemming van de scholen. Hoe die laatste in de praktijk verloopt en gemotiveerd wordt, zal worden geanalyseerd.

Elias Hemelsoet onderstreept het belang van het brede karakter van de toets. Procesmatig betekent dit een geïntegreerde aanpak, want toetsen staan niet op zich. Dat begint bij de ontwikkeling en loopt door tot de begeleiding na de feedback, wat hij het voordeur- tot achterdeurprincipe noemt. Op curriculair vlak betekent dit dat men zich richt op de hele breedte van de eindtermen, om het zogeheten 'teaching to the test' te voorkomen. Tot slot in dit verband wordt ook resoluut de kaart getrokken van brede evaluatie, want het competentiegericht formuleren van doelstellingen vereist aangepaste evaluatievormen. Daartoe wordt gewerkt met praktijktoetsen en geïntegreerde proeven.

Tweede troef is de focus op onderwijskwaliteit, die een gedeelde verantwoordelijkheid vormt. Hierbij treedt OVSG op als ondersteuner, maar vindt de realisatie op de eerste plaats op de werkvloer plaats. Daartoe betreft OVSG het hele spectrum van actoren: leerkrachten, klassenraad en bestuurders. Het professionaliserings-effect op de begeleiders behandelt hij straks.

De toets is op maat van het veld. De actoren geven aan dat het een bruikbaar en waardevol instrument is, waarbij ze onderstrepen dat het in de taal van de school

gesteld is. Verder heeft de toets een groot bereik en draagvlak. Jaarlijks wordt hij afgelegd door ongeveer 25.000 leerlingen in 913 scholen, wat neerkomt op vrijwel het gehele officiële onderwijs.

Hij vermeldt ook nogmaals dat het een snelle toets is, waardoor de school meteen met de resultaten aan de slag kan, al kosten de individuele schoolfeedbackrapporten wat meer tijd. Niet onbelangrijk voor het beleid is ook zijn betaalbaarheid, voor OVSG dankzij de integratie in begeleiding en koepelwerking, voor de scholen dankzij de lage prijs per leerling (ongeveer 3,5 euro).

Verder is de toets een stimulans voor innovatie. Deelname leidt tot andere praktijken op de werkvloer, zowel in de evaluatie als de didactiek. De digitalisering stimuleert scholen om verder na te denken over de digitale evaluatiepraktijk. Geïntegreerde proeven zetten aan tot reflectie over het eigen competentiegericht toetsen en evalueren. Tegelijk wordt een achterdeur geopend voor een snellere implementatie van een nieuw curriculum.

Tot slot gaat hij zoals aangekondigd in op het professionaliseringseffect op de pedagogische begeleiders. Hun betrokkenheid bij de ontwikkeling leidt tot een verandering in hun omgang met evaluatiebeleid. Hun aandacht voor datagericht werken, wordt aangescherpt, ook op andere domeinen dan alleen toetsen. De versterking van de datageletterdheid en het intern professionaliseringseffect zijn heel belangrijk, besluit hij.

Wat de toekomst betreft, vermeldt hij opnieuw de regeerverklaring en de beleidsnota over de invoering van gestandaardiseerde, gevalideerde en genormeerde toetsen. In dat verband pleit hij ervoor de expertise in het veld te benutten. Elk van de bestaande toetsen heeft zijn kracht, troeven en beperkingen. Complementariteit opzoeken is dan ook belangrijk. OVSG biedt zich in elk geval aan als gesprekspartner bij de uitrol. De spreker benadrukt daarbij hoe belangrijk een duidelijke definitie van leerwinst is.

Tot slot attendeert hij nog op de toetsenwijzer, met toetsen die in samenwerking met lokale actoren ontwikkeld zijn voor het eerste tot en met zesde leerjaar. De spreker ziet mogelijkheden in de professionalisering daarvan, in samenhang met de OVSG-toets, in het perspectief van leerwinstmonitoring.

II. Uiteenzetting door Katholiek Onderwijs Vlaanderen

Lieven Boeve leidt de voorstelling van de toetsen van het katholiek onderwijs in, die inmiddels 52 jaar bestaan.

1. Doel en inhoud IDP

Anne Verhoeven behandelt doel en inhoud van de IDP-toetsen. Het doel is dubbel: de interne kwaliteitsontwikkeling van een school en de interne kwaliteitszorg van Katholiek Onderwijs Vlaanderen. Het eerstgenoemde gebeurt door de interpretatie van de eigen schoolresultaten in vergelijking met andere scholen met een vergelijkbare leerlingenpopulatie en het bekijken van leerwinst op schoolniveau, het laatstgenoemde door het interpreteren van resultaten voor de netwerkorganisatie en het bekijken van de leerwinst op systeemniveau, naast het bewaken van de kwaliteit.

1.1. Situering in de evaluatiebox

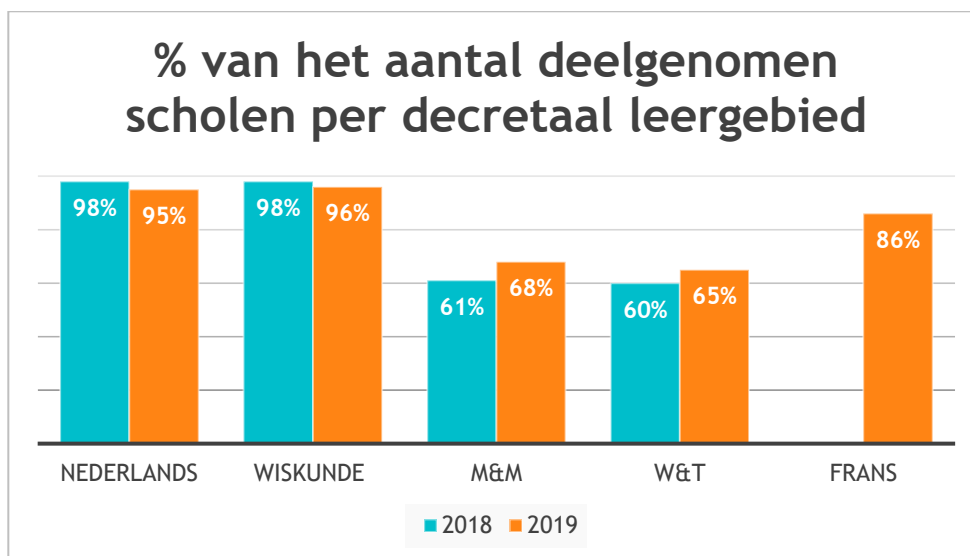
De gevalideerde toetsen maken deel uit van een evaluatiebox. In die doos zitten veel instrumenten voor een brede evaluatie. Zij bevat verwijzingen naar

gestandaardiseerde proeven zoals IDP4 voor het vierde leerjaar, SALTO en de G-toets, de geïntegreerde of praktische proeven, en tot slot de gevalideerde toetsen, die in uitvoering van het decreet basisonderwijs van 25 februari 1997 apart werden gehouden.

Vorig schooljaar hebben 37.150 leerlingen deelgenomen aan de gevalideerde toetsen, voor een groot en toenemend deel online in de plaats van op papier.

1.2. Breed aanbod met scherpe focus

Met een breed aanbod probeert men op alle decretale leergebieden in te zetten. Aanvankelijk ging het om Nederlands en wiskunde, maar sinds vorig jaar tevens om Frans. Mens en maatschappij, en wetenschappen en techniek zaten er al eerder in. Dit jaar komt er muzische en volgend jaar lichamelijke opvoeding bij. Daarnaast zijn er IDP's voor het vierde leerjaar en praktische proeven. De spreker onderstreept daarbij de scherpe focus van het brede aanbod.



In bovenstaande grafiek over het aantal scholen per leergebied merkt men dat er – terecht – veel aandacht gaat naar Nederlands en wiskunde, maar ook de nieuwe toets Frans heeft veel scholen gecharmeerd, wat heeft geleid tot een verviervoudiging van het aantal vragen om begeleiding. De scherpe focus brengt telkens één stuk van het leergebied in kaart, behalve voor begrijpend lezen, waarvoor die wisseling de afgelopen jaren niet plaatsvond. De spreker merkt daarbij op dat men niet altijd voor de meest voor de hand liggende leerinhouden kiest. In muzische vorming dit schooljaar is dat bijvoorbeeld dans, omdat de Onderwijsspiegel aangeeft dat vooral gefocust wordt op beeld en muziek.

1.3. Resultaten

Vervolgens gaat Anne Verhoeven in op de evolutie van de resultaten voor Nederlands en meten en metend rekenen over meerdere jaren. Omdat de inhoud telkens licht anders is, kan men echter geen lijn trekken.



Resultaten

Groep	Betekenis	%
paarse groep	Deze scholen hebben een OKI-waarde < 0,5.	57,4
oranje groep	Deze scholen hebben een OKI-waarde $\geq 0,5$ en < 1.	18,5
blauwe groep	Deze scholen hebben een OKI-waarde ≥ 1 en < 2.	14,4
groene groep	Deze scholen hebben een OKI-waarde ≥ 2 .	9,7



6

1.4. Evolutie Nederlands – begrijpend lezen

Zoals gezegd, is drie jaar na elkaar begrijpend lezen in de focus gebleven, waardoor op populatieniveau een vergelijking mogelijk is. Die leert dat het leergebied nog steeds achteruitgaat in weerwil van de inspanningen.

2. Kwaliteitsontwikkeling als organisatie

Met de IDP-toetsen die in mei-juni 2020 worden afgenomen, is men gestart in september-oktober 2018, licht de spreker toe. Betrouwbaarheid vereist een systematische aanpak, die ook onderzocht werd door professor Janssen in het perspectief van de validering. Het hele proces van opgaven ontwikkelen, testen, herwerken en door experts laten beoordelen, vereist die lange doorlooptijd.

Daarnaast vindt heel wat onderzoek op de resultaten plaats. Universiteiten worden jaarlijks aangeschreven met onderzoeksvragen. Anne Verhoeven verzekert dat de resultaten worden gebruikt om het systeem te verbeteren. Ze zijn op aanvraag beschikbaar, vult ze nog aan.

Interne kwaliteitsontwikkeling betekent ook professionalisering van de pedagogische begeleiding. In het traject 'coach de coach' begeleiden begeleiders elkaar. Een breed overleg over heel Vlaanderen komt geregeld samen over (visie op) evaluatie. Uiteraard vindt data-analyse plaats over de output van de didactiek en het leerplan. Ondertussen wordt achter de schermen de vernieuwing van het digitale toetsenplatform voorbereid, dat inmiddels tien jaar oud is. De vormgeving van leerwinstmeting wordt daaraan gekoppeld, samen met het CLB. Zij onderstreept daarbij dat de namen van de leerlingen versleuteld worden, omdat het niet de bedoeling is de leerwinst van individuen te meten, wel die van scholen en klasgroepen. Op termijn zit daar ook de uitbreiding van de leeftijdsgroepen in.

Tot slot is men bezig met het evenwicht tussen aanbod- en vraaggestuurd werken. De spreker is er zich van bewust dat het beleid de nadruk legt op dat laatste, maar wijst erop dat er scholen blijven die na ontvangst van de resultaten geen vragen richten aan de koepel. Dat kan zijn omdat het beleidsvoerend vermogen zo groot

is dat ze geen begeleiding nodig hebben, maar als dat niet het geval is, kan een begeleidingsaanbod wenselijk zijn.

3. Kwaliteitsontwikkeling op school

Wat het niveau van de school betreft, vindt er vanzelfsprekend voor en tijdens de afname veel communicatie plaats. Er zijn infosessies over toetsen en over datageletterde verwerking van een rapport, naast het materiaal dat doorlopend in de genoemde evaluatiebox beschikbaar is, en een helpdesk. Wie online de toets afneemt, is verplicht eerst de 'even-proberen-toets' te doen, als test van het digitale systeem op school.

Na afname ontvangt de school haar resultaten eind juni per leergebied, per scherpe focus en per afzonderlijke vraag. Tegelijk krijgen ze die op Vlaams niveau, volgens hun OKI-groep en op school- en klasniveau. De leerkracht krijgt ook het individuele leerlingenniveau, al onderstreept de spreker dat het niet de bedoeling is op grond daarvan wel of niet een getuigschrift toe te kennen. Daarnaast ontvangt de school nog een analysedocument met opgaven en antwoorden, de leerplandoelen en de scores per opgave op Vlaams niveau.

Bij het begin van het nieuwe schooljaar zijn de samenvattingen en handvatten per leergebied klaar op basis van de Vlaamse resultaten, met als titel 'Wat doet jouw school met de resultaten van IDP 2.0?' Geïnteresseerden kunnen die voor de leergebieden van vorig schooljaar op de website vinden. Doorheen het schooljaar blijven ondersteunend materiaal en pedagogische begeleiding beschikbaar.

Hilde Hendrickx, die als schoolbegeleider actief is in de regio Antwerpen, vult aan dat de pedagogische begeleiding er scholen ondersteunt op twee niveaus bij de analyse en de opvolging van de IDP-rapporten.

Schooloverstijgend kan men tegelijk een groot aantal scholen bereiken. In september-oktober gaat men aan de slag met kernteams uit verschillende basisscholen in werksessies. Aan de hand van een vast stappenplan worden de data van de eigen school bekeken, om de teams te leren welke interessante informatie in de rapporten te vinden is. De resultaten worden bekeken op leergebied en -domeinniveau, vergeleken met de referentiegroep en de evolutie van de laatste vijf jaar daarin. Ze onderstreept daarbij wel dat IDP een signaalfunctie heeft. Op de analyse volgt een eerste besluit, waarin de elementen voor borging en bijsturing geïdentificeerd worden, en waarmee de kernteams terugkeren naar hun school om eventueel verdere acties te plannen.

Scholen worden ook aangemoedigd om het genoemde stappenplan elk jaar ook op de school zelf te doorlopen om systematischer van data gebruik te leren maken. Soms wordt ook gewerkt op het niveau van de scholengemeenschap, bijvoorbeeld over het ROK-kader of om van elkaar te leren. Verder is er ook in de opleiding voor nieuwe directeurs een sessie over datageletterdheid, waarin ze kennis maken met IDP-rapporten, analysedocumenten en het stappenplan.

De behoefte aan ondersteuning op schoolniveau blijkt sterk af te hangen van het beleidsvoerend vermogen. Ook de werking op dat niveau start met de samenstelling van een kerngroep, die eveneens leergebieden en -domeinen analyseert en het verband legt met andere data op school om te zien of daarin dezelfde signalen optreden. Op grond van de resultaten wordt uitgemaakt of het nodig is om met behulp van het genoemde analysedocument ook voor het vraagniveau te kijken wat de specifieke resultaten zijn, wordt samen met de school een actieplan opgesteld, met operationele doelen, acties, opvolging en effectmeting.

De pedagogische begeleiding werkt met ankerbegeleiders, waardoor een vertrouwd contact ontstaat en ook bij andere prioriteiten de band met de IDP-data kan worden gelegd. Uiteraard worden directeurs gestimuleerd hun data-analyse en actieplanning te delen met het schoolteam. Betere resultaten zijn immers een gezamenlijke verantwoordelijkheid.

Lieven Boeve nodigt de commissie uit om de afname van een IDP-toets, de verwerking van de resultaten en de begeleidingstrajecten eens mee te maken. Naar aanleiding van de PISA-resultaten spraken de onderwijsverstrekkers met de minister af om aan de slag te gaan voor ambitieuze eindtermen en gestandaardiseerde proeven. Zijn koepel hoopt met zijn ervaring daarin van nut te kunnen zijn.

III. Uiteenzetting door het Steunpunt voor Toetsontwikkeling en Peilingen

Rianne Janssen vertegenwoordigt STEP, een interuniversitaire samenwerking tussen KU Leuven en de Universiteit Antwerpen.

1. Peilingen

Peilingen zijn er in alle OESO-landen behalve in Griekenland en Portugal. Het gaat om grootschalige, gestandaardiseerde toetsafnames bij leerlingen in het leerplichtonderwijs. In Vlaanderen bestaan ze sinds 2002. Ze zijn netoverschrijdend en gericht op de eindtermen, bij een voor Vlaanderen representatieve steekproef van scholen en leerlingen. Ze gebeuren voor diverse leer- en vakgebieden alsook voor leergebiedoverstijgende domeinen en dit in verschillende onderwijsniveaus en -vormen.

De spreker geeft een overzicht van de peilingen die sinds 2002 hebben plaatsgevonden in basis- en secundair onderwijs. Zij laten een breed palet van eindtermen zien.

2. Doelstellingen

2.1. Systeemniveau

De hoofddoelstelling bevindt zich op systeemniveau. In eerste instantie wil men in een criteriumgerichte meting nagaan in hoeverre leerlingen de eindtermen behalen. De peiling vindt plaats op een bepaald tijdstip met oog voor verschillen tussen leerlingengroepen. Bovendien wordt aan de hand van herhalingspeilingen de evolutie nagegaan. Daarnaast wordt vergelijkenderwijs naar de verschillen tussen scholen gekeken, zowel qua feitelijke prestaties als qua prestaties na rekening te hebben gehouden met de achtergrondkenmerken van de leerlingen op een school. Peilingen passen in het systeem van externe kwaliteitszorg op systeemniveau. In de cyclus 'plan-do-check-act' zitten ze op checkniveau. Ze bieden evaluatie en feedback aan onderwijsactoren, die in de volgende cyclus opnieuw kunnen starten met borging en bijsturing.

2.2. Schoolniveau

Op schoolniveau wordt bij een peiling feedback gegeven aan de deelnemende scholen, zowel over de prestaties van hun leerlingen als over de prestatie van de school in vergelijking met andere scholen. Daarbij wordt kwantitatief in kaart gebracht welke bijdrage een school levert aan het leren van de leerlingen. Daarnaast zijn paralleltoetsen ter beschikking, die scholen op eigen initiatief kunnen inzetten in het kader van hun interne kwaliteitszorg. Het steunpunt onderhoudt daarover een website en stuurt vier keer per jaar een nieuwsbrief.

3. Onderzoeklogica

Daarop gaat Rianne Janssen nader in op de onderzoeklogica. Peilingen worden ontwikkeld in het kader van een project dat over bijna vier jaar loopt. Het peilingsjaar wordt immers voorafgegaan door voorbereiding en (toets)ontwikkeling, en gevolgd door valorisatie. Elk van de vier fasen kan verder worden onderverdeeld. Via resonantie-, opvolgings- en stuurgroepen (inclusief onderwijsinspectie, PBD's van de verschillende onderwijsverstrekkers, Vlor, onderwijsbeleid, lerarenopleiders) wordt samengewerkt met het veld. Daarnaast worden continu inhoudelijke experts betrokken.

De overheid bepaalt het domein waarover de peiling gaat. De stuurgroep bakent het verder af op basis van een selectie van eindtermen binnen het leer- of vakgebied. Dat gebeurt netoverschrijdend, waarna in samenspraak met de onderwijsverstrekkers en de opdrachtgever een toetsmatrijs wordt gemaakt. Die toetsmatrijs doet dienst als blauwdruk en heeft als doel de borging van de inhoudsvaliditeit. Verder vermeldt de spreker de ontwikkeling van items in het kader van grootschalige, gestandaardiseerde toetsenafnames, het voorkomen van zowel open als gesloten antwoordvormen, de evolutie naar digitale afname en de ontwikkeling van praktische proeven naast de schriftelijke voor sommige eindtermen. Die laatste worden niet meegenomen bij de paralleltoetsen maar aan de scholen aangeboden via brochures en KlasCement.

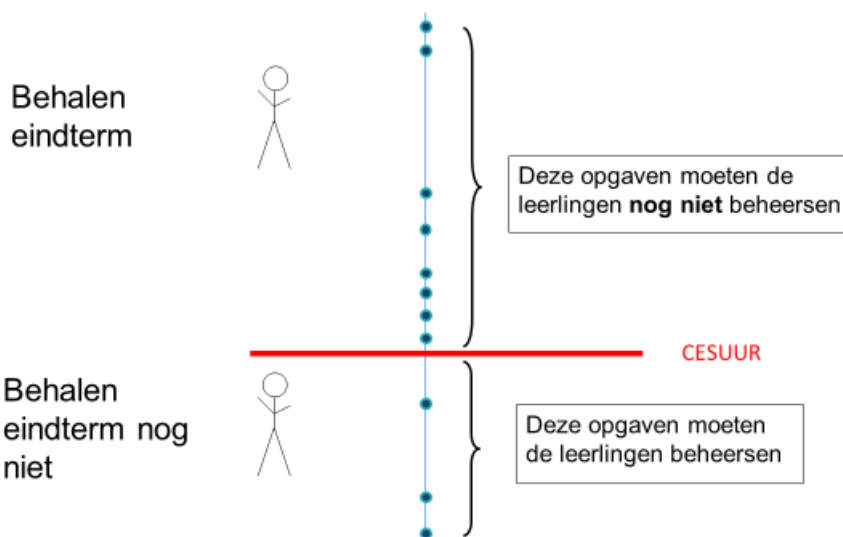
De toetsontwikkeling vindt plaats in het schooljaar voorafgaand aan de peiling met gedetacheerde leerkrachten en lerarenopleiders. Er gebeuren verschillende vooronderzoeken. Er vindt tevens een grootschalig kalibratieonderzoek plaats zodat men bij de start van de peiling al over een meetschaal beschikt.

3.1. Leerprestaties leerlingen

De meetschaal vormt de kern van de onderzoeklogica en is te vergelijken met een ladder waarop de opgaven in toenemende moeilijkheidsgraad gesitueerd worden. Leerlingen worden op dezelfde meetschaal gesitueerd in toenemende mate van vaardigheid, op basis van een probabilistisch model, dat wil zeggen volgens de slaagkans van de leerling voor een toetsvraag. De opgaven onder zijn locatie op de meetschaal beheerst de leerling (hoge succeskans), die erboven nog niet (lage succeskans).

Werken met een meetschaal biedt vier grote voordelen. Om te beginnen kunnen op basis van de meetschaal paralleltoetsen en peilingstoetsen worden ontwikkeld die equivalent zijn naar inhoud (conform de eindtermen) en moeilijkheidsgraad. Ten tweede kan men aan de hand van cesuurbepaling komen tot een criteriumgerichte meting. Een groep van experts (lerarenopleiders, inspecteurs, pedagogische begeleiders, leerkrachten, beleidsmakers) bepaalt welk niveau op de meetschaal minimaal nodig is om de eindtermen te behalen, waardoor een tweedeling ontstaat tussen opgaven die de leerlingen moeten beheersen en degene die zij nog niet hoeven te beheersen, hoewel die laatste opgaven ook binnen het domein van de eindtermen vallen. Door de cesuur aan de itemkant kan meteen ook een tweedeling worden gemaakt tussen leerlingen die wel en nog niet de eindtermen halen.

CESUURBEPALING



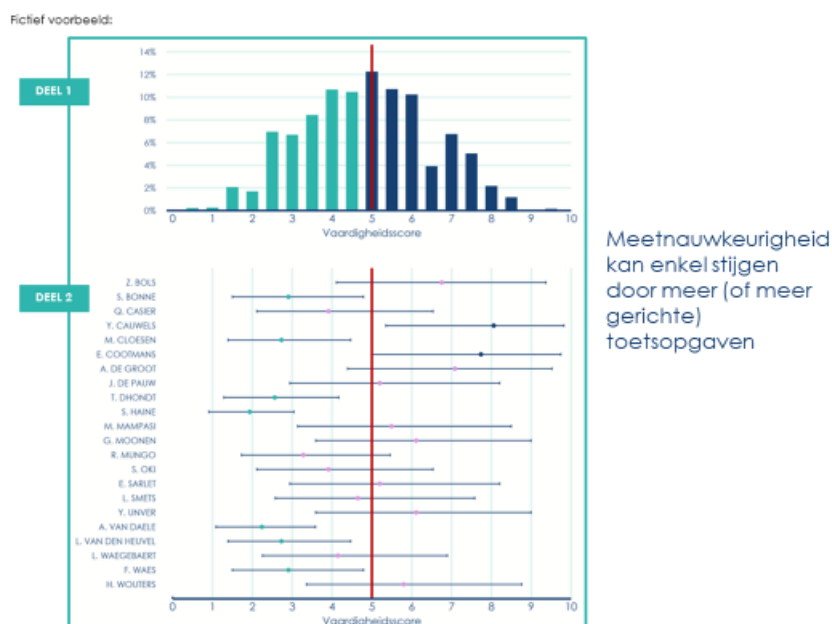
28

ST²P STEUNPUNT TOETSONTWIKKELING EN PEILINGEN

De spreker onderstreept dat feedback wordt gegeven op het niveau van eindtermen, maar niet per opgave. Voor wiskunde loopt het aantal deeltoetsen op tot bijna twintig. Een leerling die de eindterm net haalt, krijgt als score 5 op 10.

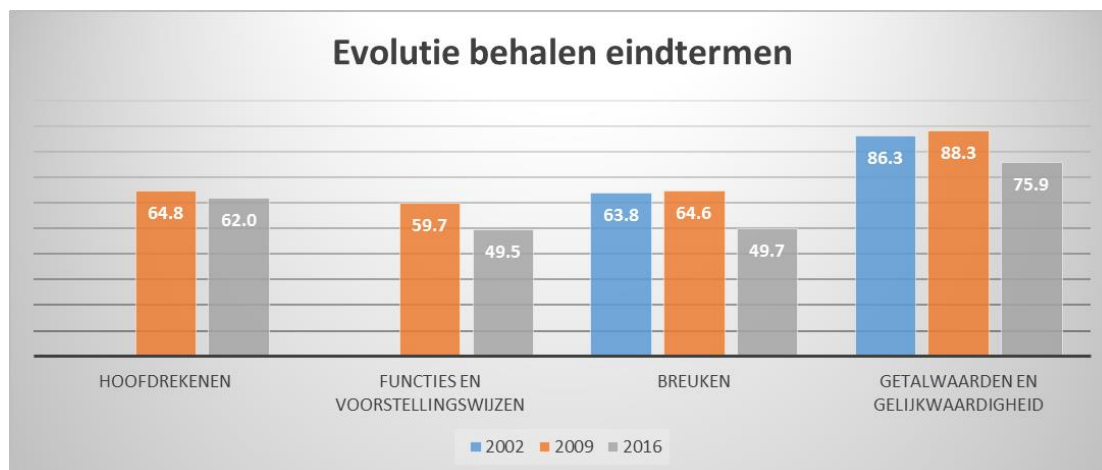
Een derde voordeel van de schaal is de mogelijkheid om de betrouwbaarheid aan te geven van het behalen van de eindtermen. In het onderstaande voorbeeld wordt de meetnauwkeurigheid per leerling aangegeven met lijntjes. Het toont dat slechts voor enkelen met statistische zekerheid kan gesteld worden dat ze de eindtermen wel of niet halen.

BETROUWBAARHEID OP LEERLINGNIVEAU



Het laatste voordeel is het meten van evoluties. Als voorbeeld verwijst de spreker naar de meetschalen van twee verschillende toetsen. Als er geen overlap is tussen

meetschalen zijn de resultaten niet vergelijkbaar. Het is immers niet geweten of een evolutie in de resultaten te wijten is aan een evolutie in het prestatieniveau van de leerling of aan een evolutie in de moeilijkheidsgraad van de toetsen waarop de vergelijking is gebaseerd. Om vergelijking toe te laten, worden ankeropgaven toegevoegd die gemeenschappelijk zijn aan de toetsen tussen verschillende meetmomenten. Daardoor kunnen wel uitspraken over trends gedaan worden. De meetschalen maken ook koppeling aan internationaal onderzoek mogelijk, zoals recent gebeurde met PIRLS. Als voorbeeld van de genoemde evolutie toont ze de opeenvolgende peilingen voor wiskunde in het basisonderwijs, waarbij voor sommige eindtermen een daling moet worden vastgesteld. Bovendien kan ook worden vastgesteld in welke mate een daling of stijging betekenisvol is dan eerder te wijten is aan het toeval.



Rianne Janssen vat samen dat de meetschaal cruciaal is voor een gestandaardiseerde, netoverschrijdende en criteriumgerichte meting. Samen met ankeropgaven garandeert zij bovendien vergelijkbaarheid doorheen de tijd. Ze vormt dan ook de basis voor een uitbreiding naar leerwinstonderzoek.

3.2. Schoolprestaties

Ook prestaties op schoolniveau kan men afleiden uit de peilings- en parallelproeven. Het model maakt het mogelijk om niet alleen de schoolgemiddelden zelf maar ook de betrouwbaarheid weer te geven, evenals welke scholen significant meer of minder presteren dan het Vlaamse gemiddelde.

Verschillen in leerprestaties van leerlingen zijn echter niet alleen toe te schrijven aan de school, maar ook aan omgevings-, fairness- en schoolcontextkenmerken. Daarom proberen de onderzoekers in tweede instantie de prestatieverschillen op te splitsen, om te kunnen aangeven welke de toegevoegde waarde of de bijdrage van de school zelf aan de toetsresultaten van de leerlingen is.

In het CIPO-model staat O voor output, die men wil corrigeren door rekening te houden met context (C), input (I) en proceskenmerken (P). De spreker overloopt daarbij de gehanteerde kenmerken. Leerlingenmerken: geslacht; eventuele schoolse voorsprong of achterstand; eventuele diagnose voor dyslexie, dyscalculie, ADHD, autisme of een andere beperking of stoornis; thuistaal; aantal boeken thuis; opleidingsniveau van de moeder; en het al dan niet ontvangen van een studietoelage. Schoolkenmerk: gemiddelde score op de financieringscriteria.

Op basis daarvan kan naast het feitelijke ook het verwachte schoolgemiddelde berekend worden. Op basis van het verschil tussen feitelijke en verwachte prestaties,

kan een nieuwe vergelijking van scholen gemaakt worden op basis van de toegevoegde waarde.

Naast het behaalde niveau van de leerling kan men ook zijn vooruitgang bekijken. De meting van dergelijke leerwinst vereist een longitudinaal onderzoeksopzet. De vraag is dan hoe goed een leerling het doet in het zesde leerjaar in vergelijking met de prestaties in eerdere leerjaren (vooruitgang op de meetschaal) en in vergelijking met leerlingen (van andere scholen) die dezelfde startpositie hadden. Kortom, de vandaag voorgestelde methodologie kan uitgebreid worden naar leerwinstmeting, zowel op leerling- als schoolniveau.

De spreker vat ook dit deel van haar uiteenzetting samen. De verschillen tussen scholen worden eveneens nagegaan op basis van een statistisch model. Niet alleen hun prestaties op zich worden vergeleken maar ook hun toegevoegde waarde, door rekening te houden met de kenmerken van de deelnemende leerlingen.

4. Schoolfeedback bij paralleltoetsen

Naast een uitgebreide leeswijzer, krijgen de scholen om te beginnen een overzicht van de deelnemers en van hoe goed hun leerlingen de eindtermen bereiken, in de vorm van percentages in een tabel, die ze kunnen vergelijken met de Vlaamse gemiddelden. Verder krijgen ze individuele leerlingenfeedback en een overzicht van hun ruwe schoolresultaten en hun toegevoegde waarde (in vergelijking met andere scholen).

5. Gebruik

Tijdens het schooljaar 2018-2019 nam 8 procent van de scholen van het zesde leerjaar van het basisonderwijs minstens één paralleltoets af. Men kan daarvoor verwijzen naar de context dat de koepels hun eigen toetsen inrichten, maar voor het secundair onderwijs, waar die niet beschikbaar zijn, haalt men eveneens slechts 8 procent in de derde graad aso/tso, en elders nog minder. Scholen geven als motief het ontbreken van een verplichting en van een toetscultuur. Overigens zijn er geen paralleltoetsen voor alle leer- en vakgebieden op alle niveaus. Bovendien besteden leraren de tijd liever aan eigen lessen.

De invoering van verplichte toetsen op het einde van het zesde leerjaar heeft wel geleid tot een duidelijke toename. De gelijktijdige maar beperktere toename in het secundair onderwijs is vorig jaar gestagneerd.

Mede op basis van een gebruikersonderzoek werd een stappenplan ontwikkeld voor de begeleiding van scholen. Daarin worden verschillende fasen onderscheiden om aan de slag te gaan met schoolfeedback van peilingen en paralleltoetsen. Het werd opgesteld in samenwerking met de pedagogische begeleidingsdiensten. Het steunpunt biedt ook vorming aan scholen in de vorm van workshops en pedagogische sessies voor begeleidingsdiensten. Het steunpunt heeft zelf geen bevoegdheid op het vlak van begeleiding, benadrukt de spreker, maar faciliteert en ondersteunt. Het doet wel onderzoek naar de processen die een rol spelen bij effectief gebruik van de peilingsresultaten.

6. Besluit

Tot besluit noemt Rianne Janssen de peilingen een uitgebouwd systeem van kwaliteitszorg voor het Vlaamse onderwijspeil, in samenwerking met verschillende actoren. Ze hebben ook geleid tot paralleltoetsen die scholen in het kader van hun eigen interne kwaliteitszorg kunnen inzetten voor een specifiek leer- of vakgebied. De toekomst is nog een vraagteken in het licht van de enigszins enigmatische zinsnede in de beleidsnota 'ter vervanging van bestaande proeven'. De implementatie

van de gewenste leerwinstmeting vergt gevalideerde, gestandaardiseerde en genormeerde proeven, niet alleen voor het bereiken van de eindtermen maar ook voor de leerwinst van leerlingen en de toegevoegde waarde van scholen. Tot slot vraagt de minister ook internationale benchmarking. Het steunpunt denkt na over de haalbaarheid en is bereid tot overleg.

IV. Bespreking

1. Tussenkomsten van de leden

Loes Vandromme vraagt wat voor de sprekers een werkbare definitie van leerwinst is. Wat levert de meting daarvan op voor de scholen? Zij wordt in de paralleltoetsen doorgetrokken van groeps- naar leerlingenniveau maar hoe staan de andere ontwikkelaars daar tegenover? Hoe kan de aandacht gevestigd worden op andere dan cognitieve aspecten zoals creativiteit en empathie, die evenzeer cruciaal zijn voor kwaliteitsvol onderwijs? Hoe kan men ervoor zorgen dat in scholen die minder goed scoren op bepaalde punten, effectief begeleiding volgt? Is men voorstander om daar zelf naartoe te stappen, naast het vraaggestuurde aanbod? Hoe gaan scholen om met hun feedback? Gaan ze innovatief nadenken? Waar kan volgens de sprekers nog kwaliteitswinst geboekt worden? Wat zijn de toekomstperspectieven van de OVSG-toetsenwijzer? Tot slot suggereert het lid dat de commissie zichzelf laat toetsen om daaruit te leren.

Jo Brouns informeert of de toetsenmakers hun expertise over ontwikkeling, organisatie en aanwending met elkaar delen. Is men daartoe bereid? Werden na de engagementsverklaring al afspraken gemaakt met de minister? Hoe ziet men zijn rol daarin? Hij vraagt ook meer uitleg over het digitaliseringsproject. Van OVSG wil hij een voorbeeld over hoe men met lokale besturen aan de slag gaat met de toetsenresultaten.

Steve Vandenberghe mist de situering van de gestandaardiseerde toetsen in het geheel van het evaluatiebeleid van scholen. Ze worden immers afgelegd onder stress waardoor niet alle kinderen hun gebruikelijke resultaten behalen. Kwaliteitsbewaking vereist objectieve outputmeting maar mag daartoe niet beperkt zijn. Hoe zien de sprekers dat?

Koen Daniëls had nooit de vandaag gebleken algemene waardering voor gevalideerde toetsen verwacht. Nog niet zo lang geleden werd enkel op gevaren als 'teaching to the test' gewezen, en niet op de opportuniteiten. Verder wil hij weten waarom het puntenniveau wordt geplafonneerd op het eindtermniveau in de plaats van ruimte te laten voor excelleren. Zou het niet zinvol zijn om ook in peilingsproeven verder te gaan dan de eindtermen, zoals in de paralleltoetsen of in het PISA-onderzoek?

Neemt elke leerling deel of zijn er scholen die op het moment van de afname differentiëren en sommige leerlingen iets anders laten doen, wat natuurlijk de resultaten beïnvloedt? Wat is er mis mee om de resultaten van gevalideerde, betrouwbare toetsen niet ook op leerlingenniveau te gebruiken, als er toch tijd in geïnvesteerd wordt? Waarom niet, in het licht van de gemeenschappelijke eindtermen, één toets voor alle Vlaamse scholen? Is men het oneens over excellentie? Of is er sprake van een leerplaneffect? Probeert men op dat vlak invloed uit te oefenen?

De laatste vraag van het lid gaat over de verhouding tussen de eindtermen en de extra doelen die men meet. Is het in het licht van het vele voorbereidingswerk mogelijk om alles op te nemen in de toetsen? Of is het beter een aantal thema's te selecteren, waarvan men dan over de jaren heen de resultaten kan vergelijken?

Jan Laeremans sluit zich aan bij de vraag naar een gemeenschappelijke toets. Zit er geen overlapping in het werk dat men levert voor de lagere school? Verder vraagt hij naar een vergelijking van de tijd die nodig is voor de toetsen van de verschillende koepels en het steunpunt. Wordt de door het katholiek onderwijs gesignaleerde en door PISA bevestigde achteruitgang van de leesvaardigheid ook gemeten door OVSG? Wat zijn hun cijfers over het laatste decennium?

Ook *Hannelore Goeman* vraagt welke voor- en nadelen men ziet aan de ontwikkeling van een gecentraliseerde toets voor basis- en secundair onderwijs. Het beleid zou er alvast meer algemene conclusies uit kunnen trekken. Hoe kan men garanderen dat de meting van leerwinst en toegevoegde waarde op schoolniveau niet leidt tot de opstelling van een rangorde die de schoolkeuze beïnvloedt, zonder dat men de metingen zelf opgeeft?

Kathleen Krekels peilt naar de hogere principes van de toetsenontwikkelaars. Hoe verhoudt zich de afhankelijkheid van de leerplandoelen tot die van de eindtermen? Welke mogelijkheden tot versmelting van de bestaande toetsen ziet men, en wat zijn de hindernissen? Hoe komt het dat OVSG en het katholiek onderwijs, anders dan het steunpunt, nooit het initiatief namen om ook toetsen voor het secundair onderwijs te ontwikkelen? Worden de OVSG-toetsen ook twee jaar voorbereid zoals de IDP? Waarom worden die laatste uitdrukkelijk niet gekoppeld aan de uitreiking van een getuigschrift? Kunnen ze niet in bepaalde mate gebruikt worden? Is het niet zo dat leerwinst op leerlingenniveau alleen kan worden gemeten met de toetsen die de leerkracht doorheen het schooljaar afneemt? Hoe zien de ontwikkelaars dat precies?

2. Antwoorden

Patriek Delbaere verzekert dat OVSG er vanaf het begin voor gekozen heeft ook andere domeinen en zaken te meten dan alleen Nederlands, wiskunde, wetenschappen en het cognitieve, al is het niet altijd evident. Als voorbeeld noemt hij de fietsproef. Eens te meer daagt hij de commissie uit om de OVSG-toets mee te maken en vast te stellen dat men voor de opdrachten naast kennis ook andere vaardigheden nodig heeft. De praktijk kan erg leerzaam zijn.

Omgaan met scholen die falen op het vlak van onderwijskwaliteit, is niet eenvoudig. Ze vragen niet spontaan om hulp, doch vaak pas na een zeer negatieve doorlichting door de inspectie. De toets is dan ook een belangrijk instrument bij proactief optreden. Naast de klassieke driehoek van school, inspectie en pedagogische begeleiding, moet ook het schoolbestuur geresponsabiliseerd worden om te zorgen voor kwaliteit, en niet alleen voor de gebouwen en het personeelsbeleid. Daar is vroeger te weinig aandacht aan besteed. Vraag- en aanbodgestuurd begeleiden, is een evenwichtsoefening, want men moet het vertrouwen behouden. Pedagogische begeleiding is geen inspectie.

Naast de centraal ontwikkelde OVSG-toets is er de toetsenwijzer. Deze is van onderuit gegroeid bij scholen(groepen) die in eigen interne toetsengroepen hun evaluatiebeleid hebben geënt op de eindtermen en de OVSG-toetsen voor het zesde leerjaar. In die regio's werd elk jaar een batterij vragen ontwikkeld, die in de toetsenwijzer naar boven zijn getild. Op dat vlak zijn nog veel mogelijkheden, die evenwel mensen en middelen vereisen. Dergelijke lokale initiatieven vereisen bij veralgemening wel enige centrale validering, waarschuwt hij.

Voor haar toets werkt OVSG al een aantal jaar samen met het GO!, dat de toets vertaalt naar haar eigen leerplan. OVSG is bereid de expertise te delen, net zoals het is ingegaan op de suggesties van het steunpunt na de validering van de OVSG-toetsen. De bestaande, sterke vormen van toetsenbeleid zijn inderdaad historisch naast elkaar ontwikkeld. Voor verdere stappen kijkt hij in de eerste plaats naar het

secundair onderwijs, waar nog een hele weg af te leggen is en dus wordt vertrokken van een wit blad. OVSG had als kleine speler op dat niveau zelf nooit de middelen om daarvoor een toets te ontwikkelen.

Hij wijst erop dat de pedagogische begeleidingsdiensten van OVSG en het katholiek onderwijs met hun blik op kwaliteit een ander perspectief hebben dan de overheid met haar peilings- en parallelproeven. Hij pleit ervoor de rolverdeling scherper te stellen om overlapping te vermijden. Is het bijvoorbeeld niet beter dat de overheid bijspringt op het secundaire niveau als de onderwijsverstrekkers het daar niet waarmaken, in de plaats van de koepels te beconcurreren met eigen paralleltoetsen in het basisonderwijs – al erkent hij wel het academische kader daarvan?

De digitale omslag faciliteert accurate feedbackrapporten. Zij worden besproken met de schoolbesturen en het lokale anker van de pedagogische begeleiding, waarin goede en slechte punten benoemd worden. Hij bevestigt andermaal dat de OVSG-toets slechts een element is van een brede blik op leerlingen. Recent Nederlands tijdsbestedingsonderzoek is in dat verband veelzeggend: een leraar besteedt 8 uur per week van de 40 uur werken aan het maken, afnemen en verbeteren van toetsen.

De digitalisering maakt het mogelijk om resultaten op leerlingenniveau in kaart te brengen, maar de spreker pleit voor voorzichtigheid in de beoordeling van leerwinst. Hij waarschuwt ook voor het gevaar om scholen of leerlingen erop af te rekenen.

Steven De Laet vult aan dat elk van de zes toetsenblokken tussen 90 en 120 minuten vergt. Voor de leerlingen gaat het dus om drie halve dagen. Hij bevestigt dat ook OVSG geen topprestaties voor lezen vaststelde maar evenmin wat uit PIRLS naar voren kwam. Er waren indicaties dat het niet goed gaat, maar er ging geen alarmbel af. Er volgen nu net als in het katholiek onderwijs herhalingsmetingen om een en ander sterker in beeld te brengen.

In grote mate doen alle leerlingen mee, al kan men dat niet exact controleren. Sommige scholen schrijven leerlingenlijsten in en niet zichzelf. Men kan nooit een toets zodanig organiseren dat men alles volledig onder controle houdt. De stress voor de leerlingen probeert men communicatief te reduceren, al hangt dat ook af van scholen en leraren. OVSG onderstreept zelf altijd dat deelname geen effect heeft op de verdere schoolloopbaan. Er zijn ook mogelijkheden ingebouwd voor leerlingen met specifieke onderwijsbehoeften, over de inzet waarvan de klassenraad beslist.

De OVSG-toets is breed, bevestigt hij nogmaals. Zij neemt elk domein en alle leergebiedoverschrijdende eindtermen mee. Het bronnenboek illustreert dat. De doorlooptijd bedraagt ongeveer een jaar en twee maanden, en is daarmee iets korter dan bij de collega's. Het proces is wel hetzelfde, maar de procedure is strak.

Elias Hemelsoet sluit zich, wat leerwinst betreft, aan bij het onderscheid tussen de natuurlijke groei van een kind en het verschil dat de school daar doorheen de tijd bovenop biedt. Ook hij gewaagt van toegevoegde waarde. OVSG trekt in zijn toetsen expliciet de kaart van de leerplannen, die een concretisering van de eindtermen vormen. Door de taal van de school te gebruiken, bevordert men ook de herkenbaarheid voor de leerlingen. Bovendien beperken de plannen zich op die manier niet tot de minimumdoelen die de overheid met de eindtermen vastlegt.

Wat de garantie van begeleiding betreft voor scholen die negatief scoren, wijst hij op de privacy. OVSG heeft voor het eerst de vraag gesteld aan de scholen om hun gegevens transparant te maken voor begeleiding en bestuur. Het overgrote deel

verleende daar toestemming voor, maar het is wenselijk dat dit ook de standaard wordt. Begeleiding afdwingbaar maken veronderstelt wel een decretaal kader.

Verdere kwaliteitswinst is nog mogelijk door academisch onderzoek zoals de voorbije doorlichting door het steunpunt, die OVSG als ondersteunend heeft ervaren. Middelen zijn ook een aandachtspunt, want men krijgt geen specifieke middelen van de overheid voor de ontwikkeling van de toetsen, die men bovendien betaalbaar wil houden voor de scholen.

Lieven Boeve sluit zich aan bij het pleidooi om het eens te worden over een werkbare definitie van leerwinst, zodat daar alvast geen discussie over is. Hij kan zich bij de door het steunpunt voorgestelde definitie aansluiten omdat zij sterk focust op wat de school zelf kan toevoegen. Scholen die in moeilijke omstandigheden werken, kunnen ermee opschuiven in de rangorde. Dat is zinvolle informatie.

De reikwijdte van de evaluatiebox van Katholiek Onderwijs Vlaanderen is breed. Er loopt zelfs een experiment waarin de attitudes – en niet de kennis – ten aanzien van burgerschap worden getoetst. Daarnaast vermeldt hij de praktische proeven als Kunstroof. Dergelijke oefeningen verdienen van nabij opgevolgd te worden om te zien of ze wel het juiste toetsen. Dat voorbehoud is noodzakelijk. Dergelijke oefeningen verdienen van nabij opgevolgd te worden om erover te waken dat op het eerste zicht minder toetsbare doelen niet uit het oog verloren worden.

Wat kwaliteitscultuur betreft, wijst hij erop dat hij een netwerkorganisatie van schoolbesturen vertegenwoordigt. De gegevens van de toetsen zijn hun eigendom. Als een school geen toestemming geeft er iets mee te doen, kan men niet zoals het GO! met zijn centrale bestuursniveau, beslissen dat men toch in alle scholen dit of dat gaat doen. Eerst moeten alle scholen akkoord gaan. De paradox van kwaliteitsontwikkeling bestaat erin dat men daarvoor vooraf al heel wat beleidsvoerend vermogen, dus een begin van kwaliteitscultuur nodig heeft. Vraaggestuurd werken, botst op die grens, al kan men natuurlijk een vraag provoceren. Meer dan proberen te overtuigen, is echter niet mogelijk. Hij hoopt dat het nieuwe Kwaliteitsdecreet daar oog voor heeft, in het licht van de keuze van het regeerakkoord voor hoofdzakelijk vraaggestuurde begeleiding. Anderzijds stelde men zoals gezegd vast dat de eerste toets Frans meteen tot een verveelvoudiging van de vraag om begeleiding heeft geleid. Toetsenbeleid kan wel degelijk wegen op het kwaliteitsbeleid, besluit hij.

Het vrij onderwijs heeft een andere verhouding met de eindtermen dan het officieel onderwijs. Daarom is er ook vrijheid van onderwijs. De overheid mag minimumnormen bepalen voor scholen, maar daar wil het katholieke net vrij mee omgaan, meer bepaald door sterk in te zetten op de leerplannen. Zijn antwoord op de vraag naar complementariteit is dan ook gelijkaardig aan dat van OVSG, maar dan in het kwadraat, en voor de kleine onderwijsverstrekker is dat waarschijnlijk het geval tot de derde macht. Hij wijst erop dat de leerplannen door de inspectie goedgekeurd zijn op basis van de concordantie met eindtermen, waardoor scholen de zekerheid hebben dat ze bij het realiseren van de leerplandoelen minstens ook de eindtermen halen. Het katholiek onderwijs toetst al 52 jaar de uitvoering van zijn leerplannen en houdt daarmee de scholen een spiegel voor. Het netwerk van scholen heeft het immers altijd belangrijk gevonden data te genereren waarmee scholen aan kwaliteitsbeleid kunnen doen.

Lieven Boeve denkt dat de netten samen aan toetsenontwikkeling kunnen doen als het model ook zelf gemodelleerd is. Hij acht een toetsenbank denkbaar met vragen die iedereen kunnen toetsen op eindtermniveau, maar waarin tegelijk variatie is gebouwd op leerplanniveau. Dat zou het voor de overheid mogelijk maken een gestandaardiseerde toets in te voeren, zonder dat tegelijk analoge toetsen van het

katholiek onderwijs nodig blijven. De spreker ziet heil in een consortiumformule met mede-eigenaarschap van diverse stakeholders en universitaire expertise.

Daarop herinnert hij eraan dat een van de oorspronkelijke vragen die men met peilingstoetsen wilde beantwoorden ook was of de eindtermen zelf goed zitten. Hij beveelt aan om het instrument in te zetten om diezelfde vraag te stellen over de nieuwe eindtermen van de eerste graad en de toekomstige voor de tweede en derde graad. Er is immers niets erger voor een leerlingenpopulatie dan eindtermen die niet goed zitten. Hij pleit er dan ook voor om ook die doelstelling te integreren in de peilingstoetsen. De spreker vat samen dat samenwerking mogelijk is als de sterktes van de verschillende modellen worden opgenomen in een consortiumformule.

Digitalisering is ook afhankelijk van infrastructuur, die in het licht van de huidige werkingsmiddelen een punt van zorg is. Hij rekent op het aangekondigde actieplan.

Verder is de spreker het helemaal eens dat een evaluatiebeleid omvattend moet zijn. Toetsen werken in een breder kader. Er is altijd maatwerk nodig. Hij nuanceert in dat verband de toepassing op leerlingenniveau: de resultaten zijn expliciet niet daarvoor bedoeld en wel om de genoemde high stakes te vermijden, maar dat neemt niet weg dat men ze op dat niveau alsnog kan gebruiken, op voorwaarde dat ze in een breder kader geplaatst worden. Dat laatste is ook de reden waarom het rapport al in juni beschikbaar wordt gesteld. Doel van de nuancering is om te vermijden dat de toetsen gebruikt worden om leerlingen uit te sluiten. Dergelijk ongewenst gedrag wordt nog versterkt door vergelijking in een rangorde.

IDP heeft een signaalfunctie en werkt met wisselende focussen, waardoor niet altijd alles wordt getoetst. Daar wordt nu van afgeweken, wat overigens in toenemende mate zal gebeuren, vanwege de verrassende PIRLS-resultaten. Nu blijkt dat ze nog altijd achteruitgaan, wat echter niet zo verrassend is omdat het dezelfde cohorte betreft. Hij hoopt dat voor de volgende PIRLS de inspanningen op en buiten de school over (begrijpend) lezen vruchten afwerpen.

Toetsen duren een uur, vanwege de aandachtsspanne van de leerling.

De spreker belijdt nogmaals zijn geloof in de mogelijkheid van een model dat een netoverschrijdende focus heeft en tegelijk de verschillen naar waarde schat. Voor zijn net is dat laatste een voorwaarde om voluit aan dergelijk initiatief deel te nemen. Waarom zou men anders een beroep doen op de vrijheid van onderwijs om een eigen net in de markt te zetten?

Er zijn een aantal factoren die verklaren waarom er nog geen dergelijke toetsen zijn in het secundair onderwijs. Om te beginnen zijn basisscholen kleiner en voelen zij sneller dat ze de kritische massa ontberen om zelf voldoende data te verzamelen en te hanteren voor het voeren van een kwaliteitsbeleid. Verder bestaat in het secundair onderwijs een grote variatie aan richtingen en leerplannen. Als derde vermeldt hij de professionaliteit van de masters in de derde graad, die eigen dingen doen met het leerplan en daarom minder geneigd zijn om gestandaardiseerde toetsen te vragen. Zelfs samen examens opstellen op schoolniveau vergt veel overtuigingskracht. Onmogelijk is het op dat niveau niet, maar het model zal ook dat element van beroepsfierheid moeten waarderen. Er werd wel al nagedacht over de verruiming van de IDP-methodiek tot de nieuwe eindtermen voor de eerste graad, die immers nog voor een zeer grote cohorte gelden. In afwachting van de uitkomst van de gesprekken met de Vlaamse Regering, is dat evenwel opgeschort.

De ontwikkeling van de IDP-toetsen – zonder de begeleiding – kost het katholiek onderwijs niet meer dan 6 vte (ICT- en helpdeskmedewerkers inbegrepen) dankzij de decennialange ervaring. Er worden ook niet voor alles elk jaar nieuwe toetsen

gemaakt. De inhoud wordt aangeleverd door de pedagogische begeleiding. Voor Lieven Boeve behoort die vertaling van het pedagogische project wel degelijk tot hun kernopdracht. Zij moeten er nadien trouwens mee werken. Een en ander wordt dus betaald door het lidgeld van de scholen aan de netwerkorganisatie. De toetsen zelf zijn gratis.

Rianne Janssen antwoordt dat de definitie van leerwinst nog steeds een hangende kwestie is. In eigenlijke zin gaat het om de vooruitgang van leerlingen over schooljaren heen. Idealiter bestaat voor de evolutie van de prestaties een leerlingvolgsysteem dat over leerjaren heen eenzelfde meetschaal hanteert. Daarnaast heb je de definitie van leerwinst als de toegevoegde waarde van scholen na controle voor de beginmeting bij hun leerlingen. Als men de opvolging beperkt tot een toets in het vierde en het zesde leerjaar, bestaat het gevaar dat scholen zich nog uitsluitend richten op leerlingen die net onder de lat blijven, waardoor nivellering van in het algemeen behaalde niveau kan ontstaan.

Voor haar blijft het meten van het niveau dat leerlingen halen belangrijk. Als men die meting uitvoert voorafgaand aan het toetsen van de eindtermen kan men aan de hand van statistische modellen uitmaken in welke mate scholen ertoe doen, en meer of minder vooruitgang boeken met hun leerlingen. Methodologisch is dat echter niet eenvoudig, want hoe brengt men bijvoorbeeld leerlingen die van school veranderen in rekening? Op individueel niveau meten zonder leerlingvolgsysteem vraagt een gigantische kalibratieoefening van de meetschaal, waarschuwt ze. De uitvoering vergt heel wat denkwerk.

Koen Aesaert bevestigt dat de ambitie van de beleidsnota om niet alleen het behalen van de eindtermen te meten, maar ook de leerwinst, en bovendien de resultaten aan internationale benchmarks te koppelen, niet kan zonder statistisch model dat de drie samenbrengt. Dat vereist inderdaad voorafgaand aan de leerwinstmeting een zware kalibratieoefening.

Belangrijk is ook de vraag wat men aan de scholen zelf wil meegeven. Of hun leerlingen de eindtermen halen, kan men inderdaad meten met een peilingstoets op het einde van het zesde leerjaar, waarin een cesuur is aangebracht. Aanvullend kan men al in het derde of vierde leerjaar een toets afnemen op dezelfde schaal (mits voorafgaande kalibratie). Als men er items uit PIRLS instopt, kan men in principe ook scores op de internationale schaal berekenen. Dat volstaat echter niet om de leerwinst te meten, want daarvoor is een eerlijker model nodig, dat onderscheid maakt tussen verwachte en feitelijke leerwinst.

Rianne Janssen vervolgt dat het in het secundair onderwijs nog complexer wordt, omdat leerlingen verschillende onderwijsvormen met eigen eindtermen volgen. Hoe moet men al die verschillende trajecten populatiebreed, zoals de minister wenst, in kaart brengen?

Uiteraard is het steunpunt bereid tot samenwerking. Het gaf vanuit zijn onderzoeklogica reeds groeigerichte feedback bij de validering van de bestaande proeven van de onderwijsverstrekkers. Een van de adviezen was dat de toetsen op leerlingniveau onvoldoende betrouwbaar zijn. Betrouwbare toetsen op leerlingniveau (zoals vooropgesteld bij de leerwinstmetingen) vereisen langere toetsen. Anders blijven de betrouwbaarheidsintervallen te breed. Verder zijn de huidige toetsen die door de onderwijsverstrekkers worden aangeboden niet criteriumgericht, kan de schoolfeedback nog beter en zitten er nog geen trendanalyses in. Ze wijst er wel op dat het steunpunt opdrachten van de overheid uitvoert en dat ze zelf geen initiatief te nemen heeft.

Alleen Nederlands en wiskunde toetsen, brengt het risico van curriculumvernauwing mee. Uit onderzoek bij scholen die paralleltoetsen gebruiken, blijkt hoe

belangrijk het eigenaarschap is bij schoolfeedback, benadrukt ze verder. Scholen moeten er zelf zin in hebben en erom vragen vanuit hun eigen kwaliteitsopvolging. Als het wordt opgelegd, ontstaan ongewenste neveneffecten.

De meetschaal kan zodanig opgebouwd worden dat ook het excelleren en verder gaan dan de eindtermen in beeld komt, maar dat kost mensen en middelen. In een digitaal systeem kan men immers gedifferentieerd toetsen, door bijvoorbeeld wie klaar is bijkomend moeilijkere vragen aan te bieden, maar ook door leerlingen die het moeilijker hebben met de basisvragen preciezer te positioneren op de schaal door hen makkelijkere vragen voor te leggen. Dergelijk adaptief systeem vergt echter een grote itembank en complexe modellen om de resultaten te verwerken. Voorts is het inderdaad mogelijk om aan de eindtermgerichte toetsopgaven andere opgaven toe te voegen zoals leerplangerichte, maar ook daar is bij wijze van spreken meer dan alleen verbeeldingskracht voor nodig, terwijl de voorbereidingstijd nu al enorm lang is.

Sinds een paar jaar wordt voorzien in redelijke aanpassingen voor leerlingen die bijvoorbeeld voorleessoftware willen gebruiken. Doel is zo weinig mogelijk leerlingen uit te sluiten.

Het steunpunt is geen voorstander van de ranking van scholen, maar als de overheid een proef oplegt en organiseert, weet ze niet wat precies de juridische consequenties zijn voor de openbaarheid van de data.

De tijd die nodig is om de toetsen van het steunpunt af te leggen, varieert. Bij peilingen vraagt men een voormiddag, maar niet alle scholen doen alle toetsen. Omdat het steunpunt een onvolledig afnamesdesign gebruikt om het hele domein van de eindtermen als systeem in kaart te brengen, hebben niet alle scholen informatie over elk deeldomein.

Het systeemniveau zal ook in de toekomst belangrijk blijven. Het mag niet beperkt blijven tot het stimuleren van scholen bij hun interne kwaliteitszorg. Besluiten op systeemniveau, bijvoorbeeld over de achteruitgang in een bepaald leergebied, terugkoppelen naar concrete acties, moet behouden blijven, al dan niet met een engagementsverklaring. Dataverzameling op systeemniveau is immers een belangrijke informatiebron voor het doel van de kwaliteitsverbetering.

3. Bijkomende vragen en antwoorden

Kathleen Krekels vraagt wat verloren zou gaan als de overheid ervoor zou kiezen om het steunpunt netoverschrijdende gestandaardiseerde toetsen te laten ontwikkelen voor de niveaus die de minister wil, en daarvoor middelen af te nemen van de netten.

Rianne Janssen denkt dat er eerst een openbare aanbesteding hiervoor moet zijn.

Lieven Boeve zou daar dan aan meedoen. Wat men zou verliezen, is een vertrouwd instrument voor kwaliteitsontwikkeling, dat dicht staat bij het leerplan en de taal van de school. Hij blijft erbij dat school- en leerplanontwikkeling minstens even belangrijk blijven als de door de minister gewenste meting op leerlingenniveau. Hij gelooft alvast dat de huidige goede samenwerking met het steunpunt bewaard zal blijven. Mede-eigenaarschap vanaf het begin zou alvast vergemakkelijken om het hele veld mee te krijgen en conflictstof te verwerken.

Patriek Delbaere wijst erop dat die keuze de overheid ook meer geld zal kosten. Daar komt bij dat dergelijke proeven slechts werken als alle scholen meedoen, zoals thans het geval is. De deelname aan paralleltoetsen is veel lager dan bij de OVSG-toets. De genoemde percentages tonen dat.

Hij waarschuwt ervoor de twee verschillende systemen niet bij elkaar te gooien. Peilingstoetsen zijn een element om wetenschappelijk in kaart te brengen of het onderwijssysteem met de vele middelen die het krijgt, zijn opdracht goed uitvoert. Tegelijk blijft het zinvol om in te zetten op interne kwaliteitszorg, wat een andere finaliteit is, waarvoor toetsen zijn ontwikkeld. Dat verloopt professioneel, maar hij erkent wel dat het niet altijd even wetenschappelijk ondersteund is, bij gebrek aan middelen. De spreker pleit daarom voor de investering in de wetenschappelijke waardering van beide en voor samenwerking. Beide fuseren, zoals in Nederland, blijkt te leiden tot betere resultaten.

Rianne Janssen pareert de vraag met een wedervraag: wat valt er te winnen bij een dergelijke beslissing? Die winst is alvast niet onvoorwaardelijk. De sterkte van het steunpunt is zijn methodologische expertise, maar het kan niets zonder de medewerking van scholen en onderwijsveld. Zij suggereert om niet overhaast te werken maar met een stappenplan te evolueren naar het ideaal van de leerwinst-meting.

Steven De Laet vult aan dat de toets weliswaar zou blijven, maar dat de vervlechting met de begeleiding verloren zou gaan. Pedagogisch begeleiders getuigen dat ze beter worden door mee aan toetsontwikkeling te doen. Ze moeten daarvoor immers in de leerplannen duiken en vanuit verschillende perspectieven naar onderwerpen kijken. Toetsontwikkeling professionaliseert met andere woorden de begeleiding. Bovendien worden zij daardoor ook ambassadeurs van de toetsen.

Koen Daniëls merkt op dat evalueren en controleren niet als verloren tijd mag bekeken worden, want er valt veel uit te leren door zowel leerling als school. Hij acht een reclamecampagne denkbaar die de deelname aan de parallelproeven tot 99 procent van de scholen opdrijft. Hij gelooft in de netwerkkracht.

Doel moet zijn om via analyse alles eruit te halen wat erin zit: op systeem- en schoolniveau, maar ook op leerlingenniveau. Verder herhaalt hij zijn vraag waarom geen extra punten worden toegekend voor excelleren. Voor de bescherming van de privacy van toetsenresultaten bestaan allerlei vormen van versleuteling. Hij pleit ervoor de GDPR niet als paraplu tegen analyse te gebruiken.

Tot slot roept hij de vraag op hoe de al lang bestaande en geprezen toetsen niet aan het licht brachten dat het zo slecht ging met begrijpend lezen. Ook sommige resultaten van de peilingsproeven voor Frans waren desastreus, terwijl daarbij evenmin een rode lamp was gaan branden in de toetsen van de koepels. Als dat ligt aan het feit dat ze gebaseerd zijn op de leerplannen, zegt dat ook iets daarover. Klopt dat verband?

Lieven Boeve repliceert dat in evalueren en controleren ook eren zit. Men moet dat niet bevoogdend doen maar met respect voor de schoolontwikkeling en de fierheid om daaraan te werken. Hij pleit er andermaal voor om als consortium van mede-eigenaars te werken aan toetsen die meerdere doelen dienen. Voor zijn netwerkorganisatie zat de eigendom van gegevens al voor de GDPR bij de school, wat niet verhindert dat de gegevens gebruikt worden op systemisch niveau. Individuele scholen worden alleen op eigen verzoek vergeleken.

Er is wel een evolutie aan de gang in het functioneren en het bewustzijn van de proeven. Nieuw is dat ze nu als proces worden bekeken, tussen het vierde en het zesde leerjaar, en dat begrijpend lezen systematischer wordt opgevolgd. Een hele set van eindtermen toetsen, is iets anders dan werken volgens een signaalfunctie, met focus op bepaalde elementen, zoals IDP tot nu toe deed. Dat de toetsen niet zagen wat de peilingsproeven wel aan het licht brachten, heeft dus te maken met hun format en opzet, besluit hij. Daar is echter uit geleerd.

Steven De Laet waarschuwt dat controleren door het leven of de wereld in cijfers te vatten, tot misbruik kan leiden. Onderzoek heeft veelvuldig aangetoond dat men op die manier immers gedrag oproept om goede cijfers te halen. Voor excelleren worden geen extra punten gegeven om niet het vreemde signaal te geven dat een leerling die alle eindtermen haalt, toch een onvoldoende zou krijgen. De punten drukken louter het beheersen van eindtermen uit, maar zij worden verbaal aangevuld om geen glazen plafond te creëren. Men stelt vast dat scholen die aanpak nu ook buiten de toetsen hanteren, wat positief is.

Rianne Janssen had het, wat de GDPR betreft, alleen over het schoolniveau. Zij merkt aanvullend op dat de ouders toestemming moeten geven voor de deelname aan de toetsen, al verkreeg het steunpunt de uitzondering dat het mag werken met active opt-out, dat echter evenzeer voor complicaties kan zorgen in de toetsen die worden vooropgesteld. Zij herhaalt dat de koepelorganisaties de achteruitgang niet hadden voorspeld omdat in de meetschaal geen constantie was ingebouwd.

Patriek Delbaere wijst erop dat de nieuwe eindtermen Frans zeer concreet geformuleerd zijn. De resultaten van peilingsproeven zeggen niet alleen iets over leerplannen maar ook over de kwaliteit van de eindtermen en de mate waarin leerplannen daar nog iets aan kunnen veranderen. Volgens hem is er sprake van een veelheid van oorzaken. De aan de gang zijnde herschrijving van de eindtermen acht hij dan ook een opportuniteit.

Elias Hemelsoet wijst erop dat men er in de vergelijking van de toetsen met PIRLS, rekening mee moet houden dat in die laatste tientallen miljoenen zijn geïnvesteerd. De koepels proberen naar beste vermogen en met beperkte capaciteit toetsen te ontwikkelen in functie van kwaliteitsversterking en niet in de eerste plaats van het detecteren van dergelijke signalen, al wordt daar nu versterkt op ingezet. Ook OVSG vraagt ondersteuning om die rol mee waar te maken.

Daarnaast zorgt het ontbreken van een vast criterium voor terughoudendheid bij het trekken van conclusies. Het omgekeerde zou methodologisch onjuist zijn. Zo wordt op dit moment vastgesteld dat de resultaten voor meten drastisch dalen. Ook hier wordt een probleem vermoed, wat bij gebrek aan vaste normering echter niet hard te maken is. OVSG wil voortvarende conclusies vermijden maar wenst een kader om zijn toetsen verder uit te bouwen en te professionaliseren.

Roosmarijn Beckers merkt in dit verband op dat de vermindering van de middelen voor de pedagogische begeleidingsdiensten misschien geen goed idee was.

Koen Daniëls leidt uit het voorbeeld over Frans af dat het probleem niet door de wijziging van de eindtermen geïnstalleerd kan zijn, omdat de peilingsproeven gebaseerd zijn op de leerplannen en niet enkel op de eindtermen.

Rianne Janssen preciseert dat de peilingsproeven niet gebaseerd zijn op de leerplannen maar op de eindtermen.

Koen Daniëls is blij dat leerplannen niet louter de eindtermen herhalen, en dat scholen niet enkel werken aan de eindtermen. Dat zijn immers minima. Het is dan ook cruciaal dat men ervoor zorgt dat naar meer wordt gestreefd. Hij suggereert om voor wie beter presteert dan de eindtermen een aparte scoretabel te maken, om dat zichtbaar te maken op alle niveaus. Het is immers ook interessant om te zien welke scholen er wel en niet in slagen om verder te raken dan de eindtermen. De leraren en de teams die daarin slagen, mogen daar best van genieten.

Rianne Janssen bevestigt dat PIRLS en PISA een veel breder bereik van de vaardigheid meten dan de peilingstoetsen.

Koen Daniëls erkent dat het belangrijk is dat de overheid daarmee achterhaalt waar men staat ten aanzien van de eindtermen. Het ruimere plaatje is echter minstens even belangrijk.

Voor *Lieven Boeve* is de vraag of de eindtermen inderdaad het minimum minimum vormen. Voor de nieuwe set moet dat nog onderzocht worden. Hij bevestigt nogmaals hoe belangrijk het is dat die goed zitten en dat het onderzoek daarnaar ook een van de oorspronkelijke doelen van de peilingsproeven was.

Het toetsen van leerplandoelen betreft niet alleen excelleren. Naast verdieping zorgen de plannen ook voor verbreding. Als de toetsen focussen op andere zaken dan het louter eindtermgerelateerde, gaat het ook daarover.

Onderzoek van de KU Leuven over PIRLS en TIMSS leert dat Katholiek Onderwijs Vlaanderen meer dan andere netten achteruit is gegaan inzake excelleren. Een van de hypothesen is dat het net te veel naar de eindtermen is gaan kijken, omdat de hele onderwijsomgeving dat doet. Veel scholen krijgen van de inspectie te horen dat ze te veel vragen. Die bedoelt daarmee allicht dat men onvoldoende peilt of het niveau van de eindtermen gehaald wordt en wat daarboven uitsteekt. Bij leerkrachten komt dat echter slecht over. Dat men de lat te hoog zou leggen, werkt demotiverend. Eindtermen louter zien als de lat waarover iedereen moet, is door de overheid georganiseerde nivellering. Communicatie is dan ook zeer belangrijk, concludeert hij.

Karolien GROSEMANS,
voorzitter

Loes VANDROMME
Jan LAEREMANS,
verslaggevers

Gebruikte afkortingen

ADHD	attention deficit hyperactivity disorder
aso	algemeen secundair onderwijs
CIPO	context, input, proces, output
CLB	centrum voor leerlingenbegeleiding
GDPR	General Data Protection Regulation (= algemene verordening gegevensbescherming, AVG)
GO!	onderwijs van de Vlaamse Gemeenschap
IDP	interdiocesane proeven
ikz	integrale kwaliteitszorg
KU Leuven	Katholieke Universiteit Leuven
OESO	Organisatie voor Economische Samenwerking en Ontwikkeling
OKI	onderwijskansarmoede-indicator
OVSG	Onderwijsvereniging van Steden en Gemeenten
PBD	pedagogische begeleidingsdienst
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment (van de OESO)
ROK	Referentiekader voor Onderwijskwaliteit
SALTO	Screeningsinstrument Aanvang Lager Onderwijs Taalvaardigheid
SES	socio-economische status
STEP	Steunpunt voor Toetsontwikkeling en Peilingen
TIMSS	Trends in International Mathematics and Science Study
tso	technisch secundair onderwijs
Vlor	Vlaamse Onderwijsraad
vte	voltijdsequivalent