



Centrum voor Onderwijseffectiviteit & -evaluatie
Faculteit Psychologie en Pedagogische Wetenschappen



Onderzoeksgroep EduBRON
Faculteit Sociale Wetenschappen

Haalbaarheidsstudie gecentraliseerde proeven in Vlaanderen

Perceel 2: Organisatorisch Eindrapport

Projectmedewerkers:

Gudrun Vanlaar en Jerich Faddar

Promotoren:

Rianne Janssen, Jan Vanhoof, Jonas Dockx en Koen Aesaert

11 februari 2021

INHOUDSOPGAVE

1. INLEIDING	7
1.1. SITUERING	7
1.2. HAALBAARHEIDSSTUDIE	7
1.3. VOORSTELLING PERCEEL 2	8
1.3.1. <i>Onderzoeksvragen</i>	8
1.3.2. <i>Aanpak</i>	8
1.3.3. <i>Doelstelling rapport</i>	9
1.4. OVERKOEPELENDE BEVINDINGEN	9
1.4.1. <i>Grijze literatuur</i>	9
1.4.2. <i>Samenhang tussen onderzoeksvragen</i>	10
1.4.3. <i>What's at stake?</i>	10
1.4.4. <i>Aanpak budgetramingen</i>	10
2. BREDE AFNAME	12
2.1. INLEIDING	12
2.1.1. <i>Omschrijving van het begrip brede afname</i>	12
2.1.2. <i>Korte voorstelling van mogelijke toetsaanpassingen</i>	13
2.2. ARGUMENTEN VOOR EEN BREDE AFNAME	15
2.2.1. <i>Juridische argumenten rond inclusie</i>	15
2.2.2. <i>Onderwijskundige argumenten voor een brede afname</i>	16
2.2.3. <i>Toetstechnische argumenten voor toetsaanpassingen</i>	17
Inhoudsvaliditeit en construct irrelevante variantie	17
Differentiële effectiviteit	17
Billijkheid op leerling- én schoolniveau	18
Discussie	18
2.3. BREDE AFNAME: LEERLINGEN	19
2.3.1. <i>Overzicht</i>	19
2.3.2. <i>Over welke leerlingen gaat het?</i>	19
Zorgcontinuüm	20
Brede basiszorg (Fase 0)	20
Verhoogde zorg (Fase 1)	20
Uitbreiding van de zorg (Fase 2)	20
Individueel Aangepast Curriculum (Fase 3)	20
Besluit	21
Speciale onderwijsleermiddelen	21
Implicaties voor de gecentraliseerde toetsen	22
Mogelijke vormen van afname	22
Aantallen leerlingen?	22

Speciale onderwijsleermiddelen.....	22
(Gemotiveerd) verslag	23
Compenserende en dispenserende maatregelen op schoolniveau.....	24
Andere onderwijssettings.....	24
Privacy.....	24
2.3.3. Scenario's inzake brede afname bij leerlingen.....	24
Toetsontwikkeling	25
Universal Design for Assessments.....	25
Oplijsten en coderen van itemkenmerken.....	26
Toegankelijk maken van items.....	26
Toekenningsbeleid	27
(Deel)vrijstellingen.....	27
Voor welke leerlingen?	27
Leerlingen met een IAC	27
Leerlingen in het onthaalonderwijs voor anderstalige kinderen (OKAN).....	27
Opt out?	28
Welke (deel)vrijstellingen?	28
Scenario 1: Op alle toetsen	28
Scenario 2: (Deel)vrijstellingen op inhoudelijke gronden.....	28
Scenario 2a: (Deel)vrijstellingen op itemniveau.....	29
Scenario 2b: (Deel)vrijstellingen op toetsniveau.....	29
Bijkomende overwegingen	29
Toetsaanpassingen	30
Voor welke leerlingen?	30
Scenario 1: voor alle leerlingen	30
Scenario 2: voor welomschreven leerlinggroepen	31
Welke toetsaanpassingen?	32
Speciale onderwijsleermiddelen.....	33
Aanvraagplatform	34
Toekenningsprocedure.....	34
Drie scenario's.....	34
Verantwoordelijkheid school.....	34
In overleg met CLB.....	35
Gunningscommissie.....	35
Buitenlandse voorbeelden	35
Koppeling met toetsafnameplatform.....	36
2.4. BREDE AFNAME: ONDERWIJSSETTINGS.....	36
2.4.1. Overzicht van de onderwijssettings	36
Leerlingenaantallen per afnamemoment.....	36
Inleiding	36
Reguliere onderwijs	36
Buitengewoon onderwijs	38

Huisonderwijs	39
Over afnamemomenten heen	39
2.4.2. <i>Scenario's met betrekking tot onderwijssettings</i>	39
Regulier onderwijs	39
Buitengewoon onderwijs	40
Huis- en privé-onderwijs	41
2.4.3. <i>Bijkomende overwegingen</i>	42
2.5. OVERZICHT DEELSCENARIO'S EN BELEIDSADVIEZEN	43
2.5.1. <i>Overzicht deelscenario's</i>	43
2.5.2. <i>Beleidsadviezen</i>	44
3. TIJDSCENARIO'S	46
3.1. INLEIDING	46
3.2. PLANNING AFNAME(S)	49
3.2.1. <i>Planning in het schooljaar</i>	49
3.2.2. <i>Bijkomende afnames?</i>	50
Inhaalmoment	50
Herkansingsmoment	51
3.2.3. <i>Budgetimplicaties</i>	51
3.3. AFNAMEVENSTER	52
3.3.1. <i>Scenario's</i>	52
3.3.2. <i>Budgetimplicaties</i>	54
3.4. DOORLOOPTIJD	54
3.4.1. <i>Scenario's</i>	55
3.4.2. <i>Budgetimplicaties</i>	57
3.5. OVERZICHT DEELSCENARIO'S EN BELEIDSADVIEZEN	57
3.5.1. <i>Overzicht deelscenario's</i>	57
3.5.2. <i>Beleidsadviezen</i>	57
4. UITVOERING AFNAME	59
4.1. INLEIDING	59
4.2. TOETSLOCATIE EN INFRASTRUCTUUR	59
4.2.1. <i>Buitenschoolse toetslocaties</i>	59
4.2.2. <i>De eigen school als toetslocatie</i>	60
4.3. TAKEN EN ROLLEN	61
4.3.1. <i>Schoolniveau</i>	61
Toetscoördinator	62
Vorbereidingsfase	62
Tijdens de toetsafname	62
Na de toetsafname	63

Toetsassistenten en toetsbegeleiders.....	64
ICT-ondersteuning.....	64
Toetsinformatiemakelaar.....	65
4.3.2. Centrale niveau.....	65
4.4. VERDIEPING VAN SCENARIO'S M.B.T. TOETSASSISTENTEN EN KWALITEITSMONITORING.....	66
4.4.1. Mogelijke scenario's m.b.t. Recruitering van toetsassistenten.....	66
Scenario 1: Extern aangeworven toetsassistenten.....	66
Scenario 2: Leraren, extern aan de school.....	68
Scenario 3: Leraren, intern aan de school.....	68
4.4.2. Mogelijke scenario's m.b.t. kwaliteitsmonitoring.....	68
Scenario 1: interne kwaliteitsmonitoring.....	68
Scenario 2: Steekproefsgewijze externe kwaliteitsmonitoring.....	69
4.5. BIJKOMENDE PROBLEMATIEK: OPEN VRAGEN.....	70
4.6. TOETSVERVERSING.....	71
4.6.1. Concepten m.b.t. toetsverversing.....	71
Omschrijving van het begrip toetsverversing.....	71
Confidentialiteit van items: een categorisch en een continu perspectief.....	71
Categorische visie.....	71
Continue visie.....	72
Het belang van de confidentialiteit van de items.....	73
4.6.2. Praktische aspecten van de afname m.b.t. toetsverversing.....	75
Toetsdoel.....	75
Mogelijke beschermingsmaatregelen.....	76
Besluit.....	80
4.6.3. Scenario's m.b.t. de confidentialiteit van items en itemverversing.....	80
Scenario 1: 'low-stakes' toetsen met toetsafname op school.....	80
Scenario 2: 'high-stakes' toetsen voor scholen met de afname op school.....	81
Scenario 3: 'high-stakes' toetsen voor leerlingen met de afname in een toetscentrum.....	81
4.7. OVERZICHT DEELSCENARIO'S EN BELEIDSADVIEZEN.....	82
4.7.1. Overzicht deelscenario's.....	82
4.7.2. Beleidsadviezen.....	82
5. BELEIDSSAMENVATTING EN -ADVIEZEN.....	84
5.1. VOORSTELLING PERCEEL 2.....	84
5.1.1. Drie deeldomeinen met eigen onderzoeksvragen.....	84
5.1.2. Aanpak.....	84
5.1.3. Samenhang met de andere percelen van de haalbaarheidsstudie.....	85
5.1.4. What's at stake?.....	85
5.2. BREDE AFNAME.....	85
5.2.1. Overzicht deelscenario's.....	85

5.2.2. <i>Beleidsadviezen</i>	86
5.3. TIJDSCENARIO'S.....	88
5.3.1. <i>Overzicht deelscenario's</i>	88
5.3.2. <i>Beleidsadviezen</i>	88
5.4. UITVOERING AFNAME	89
5.4.1. <i>Overzicht deelscenario's</i>	89
5.4.2. <i>Beleidsadviezen</i>	90
6. REFERENTIES	91
7. BIJLAGE 1: ONDERZOEKSVRAGEN PERCEEL 2 VOLGENS DE OPROEP VOOR DE HAALBAARHEIDSTUDIE	96
8. BIJLAGE 2: VERSLAG FOCUSGESPREK STEP	97

1. INLEIDING

1.1. SITUERING

Het Vlaamse regeerakkoord voor de periode 2019-2024 stelt dat het Vlaamse onderwijs altijd tot de Europese top heeft behoord. De Vlaamse regering wil dat behouden: “Met regelmatige, gestandaardiseerde, genormeerde en gevalideerde net- en koepeloverschrijdende proeven en aangescherpte eindtermen kunnen we leerlingen en ouders de garantie geven dat leerlingen ook daadwerkelijk over de nodige kennis en competenties zullen beschikken die bij dat diploma horen, ongeacht de school” (Departement Kanselarij en Bestuur, 2019, p. 21). Die proeven (hierna ‘gecentraliseerde toetsen’ genoemd) hebben als doel de vinger aan de pols te houden van de Vlaamse onderwijskwaliteit en maken deel uit van het systeem van externe kwaliteitszorg. Meer specifiek wil de overheid met de toetsen de leerwinst van jongeren en scholen in het leerplichtonderwijs in kaart brengen alsook de mate waarin leerlingen de eindtermen bereiken (Departement Kanselarij en Bestuur, 2019). Daarnaast hoopt men dat de toetsresultaten door scholen als input gebruikt worden voor hun interne kwaliteitszorg. Er werd beslist te starten met de ontwikkeling van digitale toetsen Nederlands en wiskunde, terwijl andere vakken/leergebieden in de toekomst overwogen zullen worden. De gecentraliseerde toetsen zullen naar alle waarschijnlijkheid worden afgenomen op het einde van het vierde en zesde leerjaar van het lager onderwijs en op het einde van de eerste graad en in het tweede leerjaar van de derde graad in het secundair onderwijs. De concrete timing hangt samen met de uitrol van de nieuwe onderwijsdoelen in het Vlaamse onderwijs.

1.2. HAALBAARHEIDSTUDIE

Hoewel Vlaanderen via het peilingsonderzoek reeds twee decennia ervaring heeft met de ontwikkeling en afname van gestandaardiseerde toetsen die door de overheid en het Steunpunt Toetsontwikkeling en Peilingen (STEP) worden georganiseerd, is de omvang, complexiteit en focus op de nieuwe eindtermen van de toekomstige gecentraliseerde toetsen een nieuw gegeven. In een apart, kortlopend onderzoeksproject wordt de haalbaarheid van het uitrollen van een dergelijk systeem van gecentraliseerde toetsen onderzocht. Deze haalbaarheidsstudie werd onderverdeeld in drie percelen: (1) een pedagogisch en psychometrisch luik; (2) een praktisch-organisatorisch luik en (3) een juridisch en (data)technisch luik. Dit rapport vormt de rapportage van het tweede perceel dat zich specifiek richt op het uitwerken van concrete, mogelijke scenario’s voor de praktisch-organisatorische aspecten van de implementatie van de gecentraliseerde toetsen. Daarbij ligt de klemtoon op één afnamejaar van de gecentraliseerde toetsen, al wordt er doorheen het rapport ook verwezen naar organisatorische kwesties in verband met de longitudinale opvolging van leerlingen en scholen over afnamejaren heen.

1.3. VOORSTELLING PERCEEL 2

1.3.1. ONDERZOEKSVRAGEN

Bijlage 1 bevat de lijst van de onderzoeksvragen van Perceel 2 zoals aangegeven in de oproep voor de haalbaarheidsstudie. Deze onderzoeksvragen worden met de oorspronkelijke nummering in Tabel 1.1 verkort weergegeven en gehergroepeerd in drie deeldomeinen. Ten eerste worden in dit perceel van de haalbaarheidsstudie scenario's uitgewerkt inzake de *brede afname* van de toetsen: kunnen ook leerlingen met specifieke onderwijsbehoeften deelnemen aan de toetsafnames en hoe kan dat kwaliteitsvol gerealiseerd worden zonder dat de standaardisatie in het gedrang komt (Vraag 18)? Ten tweede worden verschillende *tijdsscenario's* uitgewerkt. Dit gaat over de vragen in welke periode in het schooljaar (Vraag 22) en met welk afnamevenster de gecentraliseerde toetsen worden afgenomen (Vraag 21). Daarnaast wordt uitgewerkt welke scenario's er mogelijk zijn inzake de doorlooptijd tussen het afnemen van de toetsen en het bekomen van de resultaten (Vraag 23). Ten derde worden scenario's ontwikkeld inzake de afnamemodaliteiten, i.e., de *uitvoering van de afname*. Daarbij worden de vragen beantwoord wie betrokken is bij uitvoering van de afname en welke verantwoordelijkheden deze actoren daarbij opnemen (Vragen 19 en 20) en hoe de vertrouwelijkheid van de toetsopgaven bewaakt kan worden (Vraag 17).

Tabel 1.1

Overzicht van de onderzoeksvragen van Perceel 2 binnen de drie onderscheiden deeldomeinen

Deeldomein Vraagnummer	Onderzoeksvraag
<i>Brede afname</i>	
18	Redelijke aanpassingen en eventuele vrijstellingen voor specifieke leerlingen
<i>Tijd</i>	
21	Afnamevenster
22	Planning afname in het schooljaar
23	Doorlooptijd tussen toetsafname en bekomen van de resultaten
<i>Uitvoering afname</i>	
19	Actoren betrokken bij de afname en hun rol
20	Autonomie en verantwoordelijkheid scholen
17	Bewaken confidentialiteit items en nood aan toetsverversing

1.3.2. AANPAK

De scenario's voor de onderzoeksvragen binnen de drie deeldomeinen van Perceel 2 worden ontwikkeld op basis van een inhoudelijke analyse van de vragen waarbij de literatuur over binnen- en buitenlandse praktijken inspirerend is. Anders dan bij Perceel 1, waar theoretische inzichten en methodologische referentiekaders het vertrekpunt zijn, of dan bij Perceel 3, waar een juridisch kader en de ICT-mogelijkheden richtinggevend zijn, zijn het voor de organisatorische vragen in Perceel 2 immers vooral goede praktijkvoorbeelden en de specifieke Vlaamse onderwijscontext die inspiratie kunnen bieden voor de implementatie van de gecentraliseerde toetsen. Een schematisch overzicht van deze praktijkvoorbeelden wordt gegeven in de inleiding van Hoofdstuk 3. Daarnaast werden ook de algemene richtlijnen geraadpleegd die door professionele toetsorganisaties werden ontwikkeld, zoals o.a. de

‘Standards for Educational and Psychological Testing’ (AERA, APA, NCME, 2014) en het ‘RCEC Beoordelingssysteem voor de kwaliteit van studietoetsen en examens’ (Sanders et al., 2016).

Tussentijds werd buiten de stuurgroepvergaderingen op geregelde basis overleg gepleegd met de gunninghouders van Percelen 1 en 3. Enerzijds werd daarin gesteld dat er een sterke samenhang is in de scenario-ontwikkeling tussen de percelen. Anderzijds werd eveneens vastgesteld dat de percelen ook voldoende afgebakend zijn om elk op zich vanuit een bepaald perspectief scenario’s over de aspecten van de gecentraliseerde toetsen uit te werken.

1.3.3. DOELSTELLING RAPPORT

In dit rapport worden voor de verschillende onderzoeksvragen (deel)scenario’s descriptief (en niet prescriptief) uitgewerkt, waarbij ook de kwaliteiten en beperkingen van elk (deel)scenario worden besproken. Waar mogelijk, worden daarbij ook voorzichtige budgetramingen opgesteld. Doorheen de volgende drie hoofdstukken worden zo de algemene plannen voor het inrichten van gecentraliseerde toetsen in Vlaanderen vanuit organisatorisch perspectief geconcretiseerd. Elk hoofdstuk eindigt daarbij met een oplistings van de beleidskeuzes die binnen het besproken deeldomein moeten worden gemaakt voor de implementatie van de gecentraliseerde toetsen in Vlaanderen. We geven daarbij ook beleidsadviezen. Het is aan de onderwijsminister om op basis van de gegeven informatie deze keuzes te maken. In het slothoofdstuk worden een aantal zaken uit de tekst hernomen om te komen tot een beleidssamenvatting met daarin ook de beleidsadviezen.

1.4. OVERKOEPELENDE BEVINDINGEN

Alvorens de onderzoeksvragen van de drie deeldomeinen te bespreken, worden eerst enkele overkoepelende bedenkingen geformuleerd die bij het uitwerken van de deelscenario’s aan de oppervlakte kwamen.

1.4.1. GRIJZE LITERATUUR

In tegenstelling tot wat we aanvankelijk dachten, was het niet eenvoudig om toegang te krijgen tot goede (buitenlandse) praktijkvoorbeelden inzake de organisatie van grootschalige toetsafnames. De informatie hierover is voornamelijk te vinden in de zogenaamde ‘grijze literatuur’. In wetenschappelijke tijdschriften wordt hierover weinig gerapporteerd. Om een overzicht te krijgen van welke buitenlandse overheden grootschalige nationale toetsen inzetten, werd een beroep gedaan op het overzichtsrapport van de OESO (OECD, 2013) over evaluatiepraktijken in de verschillende landen van de OESO. Daarnaast werd ook gebruik gemaakt van het eindrapport van OBPWO-13.03 (Vanhoof e.a., 2015), het bijhorende boek (Penninckx e.a., 2017) en van enkele overzichtsartikelen rond longitudinale datasystemen in het onderwijs

(González-Sancho & Vincent-Lancrin, 2016; OECD, 2010)¹. Op basis van deze informatie werd op het internet gezocht naar meer concrete informatie in nationale rapporten, die omwille van hun taal en het interne karakter ervan niet altijd toegankelijk waren. Ook was de organisatorische kant in de meeste van deze rapporten te onderbelicht om daaruit concrete voorstellen voor de Vlaamse gecentraliseerde toetsen te kunnen distilleren. Desalniettemin worden buitenlandse praktijken in de tekst aangehaald als voorbeeld van bepaalde deelscenario's of om bepaalde problematieken te illustreren.

1.4.2. SAMENHANG TUSSEN ONDERZOEKSVRAGEN

Niet alleen tussen de percelen, maar ook tussen de onderzoeksvragen van Perceel 2 is er een sterke samenhang. Dit zal zeker blijken in het eindrapport. Bijvoorbeeld, de breedte van het afnamevenster heeft repercussies op de mogelijkheden om toetsassistenten in te zetten en op het inbouwen van garanties voor het bewaken van toetsconfidentialiteit. Toch wordt in deze studie gepoogd om elke onderzoeksvraag op zich te beantwoorden aan de hand van specifieke deelscenario's. De (in)compatibiliteit tussen deze deelscenario's zal duidelijk blijken wanneer de overheid op basis hiervan het overkoepelende scenario voor de implementatie van de gecentraliseerde toetsen zal vastleggen.

1.4.3. WHAT'S AT STAKE?

Een ander probleem waarmee we geconfronteerd werden bij de uitwerking van de deelscenario's over de organisatie van de gecentraliseerde toetsen is dat (deel)scenario's minstens deels samenhangen met de vraag wat het precieze doel is van de gecentraliseerde toetsen. Om het met een boutade te stellen: het is moeilijk om wie-, waar-, wanneer- en hoe-vragen te beantwoorden los van de waarom-vraag. Ook al zijn de krijtlijnen van het doel van de gecentraliseerde toetsen omschreven, toch zijn er nog interpretaties mogelijk, bijvoorbeeld of de gecentraliseerde toetsen overheersend ondersteunings- of verantwoordingsgericht zijn of in welke mate de toetsen voor leerlingen en leerkrachten 'high stakes' zijn. Ook in Percelen 1 en 3 wordt de vraag gesteld naar wat nu precies 'op het spel' zal staan bij de gecentraliseerde toetsen. Net als in Perceel 1 wijzen we er alvast op dat het niet evident is om verschillende toetsdoelen te combineren en dat de perceptie van de toetsfinaliteit zal bepalen hoe verschillende actoren in het onderwijsveld met de gecentraliseerde toetsen omgaan.

1.4.4. AANPAK BUDGETRAMINGEN

Bij de verschillende deelscenario's worden indien mogelijk inschattingen gemaakt van de bijkomende budgettaire kost ervan. Deze schattingen zijn benaderend. Er wordt daarbij een onderscheid gemaakt tussen opstartkosten (bijvoorbeeld, de ontwikkeling van een portaal voor het aanvragen van (deel)vrijstellingen of toetsaanpassingen), recurrente kosten en bijkomende onderzoeksopdrachten die

¹ Andere interessante naslagwerken over grootschalige (inter)nationale toetsafnames zijn Blömeke en Gustafsson (2017), Harju-Luukkainen, McElvany, en Stang (2020), Rutkowski, von Davier, en Rutkowski (2014) en Wagemakers (2020).

mogelijk (deels) binnen het nieuwe steunpunt gecentraliseerde toetsen kunnen worden opgenomen. De recurrente kosten zijn enerzijds afhankelijk van het aantal leerlingen of scholen (met een geschatte kost per eenheid) en verwijzen anderzijds naar overkoepelende kosten (bijvoorbeeld een organisatiekost). De meeste kosten worden begroot in de benodigde voltijdsequivalenten (VTE) voor een periode. Voor de kosten van extra toetsontwikkeling of de ontsluiting van een digitaal toetsplatform wordt verder doorverwezen naar de rapporten van Percelen 1 en 3. Voor erg complexe scenario's wordt afgezien van een financiële berekening omdat de omvang van het scenario op zich al voldoende suggereert dat het weinig haalbaar is (zie bijvoorbeeld het deel over externe toetsassistenten).

2. BREDE AFNAME

2.1. INLEIDING

Alvorens in te gaan op de scenario's voor de brede afname van de gecentraliseerde toetsen wordt eerst een kort overzicht gegeven van het begrip brede afname en welke toetsaanpassingen mogelijk zijn. Daarna worden juridische, onderwijskundige en toetstechnische argumenten besproken rond verschillende aspecten van een brede afname. Deze argumenten kunnen worden aangehaald ter ondersteuning van een brede afname, maar wijzen tevens op de complexiteit van de beslissingen die bij de implementatie ervan moeten worden genomen.

2.1.1. OMSCHRIJVING VAN HET BEGRIP BREDE AFNAME

In principe worden de gecentraliseerde toetsen voor een bepaalde doelgroep op een gestandaardiseerde wijze afgenomen bij *alle* leerlingen van het betreffende leerjaar. Binnen het onderzoek naar de brede afname van de toetsen wordt daarbij de vraag gesteld in welke mate en op welke manier dat gerealiseerd kan worden. Vanuit het standpunt van de leerlingen zijn er immers bepaalde leerlinggroepen voor wie de afname minder evident is. Het kan daarbij gaan over leerlingen die vanuit hun specifieke onderwijsbehoeften in de klaspraktijk een beroep doen op 'redelijke aanpassingen' (bijvoorbeeld, een leerling met dyslexie gebruikt voorleessoftware) of die speciale onderwijsleermiddelen nodig hebben (bijvoorbeeld een brailregel voor een blinde leerling). Daarnaast zijn er in het reguliere onderwijs ook leerlingen die een individueel aangepast curriculum volgen en daardoor onderdelen van de getoetste leerstof van de gecentraliseerde toetsen niet onderwezen krijgen.

Verschillen tussen leerlingen in hun specifieke onderwijsbehoeften, in hun nood aan speciale leermiddelen en in het gevolgde curriculum maken dus dat 'testing accommodations' zich opdringen. Zo kunnen bepaalde leerlingen (deels) vrijgesteld worden van de gecentraliseerde toetsen of kan een – allicht grotere – groep leerlingen deze toetsen afleggen in aangepaste omstandigheden, zoals bijvoorbeeld met extra hulpmiddelen. Deze laatste mogelijkheid wordt in het Engels aangeduid als 'test modifications', wat in deze tekst vertaald wordt als toetsaanpassingen. We hebben deze term verkozen om zo het verschil te benadrukken met de 'redelijke aanpassingen' (zoals compenseren, remediëren, differentiëren of dispensereren) die reguliere scholen voorzien in hun onderwijs aan leerlingen met specifieke onderwijsbehoeften. Allicht komen leerlingen met redelijke aanpassingen ook in aanmerking voor toetsaanpassingen, maar in de verschillende scenario's van een brede afname is er niet per se een perfecte overeenstemming tussen beiden.

Samengevat, een brede afname van de gecentraliseerde toetsen vraagt dus naar scenario's over welke toetsaanpassingen en (deel)vrijstellingen toegelaten kunnen worden, hoe ze in de toetsomgeving kunnen worden geïmplementeerd en op welke manier ze aan individuele leerlingen kunnen worden gegund. Naast de organisatie van de brede afname voor de leerlingen wordt in dit hoofdstuk ook stilgestaan bij de vraag of de gecentraliseerde toetsen ook in onderwijssettings buiten de reguliere scholen afgenomen

zouden worden. Het gaat hier o.a. om het buitengewoon onderwijs en het huisonderwijs. Ook hiervoor worden in dit hoofdstuk scenario's voorgesteld.

2.1.2. KORTE VOORSTELLING VAN MOGELIJKE TOETSAANPASSINGEN

In het kader van een brede afname kan worden afgeweken van een uniforme toets (wegens vrijstellingen) of van een uniforme toetsafname (wegens toetsaanpassingen) bij alle leerlingen. Om een overzicht te geven van mogelijke toetsaanpassingen worden hieronder twee classificatieschema's voorgesteld. Merk op dat er in beide classificaties niets veranderd wordt aan de inhoud van de toets. Wanneer leerlingen bepaalde items of bepaalde toetsen niet hoeven af te leggen, gaat het over vrijstellingen en niet over toetsaanpassingen.

Volgens Thurlow e.a. (2005) zijn aanpassingen aan een toets mogelijk op de volgende vijf aspecten:

- (a) timing: het geven van extra tijd bij een item of bij de gehele toets, of het inlassen van bijkomende rustpauzes;
- (b) respons: alternatieve opties voor het antwoorden op de toets (bijv. direct markeren in het testboekje en niet overbrengen op een antwoordblad, het antwoord luidop zeggen in plaats van noteren, ...);
- (c) setting: veranderingen in de testomgeving (bijvoorbeeld testen in een rustig lokaal, aangepast meubilair voor fysieke beperkingen, ...);
- (d) uitrusting: het gebruik van extra apparaten of hulpmiddelen (bijv. rekenmachines of formularium, brailleleesregel);
- (e) presentatie: aangepaste presentatie van toetsmaterialen (bijv. voorlezen van de vragen, groter lettertype).

Een ander classificatieschema voor toetsaanpassingen is gebaseerd op het onderscheid tussen digitale en lokale toetsaanpassingen bij digitale toetsen (Smarter Balanced Assessment Consortium, 2014). Een aantal toetsaanpassingen kunnen digitaal worden ingebed in de toetsapplicatie. Voorbeelden hiervan zijn onder andere extra tijd geven, pauzes inlassen, voorleessoftware, toegang tot een formularium, woordenboek, spellingcorrector, of rekenmachine, ondertiteling bij video's, een doventolk rechts onderaan een beeldfragment. Daarnaast kunnen ook lokale toetsaanpassingen worden voorzien. Dit zijn toetsaanpassingen die ter plekke in het afnamelokaal aan bepaalde leerlingen worden aangeboden. Een aantal van deze lokale toetsaanpassingen kunnen niet digitaal worden aangeboden. Het gaat bijvoorbeeld over het afleggen van de toets in een aparte of een stille ruimte of het toelaten van medicatie of voedingsmiddelen. Andere lokale toetsaanpassingen zijn een variant van digitale toetsaanpassingen, zoals het geven van extra tijd of het gebruik van een formularium, woordenboek of zakrekenmachine. De keuze om deze hulpmiddelen lokaal of digitaal aan te bieden raakt aan de principes van standaardisering en billijkheid. Als er bijvoorbeeld digitaal een formularium wordt aangeboden, dan zal dit naar alle waarschijnlijkheid voor alle leerlingen hetzelfde zijn, wat de standaardisering van de toets(afname) ten goede komt. Indien dit niet-digitaal gebeurt, zullen de aangeboden formularia meer aansluiten bij degene

Tabel 2.1

Voorbeelden van lokale en digitale toetsaanpassingen.

Omschrijving	Universele tools	Toetsfaciliteiten	Speciale onderwijsleermiddelen
Lokaal aangeboden			
Timing		- vroeger startuur; later einduur	
Respons			- writer (iemand die antwoorden van leerling registreert)
Setting		- stille ruimte; - aparte plaats in lokaal	- aangepast meubilair; - ondersteuner of begeleider; - medicatie, speciale voeding
Equipment	- meetlat / geodriehoek ... - gebruik van kladbladen	- woordenboek; - rekenmachine	- brailleleesregel
Presentatie			- braille; - doventolk
Digitaal voorzien in toetsapplicatie (kunnen aan en uit gezet worden)			
Timing		- extra pauzes; - extra tijd	
Respons			- speech to text software
Setting			
Equipment	- zoom / vergrootglas; - formularium; - woordenboek - markeerfunctie; - digitaal kladblad	- spellingcontrole; - rekenmachine-applicatie - formularium	
Presentatie	- makkelijke navigatie; - duidelijk leesbaar lettertype - keuze tussen lettertypes; - keuze kleurcontrast; - mogelijkheid om opmerkingen bij vraag te zetten of om vraag aan te kruisen of voorlopig over te slaan ...	- voorleessoftware; - mogelijkheid om extra uitleg op te roepen via extra muisklik indien vraag niet duidelijk is.	- bij mediafragmenten: in de hoek verschijnt er een doventolk op het scherm; - ondertiteling van videofragmenten

die de leerlingen gewoon zijn om te gebruiken. De billijkheid stijgt, maar de standaardisering daalt, omdat formularia kunnen verschillen tussen scholen.

In Tabel 2.1 worden voorbeelden gegeven van toetsaanpassingen uit buitenlandse praktijkvoorbeelden onderverdeeld volgens beide classificatieschema's. Daarbij wordt ook een verschil gemaakt tussen universele tools (toetsaanpassingen die bij een digitale afname in principe aan alle leerlingen kunnen worden aangeboden, zodat er voor bepaalde leerlingen met een lichte beperking verder geen

toetsaanpassingen nodig zijn²), toetsfaciliteiten (toetsaanpassingen die voor specifieke leerlinggroepen worden voorbehouden) en speciale onderwijsleermiddelen (die specifiek verwijzen naar de aanpassingen die leerlingen met een beperking nodig hebben). Deze verschillen worden verderop in de tekst nog verduidelijkt. Alvast kan hier worden opgemerkt dat sommige hulpmiddelen afhankelijk van het toetsafnamebeleid ofwel als universele tool ofwel als speciale faciliteit worden ingeschakeld (bijvoorbeeld een woordenboek of rekenmachine). Ook dit illustreert de complexiteit van de beslissingen over het gebruik van toetsaanpassingen.

2.2. ARGUMENTEN VOOR EEN BREDE AFNAME

2.2.1. JURIDISCHE ARGUMENTEN ROND INCLUSIE

In december 2006 ondertekende ons land het VN-verdrag voor personen met een beperking. Artikel 24 handelt over onderwijs, en stelt dat een kind met specifieke onderwijsbehoeften recht heeft op inclusief onderwijs op alle onderwijsniveaus.

Teneinde dit recht zonder discriminatie en op basis van gelijke kansen te verwezenlijken, waarborgen de ondertekenende landen een inclusief onderwijssysteem op alle niveaus “om zo alle kansen te bieden voor een volledige ontwikkeling van het menselijk potentieel en het gevoel van waardigheid en eigenwaarde en de versterking van de eerbiediging van mensenrechten, fundamentele vrijheden en de menselijke diversiteit”.

Bij de uitoefening van dit recht waarborgen de ondertekenende landen dat:

- a) kinderen met een handicap niet op grond van hun handicap worden uitgesloten van gratis en verplicht basisonderwijs of van het voortgezet onderwijs;
- b) personen met een handicap toegang hebben tot inclusief, hoogwaardig en gratis basisonderwijs en tot voortgezet onderwijs en wel op basis van gelijkheid met anderen in de gemeenschap waarin zij leven;
- c) redelijke aanpassingen worden verschaft naar gelang de behoefte van de persoon in kwestie;
- d) personen met een handicap, binnen het algemene onderwijssysteem, de ondersteuning ontvangen die zij nodig hebben om effectieve deelname aan het onderwijs te vergemakkelijken;
- e) doeltreffende, aan het individu aangepaste, ondersteunende maatregelen worden genomen in omgevingen waarin de cognitieve en sociale ontwikkeling wordt geoptimaliseerd, overeenkomstig het doel van onderwijs waarbij niemand wordt uitgesloten.

Bij de ondertekening engageerden de landen er zich toe om personen met een handicap in staat te stellen praktische en sociale vaardigheden op te doen, teneinde hun volledige deelname aan het onderwijs op voet van gelijkheid te vergemakkelijken. Daartoe nemen de landen passende maatregelen, waaronder:

- a) het vergemakkelijken van het leren van braille, alternatieve schrijfwijzen, het gebruik van ondersteunende en alternatieve communicatiemethoden, -middelen en -vormen, alsmede het opdoen van vaardigheden op het gebied van oriëntatie en mobiliteit en het vergemakkelijken van ondersteuning en begeleiding door lotgenoten;
- b) het leren van gebarentaal vergemakkelijken en de taalkundige identiteit van de gemeenschap van doven bevorderen;

² Zie ook Universal Design for Assessments verderop in de tekst.

- c) waarborgen dat het onderwijs voor leerlingen die blind, doof of doofblind zijn, plaatsvindt in de talen en met de communicatiemethoden en -middelen die het meest geschikt zijn voor de desbetreffende leerling en in een omgeving waarin hun cognitieve en sociale ontwikkeling worden geoptimaliseerd.

Om de uitoefening van dit recht te vergemakkelijken, nemen de ondertekenende landen passende maatregelen om leerkrachten aan te stellen, met inbegrip van leerkrachten met een beperking, die zijn opgeleid om te werken binnen een inclusieve onderwijscontext, en leidinggevenden en medewerkers op te leiden die op alle onderwijsniveaus werkzaam zijn. Bij deze opleiding moeten studenten en leraren worden getraind in het omgaan met leerlingen met een beperking en het gebruik van de desbetreffende ondersteunende communicatie en andere methoden, middelen en vormen van en voor communicatie, onderwijstechnieken en materialen om leerlingen met een beperking te ondersteunen.

Deze eisen uit het verdrag werden in Vlaanderen in 2008 opgenomen in Artikel 19 van het Gelijkekansendecreet. Om het recht op redelijke aanpassingen te waarborgen is er een duidelijke omkadering en wetgeving nodig inzake inclusief onderwijs, zoals dat nu geregeld wordt door het M-decreet sinds 1 september 2015.

Recht op deelname aan onderwijs houdt ook in dat leerlingen met een beperking recht hebben om deel te nemen aan toetsen of om geëvalueerd te worden en feedback te krijgen over hun leerproces. Volgens de juristen van Perceel 3 is het nodig dat de overheid een nieuw decreet over de gecentraliseerde toetsen opstelt. Het is belangrijk dat de brede afname van de gecentraliseerde toetsen daarin helder omschreven en afgebakend wordt om zo voor bepaalde leerlinggroepen het spanningsveld op te lossen tussen verplichte deelname versus recht op deelname. Daarbij stelt zich ook de vraag welke modaliteiten daarvoor tijdens de toetsafname moeten worden voorzien. Mogelijk zouden leerlingen met een beperking of hun ouders bijvoorbeeld van de overheid kunnen eisen dat er aangepaste toetsen worden voorzien. Deze juridische problematiek werd tijdens de haalbaarheidsstudie aan de juristen van Perceel 3 gesignaleerd.

2.2.2. ONDERWIJSKUNDIGE ARGUMENTEN VOOR EEN BREDE AFNAME

Vanuit onderwijskundige hoek zijn er twee grote argumenten om ook leerlingen met beperkingen of specifieke onderwijsbehoeften bij de afname van grootschalige toetsen te betrekken. Ten eerste is de (aangepaste) deelname van deze leerlingen cruciaal opdat scholen en leerkrachten zich verantwoordelijk voelen om bij alle leerlingen leerwinst te boeken, en dus niet enkel te focussen op de groep leerlingen zonder specifieke onderwijsbehoeften (Buzick, 2015; Pitoniak & Royer, 2001). Ten tweede gaat het om een behoorlijk grote groep leerlingen, namelijk naar schatting tussen de 10 en 15 procent van een leerlingcohort (Buzick, 2015). Deze leerlingen niet laten deelnemen zou een vertekend beeld geven van de gehele doelpopulatie. Hen niet op een voor hen aangepaste manier laten deelnemen zou mogelijk geen juiste weergave bieden van hun prestatiemogelijkheden. Een brede afname is dus noodzakelijk om een correct beeld te krijgen op systeem- en schoolniveau van de leerprestaties van alle leerlingen.

2.2.3. TOETSTECHNISCHE ARGUMENTEN VOOR TOETSAANPASSINGEN

Binnen de toetspraktijk zijn er verschillende principes die een rol spelen bij het al dan niet toelaten van aanpassingen aan de toets: inhoudsvaliditeit, (differentiële) effectiviteit en billijkheid. Deze deels onderling samenhangende principes worden hieronder toegelicht, waarna een korte discussie volgt.

INHOUDSVALIDITEIT EN CONSTRUCT IRRELEVANTE VARIANTIE

Een belangrijke voorwaarde voor grootschalige, gecentraliseerde toetsen is een gestandaardiseerde afname. Zo krijgen alle leerlingen in principe gelijke kansen om goed te presteren op de toets. Omdat hun prestaties ook een juist beeld geven van de werkelijke mogelijkheden van een leerling, moet zoveel mogelijk 'construct irrelevante variantie' worden vermeden. Er is sprake van construct irrelevante variantie wanneer de verschillen in leerlingprestaties op de toetsen niet alleen zijn toe te wijzen aan het construct of de onderliggende vaardigheid die men wenst te meten, maar ook aan andere kenmerken waarin leerlingen verschillen. Zo is het bijvoorbeeld mogelijk dat leerlingen met dyslexie die even vaardig zijn in wiskunde dan andere leerlingen toch minder goed presteren op een wiskundetoets. Verschillen in leesvaardigheid, i.e., verschillen in een contextueel leerlingkenmerk dat losstaat van het te meten construct, zouden geen rol mogen spelen bij het vaststellen van wiskundeprestaties. Om dit te bewerkstelligen kan worden toegestaan dat leerlingen met een attest van dyslexie voorleessoftware mogen gebruiken tijdens de wiskundetoets.

Toetsaanpassingen kunnen in principe alleen aan leerlingen toegestaan worden voor het voorkomen van construct irrelevante variantie. De toetsaanpassingen zelf mogen dus het te meten construct niet in het gedrang brengen (NAEP, 2014) maar hebben enkel als doel de verschillen veroorzaakt door contextuele hindernissen op te heffen. Als een toets bijvoorbeeld wil nagaan of leerlingen de maaltafels geautomatiseerd hebben, dan wordt dat toetsdoel in het gedrang gebracht door het toelaten van een maaltafelkaart, rekenmachine of extra afnametijd. Het invoeren van toetsaanpassingen moet dus worden voorafgegaan door minstens een inschatting van hun mogelijke impact op de inhoudsvaliditeit.

DIFFERENTIËLE EFFECTIVITEIT

Het toekennen van toetsaanpassingen aan leerlinggroepen moet volgens sommigen zelfs eerst ondersteund worden door empirisch onderzoek naar het effect ervan op de toetsscores. Sommigen betwijfelen de billijkheid van een toetsaanpassing als blijkt dat reguliere leerlingen ook baat zouden kunnen hebben bij de aanpassing (Buzick & Stone, 2014). Bepaalde onderzoekers vinden een toetsaanpassing dan ook pas billijk als die de toetsscores verhoogt voor leerlingen die de aanpassing nodig hebben, terwijl de toetsscores voor leerlingen die de aanpassing niet nodig hebben ongewijzigd blijven (Scarpati & Li, 2005; Zuriff, 2000). Ze verwachten dus dat de toetsaanpassing leidt tot een interactie-effect, i.e., differentieel effectief is voor specifieke leerlingen(groepen), en niet tot een hoofdeffect. Alleen op die manier wordt ook de inhoudsvaliditeit van de toets gegarandeerd. Immers, als een toetsaanpassing niet enkel helpt om belemmerende contextfactoren te reduceren bij bepaalde leerlinggroepen, maar ook

bijdraagt aan de prestatie van de reguliere leerlingen, dan raakt de toetsaanpassing aan het te meten construct.

Anderen stellen dat het voldoende is als empirisch kan aangetoond worden dat het voordeel van de toetsaanpassing voor de leerlingen met beperking groter is dan voor de leerlingen zonder beperking (Cahalan-Laitusis, 2007). Dit betekent dat men bij het vergelijken van de toetsprestaties bij het afleggen van de toets met en zonder toetsaanpassing vaststelt dat de leerlingen met een beperking in de eerste conditie gemiddeld een duidelijk hogere score behalen met dan zonder de aanpassing dan de andere leerlingen, die echter ook hoger kunnen scoren met de aanpassing. Er is dan zowel een hoofdeffect van de toetsaanpassing als een interactie-effect. Dit betekent evenwel dat het effect van de toetsaanpassing mee moet worden opgenomen in de omschrijving van de inhoudsvaliditeit van de toets. Afhankelijk van de gekozen benadering kan een toetsaanpassing dus worden ingezet als universele tool of toetsfaciliteit.

BILLIJKHEID OP LEERLING- ÉN SCHOOLNIVEAU

Het voorbeeld van de voorleessoftware bij dyslexie illustreert waarom het 'fair' of billijk wordt geacht dat er bij (grootschalige) toetsen verschillen in de afnameprocedure tussen leerlingen worden toegestaan, ook al creëert dat op zich verschillen in de gestandaardiseerde toetsafname. Deze billijkheid om construct irrelevante variantie tegen te gaan op leerlingniveau, mag zich bij grootschalige toetsen niet vertalen in bijkomende, construct-irrelevante schoolvariantie. Dit zou het geval zijn als scholen sterk zouden verschillen in het toelaten van extra hulpmiddelen of het geven van (deel)vrijstellingen bij de gecentraliseerde toetsen. Op zich is hier weinig empirisch onderzoek naar. Lai en Berkely (2012) toonden bijvoorbeeld wel aan dat er in de VS grote variabiliteit bestond tussen staten met betrekking tot toegestane toetsaanpassingen. Bij een meta-evaluatie van het M-decreet in Vlaanderen werd alvast gewag gemaakt van verschillen tussen scholen in de mate waarin ze redelijke aanpassingen toestaan (Departement Onderwijs en Vorming, 2017).

DISCUSSIE

De verschillende principes inzake het toelaten van toetsaanpassingen bij grootschalige toetsafnames maken duidelijk dat het opstellen van een toetsafnamebeleid voor gecentraliseerde toetsen niet eenvoudig is (Bolt & Thurlow, 2004). Het principe van standaardisatie van de afname kan deels worden losgelaten vanuit een billijkheidsprincipe waarbij toetsaanpassingen toegelaten worden voor bepaalde leerlinggroepen. Het principe van billijkheid kan echter nooit opwegen tegen het loslaten van de inhoudsvaliditeit. Dit laatste kan in het gedrang worden gebracht door het toelaten van toetsaanpassingen die geen strikt differentiële effectiviteit hebben en voor alle leerlingen leiden tot een verhoging van de toetsprestaties.

Het empirisch onderzoek naar de (differentiële) effectiviteit van toetsaanpassingen is erg intensief en dan ook eerder schaars, al zijn er ook al enkele meta-analyses uitgevoerd, zoals bijvoorbeeld over voorlezen bij toetsafnames bij wiskunde- en taaltoetsen (Buzik & Stone, 2014; Laitusis et al., 2012) of over het

toekennen van extra toetstijd (Duncan & Purcell, 2020). Resultaten van dergelijke studies zijn niet zonder meer te generaliseren naar andere toetscontexten. Het is echter ook niet haalbaar om elke toetsaanpassing eerst aan de hand van een effectiviteitsstudie te verifiëren bij diverse leerlinggroepen vooraleer ze in te voeren. Dit maakt dat het invoeren van toetsaanpassingen vaak eerder is gebaseerd op ‘face validity’ en bestaande praktijken dan op empirisch onderzoek naar de effecten van bepaalde maatregelen.

Om de billijkheid te bewaken, is het ook nodig dat het toetsafnamebeleid over welke leerlingen welke toetsaanpassingen kunnen krijgen zo concreet mogelijk wordt geformuleerd en zo gestandaardiseerd mogelijk wordt toegepast. Alleen op die manier kunnen ongewenste verschillen tussen scholen vermeden worden.

2.3. BREDE AFNAME: LEERLINGEN

2.3.1. OVERZICHT

In deze paragraaf worden de organisatorische aspecten van de uitrol van de brede afname op leerlingniveau besproken. Eerst wordt besproken welke leerlingen daarvoor in aanmerking zouden kunnen komen binnen de Vlaamse onderwijscontext. Vervolgens worden scenario’s voorgesteld hoe de brede afname vanuit het perspectief van de aanpassing van de toetsen voor individuele leerlingen zou kunnen verlopen, waarbij ook budgetinschattingen worden gemaakt.

In deze bespreking wordt geen onderscheid gemaakt tussen de verschillende onderwijssettings waar leerlingen schoollopen (die in § 2.4 aan bod komen). De uitrol van de brede afname van de gecentraliseerde toetsen voor leerlingen uit het reguliere onderwijs kan immers ook voor leerlingen uit andere onderwijssettings worden opengesteld.

2.3.2. OVER WELKE LEERLINGEN GAAT HET?

Binnen het Vlaamse onderwijs wordt het zorgbeleid van scholen momenteel³ doorgaans uitgewerkt volgens de principes van het zorgcontinuüm (zie bijvoorbeeld Vandecandelaere e.a., 2020) om zo adequaat onderwijs aan te bieden aan alle leerlingen van de school. Dit zorgcontinuüm bestaat uit verschillende fasen, die hieronder eerst kort worden toegelicht met het oog op het omschrijven van de groep leerlingen voor wie een aangepaste afname van de gecentraliseerde toetsen wenselijk zou kunnen zijn. Bij het schrijven van deze tekst werd gebruik gemaakt van de informatie op de website van Prodia (2020). Daarna wordt ook de aanvraag voor speciale onderwijsleermiddelen voor leerlingen met een visuele, auditieve of motorische beperking besproken. Tot slot wordt op basis van al deze informatie een eerste kadering gegeven van een brede afname van de gecentraliseerde toetsen, zowel met betrekking tot het verlenen van (deel)vrijstellingen als met betrekking tot het toelaten van toetsaanpassingen.

³ Vanaf het schooljaar 2021-2022 wordt een nieuw begeleidingsdecreet voorzien, wat mogelijk deels een impact kan hebben op hoe de brede afname binnen de gecentraliseerde toetsen gekaderd gaat worden in de onderwijspraktijk.

ZORGCONTINUÛM

Brede basiszorg (Fase 0)

Elke school biedt een brede basiszorg om tegemoet te komen aan de onderwijsbehoeften van diverse groepen van leerlingen binnen het gemeenschappelijke curriculum. Krachtlijnen bij dit zorgbeleid zijn krachtige leeromgevingen, stimuleren, differentiëren en remediëren.

Verhoogde zorg (Fase 1)

Wanneer er ongerustheid is over de leerontwikkeling van een bepaalde leerling kan een school samen met de ouders en leerling beslissen om over te stappen naar verhoogde zorg. Naast differentiëren en remediëren kan men bij deze verhoogde zorg ook een beroep doen op compenserende of dispenserende maatregelen. Dit laatste gebeurt bij voorkeur in samenspraak met het CLB.

Uitbreiding van de zorg (Fase 2)

Wanneer de leerontwikkeling stagneert of een negatieve evolutie dreigt kan de school in overleg met de ouders en leerling beslissen om het CLB in te schakelen. In tegenstelling tot Fase 1 gaat het dus om leerlingen met hardnekkige (leer)problemen. Ouders of leerlingen kunnen zelf ook het initiatief nemen om bij het CLB aan te kloppen. Het CLB kan dan (a) een handelingsgericht diagnostisch (HGD) traject opstarten; (b) informatie of advies verstrekken zonder HGD-traject, bijvoorbeeld bij een onderwijsloopbaankeuze; (c) een kortdurende begeleiding starten of (d) samenwerken met een netwerk.

De inschakeling van het CLB kan leiden tot een gemotiveerd verslag voor de leerling. In een gemotiveerd verslag wordt aangegeven waarom een leerling met specifieke onderwijsbehoeften toch het gemeenschappelijke curriculum kan volgen mits ondersteuning en redelijke aanpassingen. Op basis van dit verslag kan een leerling bijvoorbeeld toegang krijgen tot ADIBib (Aangepaste Digitale Bibliotheek voor leerlingen met een beperking in de schriftelijke communicatie) en dyslexiesoftware (AGODI, 2019). Sinds 1 september 2020 geldt dat bij een gemotiveerd verslag een verslag van het handelingsgerichte advies van het CLB in het multidisciplinaire dossier van de leerling wordt opgenomen. Dit gebeurt binnen het digitale platform "Leerlingen Activiteiten Registratie Systeem" of kortweg LARS dat gebruikt wordt door alle CLB's van alle netten. De informatie wordt daarbij ingevoerd met standaard nomenclatuurfiches ('nomfiches') die automatisch ook doorgestuurd worden naar het IRIS-platform (digitale leerlingenpaspoort) van de Vlaamse overheid, waartoe ook ondersteuners en de hiertoe gemachtigde contactpersonen van de school toegang hebben. De Vlaamse overheid beschikt via het IRIS-platform dus over precieze aantallen leerlingen met een gemotiveerd verslag, zonder evenwel de inhoud van dit verslag te kennen.

Individueel Aangepast Curriculum (Fase 3)

Sommige leerlingen hebben specifieke onderwijsbehoeften die om aanpassingen vragen die niet als redelijk worden geacht voor de school of die er toch niet voldoende voor zorgen dat de leerling het

gemeenschappelijke curriculum kan blijven volgen. Het CLB-team, de leerling, zijn ouders en de school kunnen dan beslissen om een 'verslag' op te maken waarin wordt aangegeven dat een leerling nood heeft aan een individueel aangepast curriculum (IAC). Het opmaken van een IAC betekent niet automatisch dat de leerling overstapt naar een school voor buitengewoon onderwijs. De school, de ouders en het CLB beslissen samen of de school, met redelijke aanpassingen, de gepaste begeleiding kan bieden (Vlaams Ministerie voor Onderwijs en Vorming, 2020). Dit is de idee van 'inclusie' dat door het M-decreet in 2015 werd ingevoerd in het onderwijs. Net als bij een gemotiveerd verslag wordt een verslag door het CLB opgenomen in LARS. De zogenaamde nomfiche wordt automatisch geüpload naar het IRIS-platform. De overheid kan dus het precieze aantal leerlingen met een verslag bepalen voor de doelgroep van elke gecentraliseerde toets, maar dus zonder de inhoud van het verslag te kennen.

Besluit

Redelijke aanpassingen en IAC's spelen in op de onderwijsbehoeften van elke individuele leerling. Dat gegeven betekent dus dat het een goede organisatie en administratie zal vereisen bij de gecentraliseerde toetsen om de juiste toetsaanpassingen aan de juiste leerlingen toe te kennen. Daarnaast is het zorgcontinuüm voor leerlingen een dynamisch gegeven. Indien bijvoorbeeld toetsaanpassingen bij een eerste afnamemoment van de gecentraliseerde toetsen worden toegestaan, dan hoeft dat niet noodzakelijk zo te blijven bij een volgend afnamemoment.

SPECIALE ONDERWIJSLEERMIDDELEN

Leerlingen met een auditieve, visuele of motorische beperking die in het reguliere onderwijs schoollopen kunnen op basis van een medisch attest bij de Cel Speciale Onderwijsleermiddelen (CSO) van het Agentschap voor Onderwijsdiensten (AGODI) een aanvraag indienen voor de toekenning van speciale onderwijsleermiddelen die het hen mogelijk maken om deel te nemen aan het onderwijsleerproces. Dankzij deze gefinancierde hulpmiddelen worden leerlingen minder beperkt door hun handicap en kunnen zij zich beter integreren in de klas en de school (AGODI, 2019). Voor leerlingen met een motorische beperking gaat het voornamelijk om aangepaste tafels en stoelen. Hulpmiddelen voor leerlingen met een visuele beperking zijn o.a. schoolsystemen⁴, aangepaste tafels, leeslampen en braileregels. Voor leerlingen met een auditieve beperking gaat het bijna uitsluitend om FM-apparatuur. Daarnaast kunnen leerlingen met een visuele beperking omzettingen vragen van hun leerboeken en studiemateriaal. Tolkondersteuning (gebarentolk of schrijftolk) wordt uitsluitend toegekend aan leerlingen met een auditieve beperking. Jaarlijks gaat het om een duizendtal aanvragen. In 2018 waren er 428 aanvragen uit het basisonderwijs en 364 uit het secundair onderwijs. Samen vormden zij 77 procent van alle aanvragen, die daarnaast ook kwamen van studenten en cursisten in het hoger onderwijs en het

⁴ Een schoolstelsel of schoolset is een combinatie van een bordcamera, een beeldscherm en/of een leesloep. Met de bordcamera kan een slechtziende leerling het schoolbord filmen en dat beeld vervolgens vergroten en/of de kleuren of contrasten aanpassen op zijn beeldscherm. Met de leesloep kan de leerling zijn schoolboeken of werkbladen vergroten. Het is ook mogelijk om enkel de bordcamera of enkel de leesloep te gebruiken i.p.v. de combinatie van beide. Er bestaan vaste schoolsystemen en mobiele (dus verplaatsbare) varianten.

volwassenenonderwijs. Veruit het grootste deel van het budget gaat naar doventolkondersteuning. Binnen de totale doelgroep van de gecentraliseerde toetsen gaat het dus om een zeer klein percentage van de leerlingen (cf. infra).

IMPLICATIES VOOR DE GECENTRALISEERDE TOETSEN

Mogelijke vormen van afname

Op basis van de Vlaamse onderwijspraktijk inzake het zorgbeleid en het toekennen van speciale onderwijsleermiddelen in het reguliere onderwijs kunnen er vier varianten onderscheiden worden in hoe de gecentraliseerde toetsen worden aangeboden. Dit wordt schematisch weergegeven in Figuur 2.1. Op basis van het curriculum dat leerlingen volgen (gemeenschappelijk of individueel aangepast) kan de inhoud van de toets aangepast worden. Leerlingen met een IAC hoeven misschien niet deel te nemen aan (al) de gecentraliseerde toetsen. Naast inhoudelijke aanpassingen kunnen de gecentraliseerde toetsen ook contextueel worden aangepast voor leerlingen door het al dan niet toelaten van bepaalde toetsaanpassingen. Op basis van de combinatie van beide criteria zijn er zo vier mogelijke vormen van de afname van de gecentraliseerde toetsen.

		Toetsaanpassingen	
		Niet nodig	Wel nodig
Curriculum	Gemeenschappelijk		
	Individueel aangepast		

Figuur 2.1. Schematische weergave van de vier mogelijke vormen van afname van de gecentraliseerde toetsen.

Voor de gecentraliseerde toetsen zal de overgrote meerderheid van de leerlingen uit het reguliere onderwijs de toetsen afleggen binnen het gemeenschappelijke curriculum zonder toetsaanpassingen. Voor elke afname en doelgroep gaat het om zowat 70000 leerlingen (cf. infra in Tabel 2.3). Het precieze aantal leerlingen in de andere cellen in Figuur 2.1 is moeilijker te bepalen en daarvoor zal allicht ook een juridisch kader moeten worden gecreëerd. Beide elementen worden hieronder verder verduidelijkt.

Aantallen leerlingen?

Speciale onderwijsleermiddelen

In Tabel 2.2 wordt per afnamemoment van de gecentraliseerde toetsen een overzicht gegeven van het aantal leerlingen aan wie op 1 februari 2020 speciale hulpmiddelen, omzettingen naar braille of

Tabel 2.2: Aantal leerlingen aan wie speciale onderwijsleermiddelen werd toegekend op 1 februari 2021.

Onderwijsleermiddel	Basisonderwijs		Secundair Onderwijs				
	4 ^{de} lj	6 ^{de} lj.	2A	2BVL	6 aso	6 bso	6 tso
<i>Hulpmiddelen</i>							
motorisch	66	43	2	--	--	1	--
visueel	36	18	6	1	5	1	5
auditief	2	--	--	--	--	--	--
<i>Omzettingen</i>	10	5	9	2	6	1	6
<i>Tolkondersteuner</i>							
Gebarentolk	2	4	8	2	--	4	1
Schrijftolk	--	1	4	--	2	1	2

Cijfers bezorgd door de Cel Speciale Onderwijsleermiddelen, AGODI, Ministerie van Onderwijs en Vorming

tolkondersteuning (enkel bij auditieve beperking) toegekend werden⁵. De leerlingaantallen in het deel van de tabel met de hulpmiddelen betreffen telkens unieke leerlingen. Een aantal van deze leerlingen komt ook voor bij de bepaling van de leerlingaantallen die vermeld staan bij de omzettingen en tolkondersteuning. Het totaal aantal leerlingen ligt dus wat lager dan de som van alle kolomtotalen.

Zoals eerder aangegeven gaat het om een beperkt aantal leerlingen per afnamemoment. Opvallend is dat er meer leerlingen met speciale onderwijsleermiddelen in het basisonderwijs zijn dan in het secundair onderwijs. Er zijn meer leerlingen met speciale onderwijsleermiddelen in het vierde leerjaar van het basisonderwijs dan in het zesde leerjaar en mogelijk ook in 2A dan in andere leerjaren van het secundair onderwijs. De Cel Speciale Onderwijsleermiddelen kon voor deze trends geen sluitende verklaring bieden. Het aantal aanvragen en toekenningen voor speciale onderwijsleermiddelen neemt de afgelopen jaren toe. Toch lijkt het duidelijk dat de omvang van deze leerlingengroep binnen de doelpopulatie van de gecentraliseerde toetsen ook bij de effectieve implementatie ervan beperkt zal blijven.

(Gemotiveerd) verslag

De Vlaamse administratie beschikt via het IRIS-platform momenteel over de aantallen leerlingen met een verslag en een gemotiveerd verslag. Dat aantal is voor de organisatie van de afname van de gecentraliseerde toetsen opvraagbaar per leerjaar (zie Tabel 2.3 hieronder) en allicht ook per school. De overheid heeft echter geen toegang tot de inhoud van die verslagen. Er kan dus vanuit die centrale registratie zeker geen automatische toekenning gebeuren van (deel)vrijstellingen of toetsaanpassingen bij de gecentraliseerde toetsen gegeven dat daarvoor ook een inhoudelijke beoordeling van het (gemotiveerd) verslag nodig is. Het CLB van elke school heeft zicht op de inhoud van de gemotiveerde verslagen en de verslagen, gegeven dat zij mede verantwoordelijk zijn om die op te stellen. Zij zijn echter gebonden aan het beroepsgeheim en om evidente redenen met betrekking tot bescherming van de persoonlijke levenssfeer (zie ook verderop) kunnen CLB's die informatie niet ter beschikking stellen van

⁵ Met dank aan de Cel Speciale Onderwijsleermiddelen van Agodi voor het aanleveren van dit cijfermateriaal en voor de bijkomende toelichtingen (K. Liessens, persoonlijke communicatie, 2 en 8 februari 2021).

derden. Ook hier zou een juridisch (en organisatorisch) kader nodig zijn waarin de CLB's betrokken zouden worden bij het toewijzen van een leerling aan een van de vier vormen van afname.

Compenserende en dispenserende maatregelen op schoolniveau

Tot slot zijn scholen autonoom in het tegemoetkomen aan de onderwijsbehoeften van hun leerlingen in de fasen van brede basiszorg en verhoogde zorg. Scholen kunnen daarbij maatregelen nemen die gelden bij de evaluatie van die leerlingen. Zo kan een school met veel anderstalige leerlingen in het kader van hun brede basiszorg (Fase 0) toestaan dat leerlingen bij toetsen en examens steeds een woordenboek mogen gebruiken. Of binnen de verhoogde zorg (Fase 1) kan een school ook met betrekking tot de evaluatie van hun leerlingen compenserende (bv. langere werktijd of afname in een apart lokaal) of dispenserende (bv. geen spellingsfouten aanrekenen) maatregelen nemen voor bepaalde leerlingengroepen. Er is geen centrale registratie over hoeveel leerlingen dit gaat. Zoals verderop nog wordt besproken (zie toekenningsprocedure) is het allicht niet zo dat alle evaluatie-afspraken voor bepaalde leerlinggroepen binnen een school zonder meer kunnen gelden bij de afname van de gecentraliseerde toetsen (cf. de principes van standaardisatie en billijkheid).

Andere onderwijssettings

Naast de leerlingen uit het reguliere onderwijs kunnen mogelijk ook leerlingen uit andere onderwijssettings (deel)vrijstellingen bekomen en/of toetsaanpassingen nodig hebben, indien deze onderwijssettings zouden deelnemen aan de gecentraliseerde toetsen. Het is ook hier momenteel niet duidelijk om hoeveel leerlingen dat per gecentraliseerde toets zou kunnen gaan. Welk kan op basis van de gegevens in Tabel 2.4 (cf. infra) hier een maximum aantal op worden geplakt.

Privacy

Los van het inschatten van de aantallen leerlingen in elke afnamevorm van de gecentraliseerde toetsen is er ook bij het toewijzen van een individuele leerling aan een afnamevorm een belangrijk probleem met betrekking tot de privacy. Zowel de omschrijving van de specifieke onderwijsbehoeften van een leerling (cf. het verslag of gemotiveerd verslag) als de toewijzing van de leerlingen aan een bepaalde afnamevorm op een bepaald afnamemoment van de gecentraliseerde toetsen zijn gegevens uit de persoonlijke levenssfeer van een leerling, die dus gebonden zijn aan Algemene Verordening Gegevensbescherming (AVG). Deze problematiek wordt verder uitgewerkt in het juridische luik van Perceel 3. In dit perceel wordt bij de uitwerking van de verschillende scenario's voor de brede afname verder uitgegaan van een juridische oplossing voor dat probleem.

2.3.3. SCENARIO'S INZAKE BREDE AFNAME BIJ LEERLINGEN

De organisatie van een brede afname bij de gecentraliseerde toetsen begint al bij de toetsontwikkeling, zoals hieronder wordt toegelicht. Daarna volgen scenario's over het toekenningsbeleid en de -procedure.

Tot slot worden ook scenario's voorgesteld over de implementatie van de brede afname tijdens de toetsafname.

TOETSONTWIKKELING

Al van bij de toetsontwikkeling wordt de brede afname van de gecentraliseerde toetsen voorbereid, zowel in het vormgeven van de items als in het coderen van bijkomende itemkenmerken in de itembank en het eventueel aanpassen van items met het oog op hun toegankelijkheid.

Universal Design for Assessments

Binnen de onderwijspraktijk worden de principes van Universal Design for Learning (UDL; Thurlow, Warren, & Chia, 2020) vooropgesteld om leeromgevingen zodanig vorm te geven dat daarbij zoveel mogelijk tegemoet wordt gekomen aan individuele verschillen in onderwijsbehoeften tussen leerlingen. Parallel hiermee heeft het National Centre on Educational Outcomes (NCEO) zeven elementen van een 'universal design for assessments' ontwikkeld, namelijk:

- Inclusieve doelpopulatie;
- Nauwkeurig gedefinieerde constructen;
- Toegankelijke, niet-bevooroordeelde items;
- Geschikt voor toetsaanpassingen;
- Eenvoudige, duidelijke en intuïtieve instructies en procedures;
- Maximale verstaanbaarheid en begrijpelijkheid;
- Maximale leesbaarheid;

Deze elementen worden best al meegenomen bij de toetsontwikkeling (Johnstone, 2003). Om de inhoudsvaliditeit van de toets zo goed mogelijk weer te geven helpt het bijvoorbeeld om de te meten constructen zo meetbaar en nauwkeurig mogelijk te omschrijven met toetsmatrijzen (cf. volgende paragraaf). Een ander voorbeeld is dat bij het formuleren van de items grondig moet worden nagedacht over de (universele toepasbaarheid van de) inhoud (cf. niet-bevooroordeelde items) en de formulering (cf. maximale verstaanbaarheid). Goede instellingen voor een uniforme layout van de items verhogen dan weer mee de leesbaarheid van de toets. Een aantal van deze elementen kunnen de noodzaak aan toetsaanpassingen verminderen. Wanneer moeilijke woorden vermeden worden, is de toets bijvoorbeeld meer toegankelijk voor anderstalige leerlingen en zal ook de construct irrelevante variantie van de toets dalen. Met het oog op de brede afname van de gecentraliseerde toetsen is het dus zeker aan te bevelen dat de toetsontwikkelaars binnen het nieuwe steunpunt zich laten inspireren door deze elementen. Merk op dat deze elementen ook al worden toegepast binnen het Steunpunt Toetsontwikkeling en Peilingen (STEP).

Oplijsten en coderen van itemkenmerken

Om de inhoudsvaliditeit van een toets te bewerkstelligen wordt er bij de toetsontwikkeling van grootschalige toetsen per meetschaal een toetsmatrijs voorzien, waarin voor elk te ontwikkelen item wordt aangegeven welk(e) onderwijsdoel(en) het item beoogt te meten op welk verwerkingsniveau. Deze iteminformatie is cruciaal om voor leerlingen met een IAC te kunnen beoordelen welke items ze wel en niet moeten kunnen beheersen.

Met het oog op toetsaanpassingen moeten er in de itembank bijkomende itemkenmerken worden voorzien. Daarmee moeten de toetsontwikkelaars kunnen aangeven welke toetsaanpassingen voor elk item mogelijk zijn, zonder dat de inhoudsvaliditeit van de toets in het gedrang komt. Zo kan bijvoorbeeld bij wiskunde-items worden aangegeven of een rekenmachine wel of niet mag worden toegestaan voor alle leerlingen dan wel alleen voor leerlingen met bijvoorbeeld dyscalculie. Het voorzien van deze itemkenmerken als bijkomende variabelen in de itembank moet worden opgenomen in het bestek van de overheid voor de ontwikkeling van het toetsplatform. Het extra werk voor de toetsontwikkelaars kan binnen hun werkzaamheden voor de toetsontwikkeling voor het nieuwe steunpunt worden opgenomen.

Toegankelijk maken van items

Indien de overheid beslist dat de gecentraliseerde toetsen ook toegankelijk moeten worden gemaakt voor leerlingen met een visuele of auditieve handicap⁶, dan dienen de toetsontwikkelaars bij elk item aan te geven of het zonder bijkomende aanpassingen doenbaar is voor deze leerlingen, dan wel welke aanpassingen nodig zijn om het item toegankelijk te maken.

Voor leerlingen met een visuele handicap maakt een braileregelsysteem die op het toetsplatform kan worden aangesloten veel items toegankelijk. Wel zijn er specifieke aanpassingen nodig voor figuren en allicht ook voor tabellen. Bepaalde afbeeldingen kunnen bijvoorbeeld in reliëf worden omgezet. Ook de toegankelijkheid van videofragmenten vraagt extra aandacht. Via de collega's van Perceel 1 konden we een prijsopgave inkijken voor het opstellen van de brailerversie van de Nederlandse Centrale Eindtoets uit 2018 met 200 opgaven rekenen en taal. Dit ging wel om het omzetten van een toets op papier in braille. Het totale kostenplaatje bedroeg zo'n 7000 euro en omvatte zowel de opdrachtbegeleiding, het maken van reliëftekeningen, het bewerken van het tekstbestand en de nabewerking.

Voor leerlingen met een auditieve handicap vormen allicht alleen items met media-fragmenten een hinderpaal. Deze fragmenten kunnen worden ondertiteld of er kan een aparte versie van worden gemaakt waarbij in een hoek van het beeld een gebarentolk wordt toegevoegd. De kostprijs hiervoor is per item

⁶ Voor leerlingen met een motorische handicap gaat het allicht vooral om aanpassingen aan de toetsafnameprocedure, maar slechts in uitzonderlijke gevallen om aanpassingen aan een item, tenzij bijvoorbeeld een item waar iets getekend moet worden.

gelijk aan het opmaken van een nieuw mediafragment, de loonkost van de gebarentolk⁷ en de kosten voor de opname van de gebarentolk. Daarnaast is het ook mogelijk dat leerlingen met een auditieve handicap tijdens het afleggen van de gecentraliseerde toets lokaal eerder dan digitaal een beroep kunnen doen op een doventolk. Afhankelijk van de finaliteit van de gecentraliseerde toetsen en de verantwoordelijkheden die verschillende actoren hebben binnen de toetsafname, kan hiervoor mogelijk dezelfde persoon worden ingeschakeld die doorheen het schooljaar al instaat voor deze ondersteuning.

Zoals in het rapport van Perceel 1 wordt aangegeven is het in principe nodig om in een apart kalibratieonderzoek na te gaan of de items in de toegankelijk gemaakte toets op een vergelijkbare manier functioneren als die van de oorspronkelijke toets en dus geen differential item functioning (DIF) vertonen. Dit is een extra kostprijs, waarbij bovendien verwacht kan worden dat het aantal leerlingen met een auditieve of visuele handicap te beperkt is om dit grondig te onderzoeken. Het niet uitvoeren van deze controle op DIF houdt dan wel een risico in bij het gebruik van de aangepaste toets. Post-afname kunnen tijdens de analyse en rapportage wel een aantal maatregelen genomen worden om dit risico te reduceren, zoals bijvoorbeeld multiple-group analyse indien met IRT wordt gewerkt en de toetsen tussen de leerlingengroepen voldoende geankerd zijn of rapportage op groepsniveau.

TOEKENNINGSBELEID

In deze paragraaf worden eerst scenario's voorgesteld voor het toekenningsbeleid van (deel)vrijstellingen en toetsaanpassingen. Daarbij worden telkens twee vragen beantwoord: (a) welke leerlingen komen in aanmerking? en (b) welke (deel)vrijstellingen of toetsaanpassingen worden toegestaan? Daarna wordt ook ingegaan op de hulpmiddelen van leerlingen die speciale onderwijsleermiddelen gebruiken.

(Deel)vrijstellingen

Voor welke leerlingen?

Leerlingen met een IAC

Leerlingen die een IAC volgen, komen in aanmerking voor het aanvragen van (deel)vrijstellingen op de gecentraliseerde toetsen aangezien zij niet het gehele gemeenschappelijke curriculum volgen. Al deze leerlingen hebben een verslag, waarmee hun aanvraag gestaafd kan worden.

Leerlingen in het onthaalonderwijs voor anderstalige kinderen (OKAN)

Via het onthaalonderwijs voor anderstalige nieuwkomers krijgen leerlingen van niet-Belgische of niet-Nederlandse nationaliteit die nieuw binnenkomen in ons land en het Nederlands onvoldoende beheersen de kans om zo snel mogelijk Nederlands leren. In het basisonderwijs gebeurt dit via één of twee

⁷ De Beroepsvereniging van Vlaamse Gebarentaaltolken (BVG, 2021) spreekt van een bruto-uurloon van €37,61 per begonnen tolkuur voor alle gesubsidieerde uren van de overheden Welzijn, Werk en Onderwijs. Voorbereidingstijd en reistijd worden niet vergoed.

onthaaljaren (met eventueel een vervolgjaar). De school heeft een grote autonomie in hoe ze dit inricht. In het secundaire onderwijs krijgen deze anderstalige leerlingen 1 jaar taallessen (de zogenaamde OKAN-klas) waarna ze begeleiding krijgen in vervolgtraject (Departement Onderwijs en Vorming, 2021).

Er valt te overwegen om anderstalige nieuwkomers vrijstelling te geven op de gecentraliseerde toetsen van zowel wiskunde als taal. Dit is een praktijk die bijvoorbeeld wordt toegepast bij de eindtoets basisonderwijs in Nederland. Ook in Australië worden niet-Engelstalige leerlingen die minder dan een jaar in Australië wonen vrijgesteld van deelname aan NAPLAN (National Assessment Program – Literacy and Numeracy). Binnen het Vlaamse peilingsonderzoek worden anderstalige nieuwkomers geweerd van de analyses om zo het probleem van construct irrelevantie te minimaliseren. Scholen beslissen zelf of ze deze leerlingen laten deelnemen maar de data van deze leerlingen worden niet verrekend in de resultaten.

Opt out?

In Australië kunnen ouders weigeren dat hun kind deelneemt aan NAPLAN wegens 'religieuze overwegingen of filosofische tegenkanting tegen toetsen' (Santiago et al., 2011; Vanhoof et al, 2015). Bij de Vlaamse peilingen wordt sinds het in voege zijn van de Algemene verordening gegevensbescherming (AVG) aan de ouders gevraagd of ze weigeren dat hun kind deelneemt. Bij de peilingen die in het voorjaar van 2021 doorgaan is er bij STEP bijvoorbeeld bij 0.9% van de leerlingen om opt out gevraagd voor de peiling wiskunde in het basisonderwijs en bij 2,3% van de leerlingen bij de peiling kritisch denken, mediawijsheid en probleemoplossen in de derde graad secundair onderwijs (in combinatie met PAV voor het bso). Het is de vraag in welke mate een dergelijke 'opt out'-mogelijkheid ook noodzakelijk is bij de gecentraliseerde toetsen. Deze problematiek wordt in Perceel 3 behandeld.

Welke (deel)vrijstellingen?

Scenario 1: Op alle toetsen

Voor de anderstalige nieuwkomers is een mogelijke piste om hen vrij te stellen van alle gecentraliseerde toetsen. Een leerling heeft het Nederlands immers niet alleen nodig voor een toets Nederlands, maar ook voor een toets wiskunde. Leerlingen wiens ouders weigeren dat ze deelnemen ('opt out') zullen aan geen enkele toets deelnemen. Ook voor leerlingen met een IAC zou men kunnen opteren om hen van alle gecentraliseerde toetsen vrij te stellen. Dat is organisatorisch eenvoudig omdat deze leerlingen bij de overheid bekend zijn via het IRIS-platform, zonder evenwel de inhoud van dat verslag te kennen. Het opvragen van deze laatste informatie is dan in dit scenario niet nodig.

Scenario 2: (Deel)vrijstellingen op inhoudelijke gronden

Voor leerlingen met een IAC zou het verlenen van (deel)vrijstellingen op inhoudelijke basis kunnen gebeuren: leerlingen hoeven geen gecentraliseerde toets af te leggen voor de onderwijsdoelen waar ze het gemeenschappelijk curriculum niet voor volgen. Als ze wel dat deel van het gemeenschappelijk

curriculum volgen, dan leggen ze wel de betreffende gecentraliseerde toets af. Dit scenario kan op twee manieren worden geïmplementeerd.

Scenario 2a: (Deel)vrijstellingen op itemniveau

Gegeven dat van elk item bekend is welke eindterm(en) het beoogt te meten, en gegeven dat van elke leerling geweten is welke eindtermen bij die leerling worden nagestreefd, kan er in principe op itemniveau worden bepaald welke items of itemgroepen binnen het IAC van een leerling vallen. Dan zouden (deel)vrijstellingen gegeven kunnen worden op itemniveau, zodat een leerling met IAC een individueel aangepaste toets krijgt.

Dit scenario is erg complex. Enerzijds moet in de toekenningsprocedure elke toets in principe item per item worden overlopen. Anderzijds moeten er voor deze leerlingen afzonderlijke psychometrische analyses worden uitgevoerd om de positie van de leerling op de meetschaal te schatten. Het is niet uit te sluiten dat het niet altijd haalbaar blijkt om een voldoende betrouwbare vaardigheidsparameter te schatten omdat er slechts weinig items overblijven. Mogelijk heeft dit implicaties voor het aantal items dat per eindterm dient ontwikkeld te worden en dus voldoende hoog dient te zijn. Ook zou het kunnen dat de resterende items slechts een deel van de meetschaal omvatten, waardoor er mogelijk ook vertekeningen op de geschatte vaardigheid kunnen optreden. Dit betekent dat de toets inhoudelijk valide is voor het IAC van de leerling, maar niet meer ten opzichte van de getoetste eindtermen, de meetschaal en de daarop vastgelegde cesuur. Dit scenario lijkt daarom niet aangewezen, zeker niet bij de opstart van het systeem van gecentraliseerde toetsen. Na een aantal jaren ervaring kan er op basis van de ingediende casussen desgevallend overwogen worden of er toch nog mogelijkheden zijn om te werken met vrijstellingen op itemniveau.

Scenario 2b: (Deel)vrijstellingen op toetsniveau

Gegeven dat van elke toets bekend is welke eindtermen die beoogt te meten, kan men nagaan of het IAC van een leerling al deze eindtermen omvat. Indien dit zo is, kan de leerling de toets afleggen. Indien dit niet zo is, zelfs al zou het gaan om slechts één eindterm die niet opgenomen is in het IAC, dan legt de leerling de gecentraliseerde toets niet af en krijgt daarvoor een vrijstelling. Dit scenario vraagt nog altijd toetsaanbiedingen op maat. Leerlingen met een IAC leggen dan niet noodzakelijk alle gecentraliseerde toetsen af van het getoetste leer- of vakgebied⁸.

Bijkomende overwegingen

Bij het verlenen van vrijstellingen of deelvrijstellingen is de schoolfeedback niet gebaseerd op alle leerlingen van een school. Dit moet duidelijk worden aangegeven in de schoolfeedbackrapporten. Op

⁸ Mogelijk werken de gecentraliseerde toetsen met verschillende deeldomeinen voor wiskunde en Nederlands die elk in een aparte toets worden bevroegd. Indien de gecentraliseerde toetsen zouden gebruik maken van 'omnibustoetsen' (waarin verschillende deeldomeinen samen worden bevroegd), dan valt een leerling met een IAC voor een deel van de onderwijsdoelen van wiskunde of Nederlands uit voor de gehele toets.

stysteemniveau kan door het nieuwe steunpunt gerapporteerd hoeveel leerlingen (deel)vrijstellingen kregen op de toetsen en of er schoolkenmerken zijn die hiermee samenhangen.

Bij de keuze tussen de scenario's kunnen de verschillende overwegingen voor een brede afname die bij het begin van dit hoofdstuk naar voren werden geschoven worden ingeroepen: wat is juridisch, onderwijskundig en toetstechnisch van belang bij het al dan niet opnemen van leerlingen met IAC en van OKAN-leerlingen bij de gecentraliseerde toetsen? Wat is daarbij voor de scholen de beste maat voor hun interne kwaliteitszorg? Daarnaast is er ook een belangrijke organisatorische kwestie. Indien deze leerlingen 'per definitie' vrijstelling krijgen op alle toetsen, dan wordt daarmee een grote organisatorische klip van het afstemmen van een IAC met de inhoud van de toetsen omzeild.

Wanneer opt out mogelijk wordt bij de gecentraliseerde toetsen, dan heeft dit vanzelfsprekend gevolgen voor de representativiteit van de resultaten op school- en systeemniveau als deze opt out frequent zou gebruikt worden, zeker als die keuze zou samenhangen met leerling- of schoolkenmerken.

Toetsaanpassingen

Voor welke leerlingen?

Deels parallel aan de discussie over de (differentiële) effectiviteit van toepassingen zijn er verschillende scenario's mogelijk op welk niveau van individualisering toetsaanpassingen moeten worden aangeboden. Aan de ene kant van het continuüm kan men ervoor kiezen om alle toetsaanpassingen beschikbaar te stellen aan alle leerlingen. Aan de andere kant zou men de toetsaanpassingen enkel kunnen aanbieden aan leerlingen die duidelijk kunnen aantonen dat ze deze toetsaanpassing vanuit hun specifieke onderwijsbehoeften nodig hebben. Beide scenario's worden hieronder verder besproken.

Scenario 1: voor alle leerlingen

Indien alle toetsaanpassingen beschikbaar gesteld worden aan alle leerlingen, dan kunnen zij die volgens eigen voorkeuren of behoeften selecteren of gebruiken. Voorbeelden daarvan zijn digitale toetsaanpassingen, zoals een zoom-functie (vergrootglas), keuze tussen verschillende lettertypes/achtergronden, volumeknop voor geluidsfragmenten (met koptelefoon), rekenmachine en een 'text to speech' knop om voorleessoftware te activeren (cf. Tabel 2.1). Ook kan de tijdsduur van de toetsafname bijvoorbeeld zo ruim worden genomen dat die voldoende zou moeten zijn voor alle leerlingen (cf. Universal Design), dus zonder dat sommige leerlinggroepen extra tijd hoeven te krijgen.

De argumentatie voor dit scenario sluit enerzijds aan bij de studies rond de effectiviteit van toetsaanpassingen die aantonen dat alle leerlingen er mogelijk voordeel uit kunnen halen en anderzijds bij de opvattingen van een Universal Design for Assessments. In het digitale platform wordt dan automatisch voorzien welke toetsaanpassingen bij een toets zijn toegestaan. Mogelijk kan dit zelfs gespecificeerd worden per item binnen de toets met het oog op het bewaken van de inhoudelijke validiteit.

Zoals eerder al aangegeven is een belangrijk nadeel van dit scenario dat wat de toets meet (inhoudsvaliditeit), kan veranderen door toetsaanpassingen toe te laten. Een ander mogelijk nadeel is dat sommige leerlingen ook benadeeld zouden kunnen worden tijdens de toetsafname omdat ze door de mogelijkheden van toetsaanpassingen meer afgeleid zijn van de toets.

Vooraf wanneer de toetsen via meerdere modaliteiten worden aangeboden, e.g., schriftelijke toetsvragen ondersteund met auditieve verduidelijking via text-to-speech, kunnen problemen optreden. Hoewel onderzoeksresultaten niet eenduidig zijn, tonen verschillende studies naar het gebruik van duale modaliteiten in functie van leesbegrip aan dat leerlingen minder goed presteren wanneer toetsen en leermateriaal worden aangeboden via meerdere modi. De hypothese luidt dat simultane input van twee modi de cognitieve load voor leerlingen vergroot waardoor de leerprestaties dalen (Diao & Sweller, 2007; Plass, Moreno, & Brunken, 2010). Dit scenario vereist dan minstens dat leerlingen in de klaspraktijk al ervaring opdoen met het aanbod aan toetsaanpassingen. Ook is het de vraag of met het beschikbaar stellen van digitale toetsaanpassingen aan alle vragen naar toetsaanpassingen voor leerlingen met specifieke onderwijsbehoeften tegemoet kan worden gekomen. Lokale aanpassingen (cf. Tabel 2.1) zitten in dit scenario bijvoorbeeld nog niet vervat. Aanpassingen met betrekking tot bijvoorbeeld het afnamelokaal zijn ook moeilijk centraal te regelen.

Scenario 2: voor welomschreven leerlinggroepen

Men kan als toekenningsbeleid ook stellen om toetsaanpassingen enkel aan te bieden aan bepaalde leerlinggroepen waarvan men duidelijk kan aantonen dat die deze toetsaanpassing vanuit hun specifieke onderwijsbehoeften nodig hebben. De toetsaanpassing moet dan op individuele basis worden toegekend. Een leerling moet dan aantonen dat hij of zij behoort tot de omschreven leerlingengroep.

Leerlingen met een gemotiveerd verslag beschikken in principe over een handelingsgericht advies van het CLB, waarin aangegeven wordt welke redelijke aanpassingen in de klaspraktijk worden ingevoerd. Indien deze redelijke aanpassingen ook betrekking hebben op evaluaties, dan lijkt het voor de hand te liggen dat deze leerlingen dezelfde toetsaanpassingen mogen gebruiken tijdens de gecentraliseerde toetsen, tenzij dat op die manier de inhoudsvaliditeit van de toets in het gedrang komt. Deze leerlingen kunnen deze vraag in principe ook staven met het CLB-advies.

Zoals eerder aangegeven zouden scholen aan bepaalde leerlinggroepen mogelijk in Fase 0 (brede basiszorg) al faciliteiten kunnen toekennen bij evaluatie-activiteiten. Zeker binnen Fase 1 (verhoogde zorg) kunnen scholen in principe autonoom compenserende of dispenserende maatregelen nemen voor bepaalde leerlingen die mogelijk ook gelden bij evaluatie-activiteiten. Op dit moment hebben wij geen zicht op het aantal leerlingen, op de aard van de maatregelen en of er schoolverschillen bestaan in dit zorgbeleid. Een indicatie hiervan zou kunnen worden verkregen door bijkomende analyses op de bevragingen die STEP tegenwoordig bij elke peiling voert naar de aard van de redelijke aanpassingen die leerlingen bij toetsen in de klas mogen gebruiken en of ze daar een verslag voor hebben.

Het is op dit moment ook moeilijk te beoordelen of vragen naar toetsaanpassingen bij de gecentraliseerde toetsen op basis van Fase 0 en Fase 1 in het zorgbeleid van een school vanuit het oogpunt van een gestandaardiseerde afname steeds gerechtvaardigd geacht zullen worden, zeker als hierin verschillen zijn tussen scholen. Binnen resonantiegroepen bij peilingen wiskunde leidt de vraag naar het toelaten van hulpmiddelen bijvoorbeeld steevast tot discussie, zeker bij peilingen in het beroepsvoorbereidend leerjaar of in de derde graad beroepsonderwijs. Enerzijds wordt gesteld dat leerlingen deze hulpmiddelen nodig hebben om te kunnen presteren. Dit komt overeen ‘scaffolding without fading’ of ook ‘scaffolds for performance’ (Pea, 2004), waarbij leerlingen de hulpmiddelen blijven nodig hebben om te kunnen presteren op het verwachte niveau, zodat die hulpmiddelen ‘tools for living’ worden omdat de kennis onvoldoende geïnternaliseerd is (Carmien & Fischer, 2005)⁹. Anderzijds wordt aangegeven dat bepaalde hulpmiddelen (zoals bijvoorbeeld het gebruik van kaarten met een stappenplan of een overzicht van een leerstofonderdeel) wel ter ondersteuning van het leerproces kunnen worden ingezet, maar dat deze ‘scaffolding’ in principe na verloop van tijd zou moeten worden stopgezet. Het blijven toelaten van deze hulpmiddelen houdt volgens deze visie dan het risico in op overcompensatie en op een daling wat van de leerlingen in het onderwijsproces verwacht wordt. Daarom wordt gepleit om het gebruik van bijkomende hulpmiddelen bij evaluatie (en dus bij een gecentraliseerde toets) niet toe te laten om zo het daadwerkelijke prestatieniveau van de leerlingen niet te maskeren.

De discussie over het al dan niet mogen gebruiken van bepaalde hulpmiddelen bij de gecentraliseerde toetsen raakt dus ook aan de didactische opvattingen en principes van leerkrachten. Hiervoor zal wel een besluit genomen moeten worden aangezien het omwille van standaardisatie nodig is dat er over scholen heen een uniforme afnameprocedure is.

Welke toetsaanpassingen?

Hierboven werd een algemene omschrijving gegeven van mogelijke toetsaanpassingen die digitaal en lokaal kunnen aangeboden worden bij een gecentraliseerde toets. Op dit moment is er geen zicht op de mate waarin elk van deze toetsaanpassingen in de Vlaamse onderwijscontext effectief wordt ingezet. Allicht zijn er ook verschillende combinaties gangbaar. Daarnaast zal bijvoorbeeld ook moeten onderzocht worden in welke mate de voorleessoftware die door de Vlaamse overheid aan leerlingen mits attestering ter beschikking kan worden gesteld (AGODI, 2019) compatibel is met het toetsplatform, dan wel dat binnen het toetsplatform een eigen voorleesfunctie (‘text to speech’) wordt voorzien. Dit laatste heeft organisatorisch duidelijk voordelen, maar misschien moeten leerlingen dan op voorhand kunnen oefenen binnen het toetsplatform. Ook hier is een belangrijke taak voor het nieuwe steunpunt weggelegd.

Ongeacht de keuze van het scenario van aanbiedingen (voor alle leerlingen of voor individuele leerlingen), moet er voor elke gecentraliseerde toets een exhaustieve lijst opgesteld worden van de mogelijk digitale en lokale toetsaanpassingen die toegelaten kunnen worden. Dit overzicht zou enerzijds ‘bottom-up’

⁹ Indien dit zo is, dan zou dit eigenlijk bij de omschrijving van de onderwijsdoelen door de overheid moeten worden opgenomen.

opgesteld kunnen worden op basis van de huidige (toets)praktijken in het Vlaamse onderwijs. De toegelaten toetsaanpassingen zouden anderzijds ook ‘top-down’ kunnen worden opgesteld op basis van een inhoudelijke inschatting van experts. De geschatte kost voor het opstellen van deze lijst verschilt naargelang de gekozen aanpak. Dit kan een bijkomende onderzoeksopdracht zijn voor of naast het steunpunt voor gecentraliseerde toetsen. Nadien moet voor elke toets die ontwikkeld wordt, door de toetsontwikkelaars worden aangevinkt welke toetsaanpassingen bij afzonderlijke items en bij de toets in zijn geheel wel of niet toegelaten kunnen worden. Deze extra check kan worden opgenomen binnen de werkzaamheden van het steunpunt gecentraliseerde toetsen.

Alleszins vraagt de invoering van gecentraliseerde toetsen in Vlaanderen naar beleidskeuzes over welke toetsaanpassingen bij de gecentraliseerde toetsen effectief kunnen toegelaten worden en welke niet, ook al zouden die wel in de klaspraktijk in (sommige) scholen worden gehanteerd. Het is nodig dat hierover al voor het proces van toetsontwikkeling beslissingen over genomen worden, al dan niet in overleg met het onderwijsveld of door een commissie van experts. Ook moeten er voor de toetsafname duidelijke richtlijnen worden opgesteld, die ook aan de scholen worden gecommuniceerd.

Hoewel de technologie ons optimistisch zou kunnen stemmen over het gemak waarmee toetsaanpassingen zouden kunnen worden ingevoerd, is er in de praktijk allicht nog veel denk- en ontwikkelwerk nodig om toetsaanpassingen in te passen in een toets. Zo zou het bijvoorbeeld kunnen blijken dat leerlingen meer moeite hebben met de voorleessoftware van het toetsplatform in vergelijking met de voorleessoftware die ze zelf gebruiken of dat de ‘speech to text’ knop niet kan worden gebruikt bij teksten die met speciale layout werden opgemaakt. Ook wanneer het duidelijk is welke toetsaanpassingen toegelaten worden bij de gecentraliseerde toetsen, zal het nieuwe steunpunt voor gecentraliseerde toetsen hier in de praktijk toch nog extra tijd voor moeten voorzien.

Speciale onderwijsleermiddelen

Leerlingen die speciale onderwijsleermiddelen gebruiken in de klas omwille van hun auditieve, visuele of motorische handicap hebben deze allicht ook nodig bij de gecentraliseerde toets, indien de overheid beslist om de gecentraliseerde toetsen voor hen toegankelijk te maken. Het kan bijvoorbeeld gaan om het inzetten van een doventolk of ‘schrijver’ van de antwoorden of het gebruik van een brailleregel. Deze speciale onderwijsleermiddelen werden op basis van een medisch attest aan deze leerlingen toegekend.

Eerder is al besproken dat het toegankelijk maken van de toetsen extra werk meebrengt bij de toetsontwikkeling. Ook moeten de speciale onderwijsleermiddelen allicht voorafgaand aan de toetsafname door de school of leerling worden afgestemd met het digitale toetsplatform, zeker als de leerling een eerste keer deelneemt aan de gecentraliseerde toetsen. Sommige van deze leerlingen hebben ook een individuele toetsbegeleider (schrijf- of gebarentolk bijvoorbeeld), wat voor de dag van afname zelf door de school georganiseerd moet worden. Dit kan mogelijk dezelfde persoon zijn die hen ook in het dagelijkse klasgebeuren begeleidt gegeven dat het tot de professionaliteit van deze leerlingondersteuner behoort om louter de antwoorden van de leerlingen weer te geven. Tot slot dient er tijdens de data-

analyse ook rekening te worden gehouden met het feit dat bepaalde leerlingen speciale onderwijsleermiddelen hebben gebruikt tijdens de toets. De eenvoudigste manier om dit te doen is door op descriptief niveau louter weer te geven welk aandeel van de leerlingenpopulatie gebruik heeft gemaakt van bijkomende onderwijsleermiddelen (NCES, 2020). De interpretatie van de resultaten en de mogelijke invloed van de toegekende leermiddelen op die resultaten ligt dan bij de ontvangers van de feedback. Daarnaast kunnen, onder het IRT-raamwerk, via DIF- en multiple-group analyses vaardigheidsscores van leerlingen geschat worden waarbij rekening wordt gehouden met het feit dat leerlingen de toets hebben afgenomen met behulp van bijkomende onderwijsleermiddelen. Een voorwaarde is dat deze groep van leerlingen groot genoeg is, wat allicht niet geldt voor de gecentraliseerde toetsen in Vlaanderen (cf. Tabel 2.2).

Aanvraagplatform

Omdat de overheid de inhoud niet kent van de verslagen die worden opgesteld door de CLB's, lijkt het noodzakelijk dat scholen voorafgaand aan de toetsafname een aanvraag indienen voor (deel)vrijstellingen van toetsen. Indien toetsaanpassingen niet voor alle leerlingen toegankelijk zijn en indien er ook lokale toetsaanpassingen toegelaten worden, dan is er ook een aanvraagprocedure nodig voor toetsaanpassingen. Voor al deze aanvragen is een aanvraagplatform nodig, waarbij alle mogelijke vrijstellingen en alle mogelijke toetsaanpassingen worden opgelijst en per leerling aanvinkbaar zijn. Dit aanvraagplatform moet ontwikkeld en beheerd worden wat ook voor extra opstartkosten zorgt.

Zoals eerder aangegeven, rekenen we op de uitkomst van het juridische luik van Perceel 3 om uitsluitel te krijgen hoe deze persoonsgevoelige informatie van leerlingen binnen het systeem van gecentraliseerde toetsen kan worden opgenomen. De informatie over vrijstellingen en toetsaanpassingen moet niet alleen meegenomen worden bij de afname van de toetsen (en het eventueel opstellen van toetsversies indien er vrijstellingen worden gegeven op itemniveau) maar ook bij de analyses en bij het opstellen van de schoolfeedbackrapporten. Bovendien moet deze informatie worden bijgehouden over afnames heen in het kader van het eventueel meten van leerwinst.

TOEKENNINGSPROCEDURE

Drie scenario's

Er zijn verschillende scenario's mogelijk hoe de aangevraagde (deel)vrijstellingen en toetsaanpassingen voor leerlingen al dan niet gegund worden.

Verantwoordelijkheid school

In een eerste scenario krijgen scholen de verantwoordelijkheid en autonomie om hierover te beslissen. Het hierboven beschreven aanvraagplatform is dan puur een registratieplatform.

In overleg met CLB

In een tweede scenario krijgt het CLB die de school begeleidt een sleutelrol in het aanvragen van (deel)vrijstellingen en toetsaanpassingen. Zij hebben voor leerlingen in Fase 2 en 3 van het zorgcontinuüm toegang tot de individuele leerlingendossiers en de bijhorende (gemotiveerde) verslagen. Voor het multidisciplinaire leerlingdossier geldt echter geheimhoudingsplicht. Het is enkel toegankelijk voor de (relevante) CLB-medewerkers, leerlingen zelf (12+) en de ouders. Dit dossier is nooit toegankelijk voor 'niet-zorgverleners' en dus ook niet voor de scholen. Voor de bijhorende gemotiveerde verslagen geldt een andere regeling. Wanneer er in samenspraak met de ouders een gemotiveerd verslag wordt geschreven door de CLB, geven de ouders ook toestemming om dit verslag te delen met de school waar de leerling zich inschrijft. In tegenstelling tot het multidisciplinaire leerlingdossier geldt hier dus geen geheimhoudingsplicht, al behoren deze leerlinggegevens wel degelijk tot de 'privacy' van de betreffende leerlingen. Voor toetsaanpassingen behorende bij Fase 0 en 1 kan het CLB mogelijk coördinerend optreden om geen (grote) schoolverschillen toe te laten in het gebruik van bijkomende (schooleigen) hulpmiddelen. In dit scenario moeten de scholen en hun CLB samen de aanvragen indienen.

Gunningscommissie

In een laatste scenario richt de overheid een gunningscommissie op, die alle aanvragen tot (deel)vrijstellingen en toetsaanpassingen beoordeelt. Op die manier wordt de uniformiteit van toekenningen zo veel mogelijk nagestreefd. Een belangrijk nadeel van dit scenario vormt echter de omvang van dit werk. Indien we ervan zouden uitgaan dat 10 procent van de leerlingen specifieke onderwijsbehoeften heeft en dus mogelijk een aanvraag voor (deel)vrijstellingen of toetsaanpassingen zou indienen, dan zou dit bij een geschatte populatie van 70000 leerlingen per gecentraliseerde toets neerkomen op 7000 aanvragen. Indien elke aanvraag op 5 à 10 minuten beoordeeld zou kunnen worden (d.i. 12 à 6 aanvragen per uur), dan zou dit per gecentraliseerde toets neerkomen op ongeveer 600 à 1200 uren of 80 à 160 voltijdse werkdagen. Deze dagen moeten allicht gepresteerd worden in de periode maart en april als de gecentraliseerde toetsen rond eind mei (zie het hoofdstuk over de tijdsscenario's) worden afgenomen. Naast de werkomvang op zich is het ook zeer de vraag wie hiervoor als beoordelaar kan worden ingezet.

Buitenlandse voorbeelden

In het buitenland zijn er verschillende praktijken m.b.t. de gunningsprocedure. Bij de eindtoets basisonderwijs in Nederland wordt de verantwoordelijkheid volledig bij de leerkracht en de school gelaten. In een uitgebreid reglement staan de mogelijke aanpassingen opgelijst (Overheid.nl, 2020). In Angelsaksische landen zijn dan weer voorbeelden te vinden van centrale gunning via een toetscommissie (ACARA, 2016). In Nieuw-Zeeland bijvoorbeeld zijn de scholen verantwoordelijk om namens hun leerlingen met speciale behoeften voor het NCEA-examen (National Certificates of Educational Achievement) een aanvraag in te dienen bij het NZQA (New Zealand's Qualifications Authority, 2020).

KOPPELING MET TOETSAFNAMEPLATFORM

Idealiter wordt de informatie uit het aanvraagplatform ontsloten in het toetsafnameplatform, zodat elke leerling automatisch een individueel aangepaste toetsversie aangeboden krijgt bij het inloggen op de toets. Een alternatief scenario is dat de (deel)vrijstellingen en lokale én digitale toetsaanpassingen bij de afname in de scholen zelf door de toetsbegeleiders worden georganiseerd. Dit verhoogt de organisatorische complexiteit van de afname in een school en geeft meer risico's op fouten en mogelijk ook op het toelaten van niet-aangevraagde toetsaanpassingen, en is dus niet aan te raden.

2.4. BREDE AFNAME: ONDERWIJSSETTINGS

De gecentraliseerde toetsen hebben een duidelijk omschreven doelgroep in het reguliere onderwijs. Het principe dat de gecentraliseerde toetsen worden afgenomen bij alle leerlingen van een bepaald leerjaar kan vanuit organisatorisch standpunt ook vertaald worden in de stelregel dat alle scholen die dat leerjaar aanbieden aan de toetsen deelnemen. In deze paragraaf wordt eerst een overzicht geboden van de onderwijssettings waar leerlingen in Vlaanderen school kunnen lopen. Daarna worden scenario's besproken met betrekking tot de eventuele deelname van niet-reguliere scholen. Hierbij worden budgettaire implicaties gegeven en kanttekeningen geplaatst.

2.4.1. OVERZICHT VAN DE ONDERWIJSSETTINGS

LEERLINGENAANTALLEN PER AFNAMEMOMENT

Inleiding

Leerlingen kunnen in Vlaanderen schoollopen in vier onderwijssettings. De meeste leerlingen volgen onderwijs in een erkende school voor regulier of buitengewoon onderwijs. Voor regulier onderwijs zijn er in Vlaanderen 2294 basisscholen en 942 secundaire scholen. Voor het buitengewoon onderwijs zijn er respectievelijk 192 en 125 scholen in het basis- en secundair onderwijs. Er zijn daarnaast ook 14 erkende privéscholen in Vlaanderen. De leerlingen van deze scholen kunnen een getuigschrift of diploma verwerven via de examencommissie. Tot slot volgt een kleine groep leerlingen huisonderwijs. Binnen deze groep vallen ook de leerlingen die naar een niet-erkende privéschool gaan, wat ook wel collectief huisonderwijs genoemd wordt. Ook deze leerlingen behalen een diploma of getuigschrift via de examencommissie. In de onderstaande paragrafen worden de leerlingenaantallen per onderwijssetting besproken. Over de leerlingenaantallen in privéscholen zijn geen cijfers beschikbaar.

Reguliere onderwijs

Tabel 2.3 geeft een overzicht van het aantal leerlingen binnen het reguliere onderwijs voor de vier vastgelegde leerjaren waarin er nu gecentraliseerde toetsen gepland zijn op basis van de tellingen voor het schooljaar 2019-2020. Daarbij wordt een onderscheid gemaakt tussen leerlingen (a) met een gemotiveerd verslag; (b) een inschrijvingsverslag; (c) een verslag en (d) leerlingen zonder een van

Tabel 2.3: Aantal leerlingen met gemotiveerd verslag, verslag, en inschrijvingsverslag en het aantal reguliere leerlingen in het gewoon onderwijs in het schooljaar 2019-2020 voor de leerjaren waarin afnames van gecentraliseerde toetsen gepland zijn.

Doelgroep	Gemotiveerd verslag	Inschrijvingsverslag	Verslag	Reguliere leerlingen	Totaal
<i>Gewoon Lager Onderwijs</i>					
4de leerjaar	3324	183	252	70886	74647
6de leerjaar	2489	582	81	67315	70467
<i>Voltijds Gewoon Secundair Onderwijs</i>					
2de jaar SO	3073	1721	485	67507	72788
2e leerjaar eerste graad GSO A	1766	522	44	58258	60590
Beroepsvoorbereidend leerjaar	1307	1199	441	9249	12198
2de leerjaar derde graad	781	2320	30	56107	59238
ASO	160	333	3	24484	24980
BSO	304	1347	20	11422	13093
KSO	43	57	2	1373	1475
TSO	274	583	5	18828	19690
2de leerjaar modulair BSO	1	12		124	137
<i>Deeltijds Beroepssecundair Onderwijs</i>	516	1061	383	7011	8975

Cijfers opgevraagd bij Ministerie van Onderwijs, cijfers schooljaar 2019-2020

voorgaande verslagen. Een inschrijvingsverslag is nog een oude vorm van verslag van voor de start van het M-decreet. Niet al de leerlingen met een inschrijvingsverslag worden vandaag nog beschouwd als leerlingen met speciale onderwijsbehoeften. Een inschrijvingsverslag kan bijvoorbeeld nog dateren van een vroeger verblijf in het buitengewoon onderwijs, of van een vorig onderwijsniveau.

Tabel 2.3 geeft een indicatie van de verwachte aantal leerlingen dat zal deelnemen aan de gecentraliseerde toetsen, al moet er daarbij ook rekening gehouden worden met toekomstige demografische schommelingen die tot een lichte toename van het aantal leerlingen per toets zullen leiden.

De leerlingen met een gemotiveerd verslag omvatten 4 procent van de totale leerlingengroep in de twee leerjaren van het basisonderwijs en in de totale eerste graad secundair onderwijs. In die eerste graad secundair onderwijs heeft in het beroepsvoorbereidend leerjaar evenwel 11 procent van de leerlingen een gemotiveerd verslag. In de derde graad varieert het percentage leerlingen met een gemotiveerd verslag tussen de onderwijsvormen. Het gaat om 1 procent van de leerlingen in het aso en tso; 2 procent in het bso en 3 procent in het kso.

Ongeveer 3 op de 1000 leerlingen in het vierde leerjaar hebben een verslag. In het zesde leerjaar is dat 1 op de 1000. In de eerste graad secundair onderwijs stijgt dit tot 7 op de 1000 leerlingen, waarvan 3,6 procent van de leerlingen in het beroepsvoorbereidend leerjaar. In de derde graad komt een verslag bijna niet voor, behalve in het bso waar het 2 op de 1000 leerlingen betreft.

Buitengewoon onderwijs

Voor de leerlingen van het buitengewoon onderwijs is het moeilijk om op basis van de gegevens uit het statistisch jaarboek van het Vlaamse onderwijs (2020) schattingen te maken over het aantal leerlingen dat mogelijk aan de gecentraliseerde toetsen zou deelnemen. De leerlingen in het buitengewoon onderwijs zijn immers niet altijd aan een leerjaar toe te wijzen (cf. infra).

Heel erg benaderend zouden we kunnen stellen dat er 4300 leerlingen in het buitengewoon onderwijs in elk leerjaar van het basisonderwijs zitten; 3400 in de eerste graad secundair onderwijs en 890 op het einde van de derde graad secundair onderwijs (OV4).

Niet alle leerlingen in het buitengewoon onderwijs behoren tot de doelgroep van de gecentraliseerde toetsen omdat ze niet allemaal de Vlaamse onderwijsdoelen hoeven te bereiken of hoeven na te streven. Voor de leerlingen uit Type 2 (verstandelijke handicap) en Onderwijsvorm (OV) 1 (sociale aanpassing) en 2 (arbeidsgeschiktmaking) in het secundair onderwijs worden geen eindtermen vooropgesteld. Voor Type 1 (licht mentale handicap), Basisaanbod, en OV3 (buitengewoon beroepsonderwijs)¹⁰ gelden er ontwikkelingsdoelen die niet overeenkomen met de eindtermen. Als de gecentraliseerde toetsen peilen naar het behalen van de eindtermen, kan men dus niet verwachten dat leerlingen waarbij deze leerdoelen niet vooropgesteld waren deze toetsen zullen afnemen. De leerlingen uit Type 3, Type 4, Type 6, Type 7 en Type 9 (zie Tabel 2.4 voor een omschrijving) komen mogelijk wel in aanmerking voor de gecentraliseerde toetsen. In het secundair onderwijs kunnen deze leerlingen doorstromen naar OV4 (secundair onderwijs). De leerlingen uit Type 8 (ernstige leerstoornissen) worden verder niet meegenomen omdat dit type momenteel in afbouw is.

In Tabel 2.4 wordt per type het aantal leerlingen geschat die binnen het buitengewoon onderwijs als doelpubliek voor de vier afnamemomenten van de gecentraliseerde toetsen in aanmerking zouden komen. Dit betekent niet dat al deze leerlingen alle toetsen zonder meer kunnen afleggen, zoals hieronder nog zal worden besproken. Zo leert de ervaring bij het Vlaamse luik van het internationaal vergelijkend onderzoek TIMSS dat het opnemen van leerlingen uit Type 9 door de gecontacteerde scholen als allerm minst evident en vaak als onwenselijk beschouwd werd omdat deze leerlingen een ernstige beperking ondervinden bij stressvolle of nieuwe situaties. Dit moet nog verder onderzocht worden.

De leerlingen uit Type 5 (ziekenhuisschool) worden in de tellingen van het statistisch jaarboek niet opgenomen bij het buitengewoon onderwijs om dubbeltellingen te vermijden, gegeven dat deze kinderen en jongeren doorgaans ook nog ingeschreven zijn in een andere school. Gezien hun bijzondere en tijdelijke situatie is het maar de vraag of leerlingen die opgenomen zijn in de ziekenhuisschool kunnen deelnemen aan de gecentraliseerde toetsen dan wel dat ze in principe worden vrijgesteld.

¹⁰ Vanuit de stuurgroep werd opgemerkt dat na de uitrol van de modernisering in OV3 een onderwijskwalificatie Niveau 2 (getuigschrift tweede graad) behaald kan worden. Om die reden zou dan ook binnen OV3 ook gecentraliseerde toetsen kunnen worden afgenomen (al gaat het hier dan allicht om de toetsen van de eerste graad secundair onderwijs). De discussie over de deelname van het buitengewoon onderwijs wordt nog verder gevoerd in § 2.4.2 hieronder.

Tabel 2.4: *Ruwe schatting van het aantal leerlingen dat in aanmerking zou kunnen komen voor de gecentraliseerde toetsen in het buitengewoon onderwijs.*

Type	Doelgroep	Lager onderwijs		Secundair onderwijs
		4 ^{de} /6 ^{de} leerjaar	2 ^{de} leerjaar	OV4
3	Emotionele of gedragsstoornis	300	104	40
4	Motorische beperking	170	65	25
6	Visuele beperking	20	5	2
7	Auditieve beperking	200	13	5
9	Autismespectrumstoornis	850	703	268

Bron: Schatting o.b.v. gegevens uit Statistisch jaarboek van het Vlaams onderwijs (2020)

Huisonderwijs

Op basis van de gegevens uit het Statistisch jaarboek van het Vlaams onderwijs (2020) kan men gegeven het totaal aantal leerlingen in het basisonderwijs die huisonderwijs volgen, schatten dat er in het huisonderwijs ongeveer 200 leerlingen in het zesde leerjaar zitten (of tenminste de examens bij de examencommissie afleggen) en allicht minder dan 200 leerlingen in het vierde leerjaar. Volgens dat jaarboek zouden op het einde van de eerste graad 240 leerlingen die huisonderwijs volgen examens afleggen en op het einde van de derde graad is dat aantal 460 leerlingen.

OVER AFNAMEMOMENTEN HEEN

Zoals vermeld, toont Tabel 2.3 enkel een overzicht van de leerlingen op één afnamemoment. Daarnaast moeten de leerlinggegevens over afnamemomenten kunnen worden gekoppeld om zo een longitudinale opvolging van leerlingen (bijvoorbeeld voor het meten van leerwinst) mogelijk te maken. De leerlingmobiliteit van leerlingen tussen scholen, onderwijsvormen en mogelijk ook onderwijssettings vormt hierbij bijkomende uitdagingen.

2.4.2. SCENARIO'S MET BETREKKING TOT ONDERWIJSSETTINGS

REGULIER ONDERWIJS

Het spreekt voor zich dat de leerlingen uit het regulier onderwijs behoren tot de doelgroep van de gecentraliseerde toetsen. Alle leerlingen van een bepaald leerjaar leggen daarbij de gecentraliseerde toetsen af. Voor scholen die niet met een traditioneel jaarsysteem werken zou dit criterium voor deelname mogelijk niet eenduidig kunnen zijn op klasniveau, al kan de school allicht wel op leerlingniveau aangeven welke toetsen deze leerlingen moeten afleggen. Deze problematiek moet mogelijk extra voorzien worden in het systeem van het gunnen van (deel)vrijstellingen.

Mogelijk zullen ook scholen die een afwijking hebben op het opgelegde onderwijscurriculum (zoals op dit moment de Steinerscholen voor de huidige eindtermen) om een uitzondering vragen op de verplichte deelname aan de gecentraliseerde toetsen. Het is aan de overheid om hierin een beslissing te nemen als die situatie zich zou voordoen.

BUITENGEWOON ONDERWIJS

Hoewel het onderwijs voor een aantal leerlingen in het buitengewoon onderwijs gericht is op de Vlaamse onderwijsdoelen, stelt zich de vraag of deze scholen ook behoren tot de doelgroep van de gecentraliseerde toetsen. De mate waarin de huidige paralleltoetsen die door STEP worden aangeboden als toetsinstrument kunnen dienen in het buitengewoon onderwijs vormt hier een interessant vergelijkingspunt. Hierover hield STEP op 11 februari 2020 een uitgebreid focusgesprek met partners uit het buitengewoon onderwijs. Het verslag van dit overleg werd opgeleverd aan de dienst SBO van het onderwijsministerie en is ook als Bijlage 2 bij dit rapport te vinden. De focusgroep was duidelijk van oordeel dat, hoewel de paralleltoetsen voor bepaalde, individuele leerlingen mogelijk zinvol kunnen zijn (bijvoorbeeld in het kader van doorstromingsbeslissingen), de paralleltoetsen beter niet kunnen worden ingezet in het buitengewoon onderwijs. De argumenten die daarbij werden aangehaald gelden ook voor de gecentraliseerde toetsen. Het buitengewoon onderwijs geeft onderwijs op maat, waarbij niet alle leerlingen dezelfde onderwijsdoelen nastreven. Het aantal leerlingen dat in aanmerking komt voor een afname van een gestandaardiseerde toets over de eindtermen is te beperkt, waardoor het gebruik van deze toetsen voor de evaluatie van de eigen werking op klas- en schoolniveau weinig zinvol is. Het werd ook niet haalbaar geacht om vergelijkende schoolfeedback te bieden aan de scholen in het buitengewoon onderwijs. Tot slot vraagt de afname van toetsen in het buitengewoon onderwijs doorgaans om differentiatie in de afname-instructies, waardoor een gestandaardiseerde afname moeilijk haalbaar is.

Een aantal van deze argumenten (onderwijs op maat, beperkte mogelijkheden tot klas- of schoolfeedback) lijken er voor te pleiten dat de leerlingen in het buitengewoon onderwijs in eerste instantie niet worden gerekend tot de doelgroep van de gecentraliseerde toetsen, zoals die nu met het doelpubliek in het reguliere onderwijs worden ontworpen. Dit ligt in lijn met buitenlandse toetspraktijken. In Nederland, bijvoorbeeld, doen de scholen in het speciaal basisonderwijs niet mee aan de centrale eindtoets. Leerlingen met een IQ lager dan 75 zijn ook bij wet vrijgesteld van de toets (Overheid.nl, 2020). Leerlingen met een handicap zijn ook vrijgesteld van de Australische NAPLAN-toetsen.

Indien toch zou gekozen worden om scholen uit het buitengewoon onderwijs te laten deelnemen, bijvoorbeeld in het kader van de opvolging van individuele leerlingen, dan impliceert dit drie zaken. Ten eerste, zullen er zeker bijkomend middelen moeten worden ingezet om de toegankelijkheid van de toetsen te verzekeren (cf. o.a. de leerlingen met speciale onderwijsleermiddelen) en mogelijk ook om voldoende computerinfrastructuur voor de digitale toetsafnames te voorzien binnen het buitengewoon onderwijs. Ten tweede, zal het aantal aanvragen verhogen naar (deel)vrijstellingen en toetsaanpassingen binnen het systeem dat ook al voor leerlingen uit het reguliere onderwijs zal moeten worden voorzien.

Ten slotte zal er binnen het nieuwe steunpunt nieuwe sjablonen voor schoolfeedback moeten ontworpen worden.

Merk ook op dat er in de aangehaalde focusgroep ook vraag was naar toetsinstrumenten om binnen het buitengewoon onderwijs in te zetten in het kader van interne kwaliteitszorg. De vraag is of dit gerealiseerd kan worden met de gecentraliseerde toetsen die ontwikkeld worden voor het reguliere onderwijs, eventueel met specifieke afnamemodaliteiten (bijvoorbeeld breder tijdsvenster, afname op het moment dat de klassenraad geschikt acht), dan wel dat de overheid beter investeert in hiervoor speciaal ontworpen toetssystemen. Desgevallend moet hierover in het nieuwe steunpunt over worden nagedacht.

HUIS- EN PRIVÉ-ONDERWIJS

Een relatief klein, maar de laatste jaren wel stijgend aantal leerlingen volgt huisonderwijs. Zij behalen hun getuigschrift of diploma via de examencommissie. Ook leerlingen van erkende of niet-erkende privéscholen richten zich tot deze instantie. Op die manier zijn al deze leerlingen al onderworpen aan door de overheid opgelegde toetsen over de Vlaamse eindtermen voor het bepalen van hun individuele prestaties. Deelname aan de gecentraliseerde toetsen voor individuele leerlingenfeedback lijkt dan misschien niet nodig. Mogelijk kan het wel interessant zijn om ook van deze leerlingen data te hebben in het kader van de leerwinstmetingen binnen de gecentraliseerde toetsen, ook al gaat het om een kleine groep.

Merk op dat vooral voor de gecentraliseerde toetsen van het vierde leerjaar in het basisonderwijs het ook moeilijk af te bakenen zou zijn welke leerlingen hiervoor in aanmerking komen, gezien er voor deze leerlingen geen specifieke examens binnen de examencommissie worden voorzien en het niet meteen in lijn ligt met de opzet van de gecentraliseerde toetsen om de doelgroep daarvan op basis van leeftijd te definiëren. Momenteel moeten leerlingen die huisonderwijs volgen hun getuigschrift basisonderwijs behalen door te slagen op de examens in een examenschool. Er zijn daarbij twee examenscholen basisonderwijs per provincie. Die examens worden georganiseerd in juni. De examenschool bepaalt hoe de evaluatie verloopt. Mogelijk kan voor deze groep leerlingen ook deelname aan de gecentraliseerde toetsen voor het zesde leerjaar overwogen worden. Voor wat betreft het secundair onderwijs leggen leerlingen wel examens af bij de examencommissie, al vallen deze examens van de examencommissie niet allemaal op het einde van het schooljaar, waardoor leerlingen op verschillende momenten in het jaar hun diploma of getuigschrift behalen in tegenstelling tot wat het geval is in het reguliere onderwijs.

Voor de leerlingen die huisonderwijs genieten is schoolfeedback uiteraard niet aan de orde. Mogelijk zou de overheid wel geïnteresseerd kunnen zijn om de prestaties van deze leerlingen te vergelijken met die van de leerlingen in het reguliere onderwijs. Allicht is het ook niet aan de overheid om schoolfeedback te voorzien voor (niet-)erkende privéscholen en zijn deze scholen gegeven hun privé-karakter hiervoor ook geen vragende partij. Ook vanuit dit doel van de gecentraliseerde toetsen vallen deze leerlingen dus buiten de scope van de gecentraliseerde toetsen. Ook in internationaal vergelijkend onderzoek worden de Vlaamse privéscholen niet meegerekend bij de omschrijving van de doelpopulatie.

Men zou toch kunnen kiezen om de leerlingen die de examencommissie zijn ingeschreven voor de examens van het lager onderwijs (zesde leerjaar, cf. supra), de eerste graad of de derde graad secundair onderwijs te laten deelnemen aan de gecentraliseerde toetsen. In overeenstemming met de organisatie van de gecentraliseerde toetsen door de scholen in het reguliere onderwijs, zou deze afname georganiseerd kunnen worden door de examencommissie. Zij en AGODI registreren welke leerlingen huisonderwijs volgen. Het afnamevenster van de gecentraliseerde toetsen zou dan ook moeten gelden voor deze leerlingen. De toetsafname van de gecentraliseerde toetsen voor deze leerlingen kan dan in Brussel gebeuren, zoals dat ook het geval is voor de examens van de examencommissie. De examencommissie beschikt daar over verschillende pc-classes. Deze moeten dan vrijgeroosterd worden voor de afname van de gecentraliseerde toetsen. Indien er onvoldoende pc's zouden zijn voor de afname van een welbepaalde toets, dan kunnen ofwel laptops gehuurd worden ofwel kan gewerkt worden met het principe 'bring your own device', zoals in Perceel 3 wordt uitgewerkt. De bijkomende organisatorische kosten zouden vanuit de werking van de examencommissie kunnen worden gedragen, net zoals er inspanningen van de reguliere deelnemende scholen gevraagd worden. Een ander mogelijk scenario is dat het steunpunt zorgt voor de afnames van de gecentraliseerde toets bij de huisgeschoolde leerlingen. In Nieuw-Zeeland moeten deze leerlingen voor het afleggen van het NCEA examen contact opnemen met een school in de buurt (NZQA, 2020), wat allicht ook samenhangt met de zeer lage gemiddelde bevolkingsdichtheid van dit land (16,4 inwoners per km²), o.a. in vergelijking met Vlaanderen (462 inwoners per km²), waardoor het minder evident is om te werken met een of meerdere toetsafnamecentra.

2.4.3. BIJKOMENDE OVERWEGINGEN

Het is de taak van de overheid om de scholen op te roepen om deel te nemen aan de gecentraliseerde toetsen en hen hierover te informeren. In geval dat het huisonderwijs ook wordt onderworpen aan de gecentraliseerde toetsen moet ook individuele leerlingen opgeroepen worden. Deze kost wordt niet verder begroot.

Indien er gekozen wordt om ook scholen buiten het reguliere onderwijs te laten deelnemen dan leidt dit boven op de budgettaire implicaties die hierboven al werden aangegeven tot bijkomend werk voor het nieuwe steunpunt, zowel organisatorisch, als qua analyses en schoolfeedback. Zo zijn er bijvoorbeeld voor deze scholen bijkomende feedbacksjablonen nodig, gegeven dat schoolfeedback voor deze doelgroep niet standaard kan verlopen. Deze extra werkzaamheden zijn nodig voor de opstart van deze bijkomende doelgroep maar komen ook jaarlijks terug voor elk van de gecentraliseerde toetsen.

2.5. OVERZICHT DEELSCENARIO'S EN BELEIDSADVIEZEN

2.5.1. OVERZICHT DEELSCENARIO'S

In dit hoofdstuk werd ingegaan op de brede afname van de gecentraliseerde toetsen. In eerste instantie werden vanuit juridisch, onderwijskundig en toetstechnisch perspectief argumenten aangehaald om een brede afname bij de gecentraliseerde toetsen te beogen. Daarna werd eerst vanuit het perspectief van de deelnemende leerlingen en vervolgens vanuit het perspectief van de deelnemende scholen in concreto uitgewerkt wat een brede afname op organisatorisch vlak met zich meebrengt. Doorheen de tekst werden daarbij verschillende beleidskeuzes voorgesteld, die opgelijst staan in Tabel 2.5. In de volgende paragraaf worden deze keuzes vanuit beleidsperspectief besproken.

Tabel 2.5.

Beleidskeuzes voor de deelscenario's rond brede afname van de gecentraliseerde toetsen.

Deelscenario	Keuzeopties
<i>(Deel)vrijstellingen</i>	
Voor wie?	leerlingen met IAC (ja; neen) OKAN-leerlingen (ja; neen)
Op welke manier?	op alle toetsen (zonder inhoudelijke motivatie) met deelvrijstellingen - op toetsniveau - op itemniveau
Opt out?	ja; neen;
<i>Toetsaanpassingen</i>	
Voor wie?	alle leerlingen leerlinggroepen (individuele aanvraag)
Door wie bepaald?	oplijsting door experts (top-down) in samenspraak met onderwijsveld (bottom-up)
<i>Toekenningsprocedure</i>	
	school zelf CLB toekenningscommissie
<i>Toegankelijkheid</i>	
Voor leerlingen met een visuele beperking	ja; neen;
Voor leerlingen met een auditieve beperking	ja; neen;
Voor leerlingen met een motorische beperking	ja; neen;
<i>Onderwijssettings</i>	
Deelname buitengewoon onderwijs	ja; neen; aangepaste vorm van gecentraliseerde toetsen?
Deelname privé-scholen	ja; neen;
Deelname huisonderwijs	ja; neen;

2.5.2. BELEIDSADVIEZEN

De beleidskeuzes die moeten worden genomen rond het verlenen van (deel)vrijstellingen en toetsaanpassingen zijn in eerste instantie afhankelijk van de resultaten van het juridische luik van de haalbaarheidsstudie (cf. Perceel 3). In dat perceel moet worden uitgeklaard of leerlingen of hun ouders ervoor kunnen kiezen om niet deel te nemen aan de gecentraliseerde toetsen (cf. opt-out). Perceel 3 moet daarnaast ook verduidelijken of en op welke manier de privacy-gevoelige informatie rond de inhoud van verslagen of gemotiveerde verslagen kan worden gedeeld met de overheid of het steunpunt zodanig dat aan individuele leerlingen op inhoudelijke grond (deel)vrijstellingen of toetsaanpassingen kan worden verleend. Indien de inhoud van deze verslagen beschermd is door de wet op de privacy of door het beroepsgeheim dan is het enerzijds mogelijk om louter algemene maatregelen te nemen die inhoudelijk niet gericht zijn op de concrete leerling, maar gelden voor alle leerlingen met een (gemotiveerd) verslag. Leerlingen met een IAC kunnen dan vrijstelling bekomen op alle gecentraliseerde toetsen. De toetsen worden dan zoveel mogelijk ontworpen binnen een ‘universal design for assessments’ en afgenomen met toetsaanpassingen (bv. voorleessoftware) die door alle leerlingen kunnen worden gehanteerd. Bij de analyses achteraf kan bij het opstellen van schoolfeedback of het rapporteren op systeemniveau wel worden rekening gehouden met het feit of een leerling al dan niet een gemotiveerd verslag had.

Anderzijds is het in die situatie mogelijk om de gunning van (deel)vrijstellingen en toetsaanpassing aan de autonomie van de school over te laten, eventueel in samenspraak met het CLB (op voorwaarde dat deze informatie op een GDPR-conforme wijze kan worden doorgegeven binnen het toetsplatform). Dit scenario is duidelijk minder kostelijk dan het installeren van een toekenningscommissie, maar heeft als nadeel dat er verschillen tussen scholen kunnen zijn in hun toekenningsbeleid, wat de noodzakelijke standaardisatie van de afname van de gecentraliseerde toetsen in het gedrang kan brengen. Zeer duidelijke instructies aan deelnemende scholen over wat wel en niet mogelijk is, kan hiervoor een oplossing bieden. Zo gebeurt dit ook binnen de Eindtoets basisonderwijs in Nederland.

Alleszins zal de overheid moeten bepalen welke toetsaanpassingen bij de afname van de gecentraliseerde toetsen worden toegelaten. Dit is vooral een inhoudelijke discussie die deels ook raakt aan didactische opvattingen. De overheid kan opteren om de lijst van toetsaanpassingen te laten opstellen door experts (top down), dan wel om op basis van bevraging van en in overleg met het onderwijsveld deze lijst op te stellen (bottom up).

Een aantal van de te nemen beleidsbeslissingen rond brede afname worden best genomen in overleg met psychometrische experts (cf. ook Perceel 1). Zo is het verre van duidelijk of het haalbaar is om leerlingen met een IAC deelvrijstellingen te geven op items van een toets zonder dat dat de meetkwaliteit in het gedrang brengt. Ook moet in principe worden nagegaan in welke mate de itemkenmerken van toegankelijk gemaakte items nog vergelijkbaar zijn met de oorspronkelijke items, al is die controle door de kleine leerlingenaantallen van leerlingen met een auditieve, visuele of motorische beperking allicht voor de Vlaamse gecentraliseerde toetsen in de praktijk niet mogelijk. Merk op dat het beperkt aantal leerlingen met een auditieve, visuele of motorische beperking leidt tot een relatief grote kostprijs per

leerling voor het toegankelijk maken van items. De beslissing over het al dan niet toegankelijk maken van items is echter ook een principiële beslissing rond inclusie, die mogelijk door leerlingen of hun ouders langs juridische weg zou kunnen worden afgedwongen. Deze laatste problematiek werd in de looptijd van het project aan Perceel 3 gesignaleerd.

Tot slot is het aan de onderwijsminister om te beslissen over de deelname van scholen buiten het reguliere onderwijs aan de gecentraliseerde toetsen. Ook hier zou het beperkte aantal leerlingen waarover dit gaat in het privé-onderwijs of huisonderwijs mee als keuzeprincipe gehanteerd kunnen worden, gegeven ook de budgettaire kost om de afnames te organiseren in de schoot van de examencommissie. Voor het buitengewoon onderwijs zijn er argumenten gegeven waarom het niet wenselijk is om de gecentraliseerde toetsen daar op grote schaal af te nemen. Dit neemt niet weg dat de gecentraliseerde toetsen mogelijk zinvol kunnen zijn bij loopbaanbeslissingen rond individuele leerlingen. Daarnaast is er mogelijk ook een vraag vanuit scholen uit het buitengewoon onderwijs naar indicatoren om in te zetten in het kader van hun interne kwaliteitszorg.

3. TIJDSSCENARIO'S

3.1. INLEIDING

In het vorige hoofdstuk stond centraal *wie* de gecentraliseerde toetsen aflegt. In dit hoofdstuk staat de *wanneer*-vraag centraal. Daarbij worden scenario's voor de verschillende tijdsaspecten van de gecentraliseerde toetsen besproken. Eerst komt de planning van de afnameperiode in de loop van het schooljaar aan bod. Daarna volgen scenario's voor het afnamevenster. Het afnamevenster is de tijdspanne waarin alle leerlingen van één bepaald leerjaar de toets moeten afgelegd hebben. Tot slot wordt de doorlooptijd besproken. De doorlooptijd is de tijd die nodig is nadat de toetsen zijn afgelegd om de resultaten bekend te maken.

Bij dit hoofdstuk zal bij het ontwikkelen van de diverse tijdsscenario's worden verwezen naar verschillende buitenlandse voorbeelden en naar de gecentraliseerde toetsen in de Franse Gemeenschap. Deze worden in Tabel 3.1 kort voorgesteld. Merk op dat de besproken gecentraliseerde toetsen in het buitenland in het basisonderwijs doorgaans schoolondersteunend zijn. Op leerlingenniveau wordt in Denemarken en Australië ook leerwinst gemeten. In Nederland leidt de Centrale eindtoets op het einde van het primaire onderwijs tot een toetsadvies dat naast het schooladvies wordt gebruikt voor de advisering over de onderwijsvorm in het voortgezet onderwijs. In de Franse Gemeenschap is het CEB samen met de schoolresultaten determinerend voor het al dan niet halen van het getuigschrift basisonderwijs. De getoetste leergebieden omvatten steevast rekenen en aspecten van moedertaalonderwijs. Andere vakken komen minder vaak aan bod. De toetsen worden op papier of digitaal afgenomen. In Nederland en Denemarken worden de digitale toetsen adaptief aangeboden. De voorbeelden voor het secundair onderwijs zijn gericht op laatstejaars die gecentraliseerde examens afleggen voor verschillende vakgebieden. In Nieuw-Zeeland werkt men daarnaast ook met praktische proeven, zowel voor de 'internal assessments' (voor het verwerven van credits) als voor de 'external assessments' (voor het verwerven van een diploma).

In Tabel 3.2 worden de afnameperiode, het afnamevenster en de doorlooptijd van deze buitenlandse voorbeelden in kaart gebracht. Daarnaast wordt verderop in deze tekst nog verwezen.

Tabel 3.1
Overzicht buitenlandse praktijkvoorbeelden

Land	Toets	Toetsdoel	D/P ^a	Doelgroep	Leer- of vakgebied
Basisonderwijs					
B ^b	CEB	Certificering	P	laatstejaars	wiskunde, Frans, wereldoriëntatie
NL	Centrale Eindtoets ^c	Studie-oriëntering	P/DA	laatstejaars	lezen, taalverzorging, schrijven, rekenen en eventueel wereldoriëntatie (naar keuze)
DK	10 verplichte nationale testen	Meten leerwinst Ondersteunend	D A	Groep 2, 4, 6, 8: Groep 3 en 6: Groep 7 en 8:	lezen wiskunde andere vakken
AUS	NAPLAN	Metten leerwinst - Ondersteunend	D	3 ^e , 5 ^e , 7 ^e , 9 ^e leerjaar	schrijven, lezen, 'language conventions' (grammatica, spelling) en rekenen
Secundair onderwijs					
B	CED1	Certificering (samen met schoolexamens)	P	Eind graad	eerste Frans, wiskunde, vreemde talen, wetenschappen
B	CESS	Certificering (samen met schoolexamens)	P en mondeling	Eind graad	derde Frans ("qualification" of "transition") en geschiedenis
NL	Centrale eindexamens	Diplomering (samen met schoolexamens)	P	laatstejaars	Nederlands, Engels, wiskunde, algemene vakken volgens studiekeuze
FIN	Matriculation exam	Diplomering	D	laatstejaars	alle algemene vakken, min. 4 te kiezen volgens studiekeuze, Fins is verplicht
NZ	NCEA internal assessments	Certificering (credits)	D, P en praktische proeven	laatstejaars	alle vakken
NZ	NCEA external assessments	Diplomering (centrale examens)	D, P en praktische proeven	laatstejaars	alle vakken

CEB = Certificat d'études de base; CED1 = Certificat d'études du 1er degré; CESS = Certificat d'enseignement secondaire supérieur; NAPLAN = National Assessment Program — Literacy and Numeracy; NCEA = National Certificate Educational Achievement

^a D = digitaal; P = papier; A= adaptief

^b Franse Gemeenschap van België

^c Scholen kunnen ook kiezen voor een andere eindtoets die door de minister is toegelaten

bronnen: ACARA, 2016 ; Beuchert & Nandrop, 2017; College voor Toetsen en Examen (2020); Enseignement.be (2021); Ministry of education, 2020a; 2020b; Shewbridge et al., 2011; Ylioppilastutkinto, 2020.

Tabel 3.2

Overzicht van diverse tijdsaspecten van de buitenlandse praktijkvoorbeelden

Land	Toets	Afname- periode	Afnamevenster	Inhaalmoment	Herkansing	Doorlooptijd
Basisonderwijs						
B	CEB	Einde schooljaar	1 vast moment per vak	neen	neen	1 week
NL	Centrale eindtoets	15 april – 15 mei	P: Vaste dag per vak; DA: 14 dagen	enkel digitaal	neen	enkele weken (mei)
DK	10 verplichte nationale testen	Midden januari tot april	naar keuze; Test hoger/ lager dan doel-niveau kan op elk moment	juni	neen	1 dag
AUS	NAPLAN	4 ^{de} maand van het schooljaar	3 vaste dagen	dag 4	neen	--
Secundair onderwijs						
B	CED1	einde schooljaar	1 vast moment per vak	neen	neen	1 week
B	CESS	einde schooljaar	1 vast moment per vak	neen	neen	1 week
NL	Centrale eindexamens	einde schooljaar	1 vast moment per vak;	idem	idem	binnen de maand
FIN	Matriculation exam	maart	1 vast moment per vak	Valt samen met herkansing	1 vast moment per vak in september	4 maanden? vanaf 20 juli/ 20 dec. kan men test inkijken
NZ	NCEA internal assessments	doorheen schooljaar	wisselend	--	--	--
NZ	NCEA external assessments	einde schooljaar (november)	vaste dag	--	--	Drie maanden (februari)

CEB = Certificat d'études de base; CED1 = Certificat d'études du 1er degré; CESS = Certificat d'enseignement secondaire supérieur; NAPLAN = National Assessment Program — Literacy and Numeracy; NCEA = National Certificate Educational Achievement ; -- = deze informatie was niet af te leiden uit de beschikbare documenten.

bronnen: ACARA, 2016 ; Beuchert & Nandrop, 2017; College voor Toetsen en Examens (2020); Enseignement.be (2021); Ministry of education, 2020a; 2020b; Shewbridge et al., 2011; Ylioppilastutkinto, 2020.

3.2. PLANNING AFNAME(S)

3.2.1. PLANNING IN HET SCHOOLJAAR

Zoals uit Tabel 3.2 blijkt, zijn alle afnames van gecentraliseerde toetsen in de buitenlandse voorbeelden voor het secundair onderwijs gepland op het einde van het schooljaar. Dat hoeft natuurlijk niet te verbazen gegeven hun toetsdoel van certificering. Voor het basisonderwijs liggen de afnames in het derde trimester van het schooljaar. In Australië worden de NAPLAN-toetsen evenwel al na vier maanden in het schooljaar afgenomen.

Gezien de gecentraliseerde toetsen in Vlaanderen de onderwijsdoelen voor wiskunde en Nederlands bevragen, lijkt het voor de hand te liggen dat de afname ervan tegen het einde van het schooljaar wordt gepland. Grofweg zouden daarbij drie periodes kunnen worden onderscheiden: (a) van 15 april tot 15 mei (of in de periode vlak na de paasvakantie); (b) van 15 mei tot half juni (voor de examenperiode die door de scholen zelf wordt ingericht) en (c) op het einde van juni (na de examenperiode). Bij de afweging van deze verschillende scenario's kunnen de volgende overwegingen spelen:

- De keuze voor de periode hoeft niet per se gelijk te lopen voor het basisonderwijs en het secundair onderwijs. Binnen dit laatste onderwijsniveau zouden de toetsen voor de tweedejaars ook niet op hetzelfde moment hoeven te vallen als die voor de zesdejaars.
- Hoe later op het schooljaar de gecentraliseerde toetsen worden afgenomen, hoe meer leeransen de leerlingen krijgen om deze toetsen op een goede manier af te leggen.
- Wanneer de toetsen na de schoolexamens worden afgenomen, hebben leerlingen door dat evaluatiemoment op school mogelijk het leer- of vakgebied beter onder de knie. Daar staat tegenover dat de toetsmotivatie van leerlingen na de schoolexamens lager kan zijn. Scholen en leerkrachten zijn tegen het einde van het schooljaar ook zelf druk bezig met de leerlingevaluaties en –deliberaties, zodat dit afnamemoment vanuit dat oogpunt ook minder geschikt lijkt.
- De keuze van het afnamemoment dient ook afgestemd te worden op de benodigde doorlooptijd als het zo is dat minstens bepaalde informatie van de gecentraliseerde toetsen beschikbaar moet zijn tegen het einde van het schooljaar. Dit lijkt gesuggereerd te worden in beleidsdocumenten waarin gestipuleerd wordt dat de resultaten op de gecentraliseerde toetsen niet de enige bron mag zijn waarop een klassenraad zich baseert bij beslissingen over certificering of diplomering. Naast de schoolresultaten van leerlingen zouden dan ook hun toetsresultaten op de gecentraliseerde toetsen op het einde van het schooljaar beschikbaar moeten zijn.

3.2.2. BIJKOMENDE AFNAMES?

INHAALMOMENT

Ongeacht het afnamevenster (cf. infra) moet ook beslist worden of er een inhaalmoment wordt voorzien voor de afname van de gecentraliseerde toetsen. Zo'n inhaalmoment moet dan niet lang na het oorspronkelijke afnamemoment worden georganiseerd om de doorlooptijd niet in het gedrang te brengen. Het op voorhand vastleggen van een mogelijk collectief inhaalmoment kan zinvol zijn als reserveafnamemoment bij onvoorziene technische problemen. Het inschatten van de kans daarop zal de beslissing mee bepalen. Mogelijk hangt die kans samen met de breedte van het afnamevenster. Indien het afnamevenster van een gecentraliseerde toets beperkt is (bijvoorbeeld 1 dag), dan wordt immers de kans groter dat er bijvoorbeeld server-problemen kunnen ontstaan (cf. Perceel 3). Ook bij de toetsafname in de scholen zelf kunnen er onverwacht problemen opduiken (bijvoorbeeld problemen met het draadloos internet), waardoor een collectief reserveafnamemoment voor die school in zo'n situatie een uitweg biedt. Bij de centrale eindtoets kunnen Nederlandse basisscholen zelf zo'n inhaalmoment inplannen in de periode van het afnamevenster. Scholen die kozen voor een papieren afname moeten bij het inhaalmoment wel overschakelen op de digitale afname.

Daarnaast kan eventueel ook een individueel inhaalmoment voorzien worden voor leerlingen die op het oorspronkelijke afnamemoment afwezig waren¹¹. Tabel 3.2 leert dat dit in het buitenland doorgaans georganiseerd wordt zodra individuele toetsresultaten van belang zijn voor bijvoorbeeld studieoriëntering of het behalen van een eindexamen. In de Franse Gemeenschap is dat echter niet het geval: als de school oordeelt dat de leerling gewettigd afwezig was kan de leerling alsnog het betreffende 'certificat' behalen op basis van de schoolresultaten.

De resultaten op de gecentraliseerde toetsen in Vlaanderen zullen niet alleen op leerling-, maar ook op schoolniveau gerapporteerd worden. Als een leerling afwezig is op de dag van de toetsafname, kan dat ook andere redenen hebben dat wat in statistische termen als 'missing at random' wordt geduid. Het kan ook gaan om niet-verwaarloosbare ontbrekende gegevens, als de afwezigheid verband houdt met bepaalde leerlingkenmerken of als er duidelijke schoolverschillen optreden¹². Het is moeilijk te voorzien hoe dit eventuele probleem van selectiviteit van de leerlingparticipatie binnen een school kan aangepakt of voorkomen worden (cf. budgetimplicaties hieronder).

¹¹ Louter als vergelijkingsbasis: bij Vlaamse onderwijspeilingen in het basisonderwijs in 2018 en 2019 waren respectievelijk 1.48% en 3.52% afwezig op de dag van afname. In 2018 was er bij de peiling wiskunde in de eerste graad A-stroom 10.32% afwezige leerlingen. Dit hoge percentage had te maken met het feit dat er die dag een staking was van De Lijn. Ook zulke oncontroleerbare elementen kunnen dus een rol spelen. In 2019 was 5.30% van de leerlingen in het beroepsvoorbereidend leerjaar afwezig bij de peiling wiskunde.

¹² Dit probleem doet zich ook voor m.b.t. de mogelijkheid van 'opt out', zoals in het vorige hoofdstuk werd aangegeven.

HERKANSINGSMOMENT

Indien het resultaat van een leerling op de gecentraliseerde toetsen een invloed kan hebben op schoolloopbaanbeslissingen en de toetsen voor de leerlingen als 'high stakes' worden ervaren, dan zouden er stemmen kunnen opgaan dat er voor leerlingen (maar niet voor scholen) ook een herkansingsmoment moet worden voorzien. De toetsresultaten op het herkansingsmoment worden dan evenwel niet meegenomen in de schoolfeedbackrapporten. De eventuele planning van zo'n herkansingsmoment zou dan moeten gebeuren nadat de resultaten bekend zijn (cf. doorlooptijd), maar voor de start van het nieuwe schooljaar, gezien het nieuwe resultaat de beslissing rond de studiekeuze zou kunnen herzien. In Nederland worden bijvoorbeeld herkansingen voor de centrale eindexamens georganiseerd samen met het eerste inhaalmoment en voor de leerlingen die aan dat inhaalmoment deelnamen in begin juli. Leerlingen kunnen zelfs herkansen als ze toch geslaagd waren, maar een hoger toetsresultaat ambiëren om bijvoorbeeld hun kansen op bepaalde studies in het hoger onderwijs te verhogen. Het gaat voor hen dus effectief om een situatie van 'high-stakes testing'.

3.2.3. BUDGETIMPLICATIES

De keuze voor de periode waarin de gecentraliseerde toetsen in het schooljaar worden afgenomen heeft op zich weinig implicaties voor de kostprijs. Wel zijn er bijkomende financiële kosten voor het organiseren van collectieve en individuele inhaalmomenten en het inrichten van een herkansingsmoment.

Het inrichten van een inhaalmoment als collectief reserve-moment voor alle scholen (bv. bij algehele serverproblemen) of voor een beperkt aantal scholen (met plaatselijke problemen) heeft een aantal budgetimplicaties. De server moet een bijkomende dag ter beschikking worden gesteld. Als het gaat om een algehele afgelasting op de voorziene afnamedag, zal de kost daarvan samenhangen met de ingeschatte verantwoordelijkheid daarvoor van de aanbieder van het toetsplatform. Binnen elke school moet de afname opnieuw worden georganiseerd, met bijkomende inzet van personeelsmiddelen of desgevallend (zie volgend hoofdstuk) van (externe) toetsassistenten. Mogelijk gaan er ook bijkomende uren verloren. Het lijkt niet nodig om een bijkomende toets te voorzien voor het inhaalmoment, al zou het strikt genomen kunnen dat een deel van de toetsen tijdens de niet-geslaagde toetsafname werden vrijgegeven bij een bepaald deel van de leerlingen of scholen. In zo'n situatie zou het beter zijn dat er ook een reservetoets voorzien wordt.

De keuze voor het inrichten van een bijkomend individueel afnamemoment (inhaal- of herkansingsmoment) heeft wel duidelijke financiële gevolgen. Er is de kost van een bijkomende organisatie van een afname voor een deel van de leerlingen. Mogelijk moeten er bijkomende data-analyses worden uitgevoerd, zeker voor herkansingsmomenten en mogelijk ook voor inhaalmomenten als die niet vlak na de eigenlijke afname worden ingepland. Daarnaast kan de toets die op een bijkomend afnamemoment wordt aangeboden niet exact gelijk zijn aan de toets op het oorspronkelijke toetsmoment, vanwege herkenbaarheid van de toets voor de leerling die deze opnieuw aflegt en de kans op toetsfraude. Allicht is dat de reden waarom er in sommige buitenlandse toetspraktijken geen

bijkomende afnamemomenten worden voorzien, zoals bijvoorbeeld in de Franse Gemeenschap, of waarom een inhaalmoment samenvalt met een herkansingsmoment, zoals bijvoorbeeld in Finland. Wanneer de toetsen digitaal en adaptief worden afgenomen en leerlingen ook al bij het oorspronkelijke afnamemoment de facto verschillende toetsen afleggen (zoals bij de toetsen in het basisonderwijs in Nederland en Denemarken), kunnen wel inhaalmomenten worden voorzien, binnen (Centrale eindtoets) of buiten (Deense volgsysteem) het afnamevenster. Dit houdt echter in dat er vooraf al sterk geïnvesteerd is in item- en toetsontwikkeling om een adaptieve procedure mogelijk te maken, zoals in Perceel 1 verder wordt verduidelijkt.

De mogelijkheid dat leerlingen ongewettigd afwezig zijn op de dag van de afname van de gecentraliseerde toetsen zou kunnen roepen om een streng protocol in het geval van afwezige leerlingen. Op zich hebben scholen al heel het schooljaar de taak om op te volgen of de afwezigheid van leerlingen gewettigd is. Het is organisatorisch complex om centraal door het steunpunt gecentraliseerde toetsen bijkomend een controle op die controle uit te voeren. Het zou een bijkomende functionaliteit moeten worden in het toetsplatform (bv. opladen van doktersattesten), maar vooral een grote personeelsinzet vragen om na screening van de data voor de 70000 leerlingen van één toetsafname stappen te zetten om discrepanties op te sporen (al kan dit mogelijks deels geautomatiseerd gebeuren indien hiervoor geïnvesteerd wordt in programmatuur) en op te volgen (contact met de scholen, schoolbezoek ter plaatse, ...). Los van de bijkomende kostprijs van deze extra controles raakt het invoeren ervan ook weer aan het spanningsveld tussen autonomie en vertrouwen geven aan scholen versus hen controleren, wat dan weer kan leiden tot 'collateral damage' als scholen of leerkrachten zich juist gaan verzetten tegen de gecentraliseerde toetsen bij de perceptie van die strenge controles. Indien er een inhaalmoment wordt voorzien, dan is een ander scenario dat scholen doorgeven welke leerlingen afwezig waren op de oorspronkelijke afname (ongeacht of dit gewettigd was). Deze leerlingen worden dan automatisch uitgenodigd voor de bijkomende toetsafname. Op die manier moet er in eerste instantie geen controle van de controle op afwezigheid worden doorgevoerd, al stelt deze vraag zich opnieuw als die leerlingen dan ook op het inhaalmoment afwezig zijn.

3.3. AFNAMEVENSTER

3.3.1. SCENARIO'S

Het afnamevenster is de tijdsperiode waarin een toets moet worden afgenomen. In de buitenlandse praktijkvoorbeelden (zie Tabel 3.2) wordt dit afnamevenster beperkt tot een vaste dag per toetsonderdeel als het gaat om toetsen die belangrijke implicaties hebben voor de leerlingen zelf, maar ook voor de voor leerlingen 'low-stakes' Australische NAPLAN-toetsen worden afnamemomenten voor elke deeltoets door de overheid vastgelegd. Bij de nationale digitaal-adaptieve toetsafnames in het basisonderwijs in Nederland en Denemarken worden wel meerdere afnamedagen voorzien.

Voor de afname van de gecentraliseerde toetsen kunnen drie scenario's inzake de breedte van het afnamevenster overwogen worden: (a) een dag; (b) een week; of (c) twee weken. Bij de laatste twee

mogelijkheden kan de school de afname dan zelf inplannen. Er wordt geen breder afnamevenster voorgesteld omdat er dan al beduidende verschillen in afnamemoment kunnen optreden tussen scholen, waardoor de leerlingen niet overal evenveel leertijd hebben genoten.

Bij de keuze tussen deze scenario's spelen alvast de onderstaande elementen een rol.

- *Planning.* Een vast afnamemoment op dezelfde dag voor alle scholen schept duidelijkheid en gelijkheid. Een breder afnamevenster geeft de scholen flexibiliteit om de afname van de gecentraliseerde toetsen in de eigen schoolkalender in te passen, rekening houdend met bijvoorbeeld klasuitstappen of schoolfeesten.
- *Infrastructuur.* Een afname op één dag binnen heel Vlaanderen stelt hogere infrastructuureisen. Zoals in Perceel 3 wordt aangegeven zijn de vereisten op het vlak van het toetsplatform en de bijhorende servercapaciteit hoger bij een afname op één moment. Een breder afnamevenster verkleint de risico's op grote problemen. Ook voor de scholen stelt een afname op één moment hogere eisen aan de eigen ICT-infrastructuur. Het betekent immers dat alle leerlingen van het betreffende leerjaar op hetzelfde moment een eigen computer, laptop of tablet ter beschikking moeten hebben met een performante wifi-verbinding (zie Perceel 3 voor meer details hierover). De laatste MICTIVO-studie (2019) geeft redenen om daar niet helemaal gerust in te zijn. Bij een breder afnamevenster kan een school een afnamerooster opstellen waarbij leerlingen binnen klassen en/of klassen binnen scholen beurtelings van de computerinfrastructuur gebruik kunnen maken. Dit is de motivatie waarom de centrale eindtoets in het Nederlandse basisonderwijs van een breed afnamevenster gebruik maakt (namelijk van 15 april tot 15 mei).
- *Toetsvarianten.* Als de afname van de gecentraliseerde toets op ééndag binnen heel Vlaanderen plaatsvindt, dan kan elke leerling dezelfde toets afleggen. Dat is niet zo bij een breder afnamevenster. Net zoals bij de bespreking van de individuele inhaalmomenten al werd verduidelijkt, moeten er in die situatie verschillende, evenwaardige toetsversies worden voorzien. Dit geeft een extra kost voor de toetsontwikkeling, zowel als het zou gaan om standaardtoetsen, als wanneer men het probleem omzeilt door te werken met adaptieve toetsafnames waarvoor een grote itembank en een specifiek afnameplatform nodig is (cf. Perceel 1).
- *Inplanning eventueel inhaalmoment.* Een breed afnamevenster is meer compatibel met het inrichten van individuele inhaalmomenten. De school kan dan zelf het bijkomende inhaalmoment inplannen binnen de vooropgestelde periode van dataverzameling. Het plannen van een collectief reserve-moment staat op zich los van de breedte van het afnamevenster. Bij de afname van de gecentraliseerde toetsen op één dag kan ook vooraf één collectieve reservedag worden ingepland.
- *Doorlooptijd.* De keuze van het afnamevenster wordt ook beïnvloed door de noodzakelijke doorlooptijd. Op basis daarvan kan men bepalen wat het laatst mogelijke afnamemoment is dat

kan worden toegestaan om de doorlooptijd niet in het gedrang te brengen. Een breder afnamevenster kan dus alleen vervroegde afnames toelaten ten opzichte van dit ultieme afnamemoment. Wanneer de doorlooptijd relatief lang is, dan zou het kunnen betekenen dat een breed afnamevenster impliceert dat sommige scholen al relatief vroeg in de eindperiode van het schooljaar de gecentraliseerde toetsen zouden afnemen (zie ook 'planning afnames').

- *Eventuele inzet externe toetsassistenten.* Indien de overheid zou beslissen dat de afname van gecentraliseerde toetsen moet gebeuren met externe toetsassistenten in elke school, dan geldt dat hoe breder afnamevenster het afnamevenster is, hoe minder externe toetsassistenten men moet recruteren op voorwaarde dat die gedurende heel het afnamevenster inzetbaar zijn in hun regio. Bij een afname op één dag moet er in elke school een toetsassistent zijn. Dit zijn 2294 toetsassistenten voor het basisonderwijs en 942 toetsassistenten in het secundair onderwijs (cf. het volgende hoofdstuk). Voor de organisatie van externe toetsassistenten zou een afnamevenster van twee weken een verschil kunnen maken t.o.v. een afnamevenster van één week.
- *Onderwijsniveau.* De keuze tussen de scenario's hoeft niet per se gelijk te lopen voor het basis- en het secundair onderwijs.

3.3.2. BUDGETIMPLICATIES

De keuze voor een afnamevenster lijkt budgettair om een afweging te vragen van verschillende kosten. Een gespreide afname leidt tot een daling van de kosten voor het inzetten van een centrale toetsserver en allicht ook tot een spreiding van het risico op technische problemen (zie Perceel 3). Daarnaast geeft een gespreide afname iets meer ruimte voor de complexe organisatie van het inschakelen van een externe toetsassistent in elke school. Daar staat echter tegenover dat een breder afnamevenster duidelijk hogere kosten met zich meebrengt op het vlak van toetsontwikkeling. Ofwel moet men verschillende toetsversies voorzien, ofwel moet men vooraf investeren in een grote itembank, zodat (semi-)adaptieve afnameprocedures mogelijk worden (zie Perceel 1).

3.4. DOORLOOPTIJD

De doorlooptijd is de tijd tussen het afleggen van de toets en het opleveren van de resultaten. In Tabel 3.2 staat voor verschillende praktijkvoorbeelden aangegeven wat de periode is tussen het afleggen van de toets en het opleveren van individuele leerlingresultaten. Deze doorlooptijd varieert van één dag (geautomatiseerd toetsstelsel in Denemarken) over één week (CEB, CED1 en CESS in de Franse Gemeenschap) of enkele weken (centrale eindtoets en eindexamens in Nederland) tot zelfs drie maanden (NCEA in Nieuw-Zeeland). Deze laatste, lange doorlooptijd werd door de OESO (2013) als duidelijk werkpunt naar voor geschoven, maar hangt allicht ook samen met het gebruik van verschillende soorten toetsen bij deze leerlingevaluaties. Groepswerk, praktijkgerichte opdrachten en projectwerk vragen om een langere beoordelingsperiode dan vragen met korteantwoordvorm of dan meerkeuzevragen. Deze laatste soorten vragen zijn bij digitale toetsafnames automatisch scorebaar.

De doorlooptijd van de Vlaamse gecentraliseerde toetsen hangt dus alvast af van het soort van toetsvragen. De complexiteit van de analyses en de mate waarin die geautomatiseerd kunnen worden vormen een andere belangrijke factor. Naast de factoren die de doorlooptijd beïnvloeden, zijn er ook factoren die de wenselijkheid van een beperkte doorlooptijd vooropstellen. Beide soorten factoren worden hieronder bij de voorstelling van de scenario's in meer detail besproken.

3.4.1. SCENARIO'S

De doorlooptijd van de gecentraliseerde toetsen in Vlaanderen kan bepaald worden op basis van de duur van de doorlooptijd en op basis van de deadline voor het opleveren van de feedback. Concreet kunnen de volgende scenario's onderscheiden worden over wanneer de resultaten worden opgeleverd: (a) op de dag van afname of de dag erna; (b) binnen de week; (c) voor het einde van het schooljaar; (d) tegen het begin van het volgende schooljaar of (e) binnen een half jaar tot een jaar na de toetsafname. De scenario's (a), (b) en (c) kunnen elkaar overlappen, naargelang van de afnameperiode en het afnamescenario. De wenselijkheid van elk tijdsscenario is sterk gebonden aan de aard van de gegeven feedback, namelijk of deze zich situeert op leerling-, school- of systeemniveau. Dit zal in onderstaande bespreking duidelijk naar voren komen.

- *Verbeterijd.* De omvang van de doelgroep en ook de vereiste doorlooptijd op leerlingniveau (die scenario's a, b en c veronderstellen) nopen de toetsontwikkelaars tot het kiezen van automatisch scorebare antwoorden. Mogelijk zien sommigen daarbij een spanningsveld tussen deze vereiste en de competentiegerichtheid van de nieuwe ontwikkelingsdoelen (AHOVOKS, 2019; Willem e.a., 2019). Daarnaast vereist het eventuele toetsen van productieve taalvaardigheden bij Nederlands minstens deels om open vraagvormen, zoals bijvoorbeeld schrijfp opdrachten. Daarvoor zijn momenteel in onderzoek al automatisch scoringssystemen ontworpen, maar die zijn o.a. vanwege hun huidige kostprijs nog niet eenvoudig praktisch te implementeren (zie Perceel 1). In het hoofdstuk over 'uitvoering en afname' wordt op het scoren van open vragen nog teruggekomen.
- *Onmiddellijke feedback op leerlingniveau?* Bij digitale toetsafnames is het in principe mogelijk om leerlingen al tijdens de toetsafname feedback te geven over hun antwoord (Scheerens e.a., 2019). Daarbij krijgen ze bij een foutief antwoord een hint en wordt het aantal en de aard van de hints meegenomen bij de scoring van het antwoord¹³. Daarnaast kunnen leerlingen op het einde van de toets bijvoorbeeld ook een overzicht krijgen van het aantal juiste antwoorden of van de vragen die ze goed en fout beantwoordden. Hoewel deze onmiddellijke feedback aan leerlingen het effect van de toets als leermoment kan versterken (zie bijvoorbeeld Dirx, Kester, & Kirschner, 2014) en ook de toetsmotivatie van leerlingen en scholen zou kunnen verhogen,

¹³ Over de wenselijkheid, mogelijke nadelen (o.a. verhoging afnametijd en toetsmotivatie) en haalbaarheid van dit afnamesysteem gaan we hier niet verder in gezien dit systeem van feedback niet meteen de piste is die men met de Vlaamse gecentraliseerde toetsen voor ogen heeft.

zijn er ook duidelijke nadelen aan verbonden. Enerzijds is zo'n ruw resultaat op zich moeilijk te interpreteren. Anderzijds plaatst dit systeem een bijkomende hypotheek op de geheimhouding van de items. Dat laatste is belangrijk om het functioneren van items en van de toets in zijn geheel niet in het gedrang te brengen, zowel binnen een schooljaar (bij verschillende afnamemomenten binnen of tussen scholen) als over schooljaren heen (waarbij ankeropgaven nodig zijn om de vergelijkbaarheid van de gehanteerde meetschaal te garanderen, zie ook Perceel 1).

- *Individuele leerlingresultaten.* Als de toetsopgaven automatisch scorebaar zijn en er op voorhand al een meetschaal is opgesteld waarop een cesuur is geplaatst, dan wordt het mogelijk haalbaar om bijvoorbeeld binnen de week aan scholen de scores van individuele leerlingen en of ze al dan niet specifieke onderwijsdoelen behalen terug te koppelen samen met een beschrijvend overzicht van deze resultaten op klas- en schoolniveau. Alleszins is het noodzakelijk dat deze individuele leerlingresultaten aan de school ter beschikking worden gesteld voor het einde van het schooljaar, gezien de overheid voorziet dat de resultaten op de gecentraliseerde toetsen een rol kunnen spelen in de delibererende klassenraden.
- *Schoolfeedback.* De feedback die scholen ontvangen omvat echter meer dan alleen beschrijvende informatie over het behalen van de onderwijsdoelen door hun leerlingen. Scholen dienen ook feedback te krijgen over hun toegevoegde waarde en van zodra mogelijk over hun leerwinst. Dit vraagt om meer complexe analyses (zie bijvoorbeeld Ehren, 2013; Vanhoof e.a., 2015; Verhaeghe e.a., 2010, maar ook Perceel 1). Voor beide vormen van feedback is het ook noodzakelijk om voorafgaand aan de analyses de toetsresultaten te koppelen aan achtergrondgegevens van leerlingen en scholen. Bij het geven van feedback over leerwinst zijn ook de resultaten van leerlingen bij eerdere afnames van de gecentraliseerde toetsen in hun schoolloopbaan nodig. Het is het streefdoel van het nieuwe steunpunt gecentraliseerde toetsen om deze analyses zoveel mogelijk te automatiseren. Dit neemt niet weg dat er allicht tijd nodig is om de gekoppelde data te 'cleanen' en om bijkomende controles op de analyses uit te voeren indien ze al automatisch kunnen verlopen. Dit maakt dat deze schoolfeedback allicht ten vroegste na de zomervakantie bij het begin van het volgende schooljaar aan de school kan bezorgd worden. In het algemeen moet feedback binnen een redelijke termijn gegeven worden omdat anders de waarde ervan afneemt (Gonzalez-Sancho & Vincent-Lancrin, 2016).
- *Feedback op systeemniveau.* Ook de overheid zal de resultaten op de gecentraliseerde toetsen op het niveau van het Vlaamse onderwijs willen opvolgen. Hierbij gaat het over het algemene prestatieniveau op elk meetmoment en hoe scholen en leerlinggroepen hierin mogelijk verschillen. Daarnaast is het ook belangrijk om trends op te volgen in de cross-sectionele vergelijking van leerlingcohortes op eenzelfde afnamemoment en om op termijn ook de longitudinale leerwinst van leerlingen en scholen te monitoren. Deze resultaten op systeemniveau vragen complexe analyses die niet zomaar geautomatiseerd en dus versneld kunnen worden uitgevoerd. Ter vergelijking kan hierbij meegegeven worden dat de Vlaamse

peilingsresultaten negen maanden na de afname worden opgeleverd. Voor grootschalige, internationale (inclusief digitale) toetsafnames is de doorlooptijd bijna twee jaar.

3.4.2. BUDGETIMPLICATIES

Indien ook open vragen worden meegenomen in de gecentraliseerde toetsen, dan beïnvloedt dat niet alleen de doorlooptijd, maar ook de kostprijs gezien de noodzaak om beoordelaars in te schakelen (zie het volgende hoofdstuk). Zoals hierboven werd aangegeven lijkt enkel een gefaseerde oplevering en dus een verschillende doorlooptijd voor de verschillende niveaus van analyses haalbaar binnen de werkzaamheden van het nieuwe steunpunt. Daarbinnen is budget voorzien voor zowel het draaien van de analyses als voor het begeleiden van scholen in het omgaan met de schoolfeedback. Het is immers belangrijk dat bij het opleveren van de resultaten aan scholen de nodige houvast voorzien wordt voor directeurs en andere leden van het schoolteam zodat zij de resultaten juist leren interpreteren en de werkpunten voor hun school eruit kunnen afleiden. Dit werd als een hekel punt van bijna elk evaluatiesysteem in de door de OESO onderzochte landen aangehaald (Nusche et al., 2012; 2014; OECD, 2013; Santiago et al, 2011; Shewbridge et al., 2011).

3.5. OVERZICHT DEELSCENARIO'S EN BELEIDSADVIEZEN

3.5.1. OVERZICHT DEELSCENARIO'S

In dit hoofdstuk werden verschillende aspecten van de tijdsplanning van de afname van de gecentraliseerde toetsen in detail besproken. Achtereenvolgens werd ingezoomd op verschillende deelscenario's rond de planning van de afname in het schooljaar, de breedte van het afnamevenster en de doorlooptijd. Daarbij werd telkens ook verwezen naar buitenlandse praktijkvoorbeelden. De beleidskeuzes die binnen dit deeldomein moeten worden gemaakt staan weergegeven in Tabel 3.3 en worden in de volgende paragraaf vanuit beleidsperspectief besproken.

3.5.2. BELEIDSADVIEZEN

Een aantal van de beleidskeuzes die gemaakt moeten worden i.v.m. de timing van de afname van de gecentraliseerde toetsen zijn in overleg schijnbaar relatief eenvoudig te nemen. De beslissing over wanneer de afnames in het derde trimester van het schooljaar vallen kan best genomen worden in overleg met het onderwijsveld en hoeft niet gelijklopend te zijn voor het basis- en het secundair onderwijs. De afspraken over de doorlooptijd dienen genomen te worden in overleg met het nieuwe steunpunt. Zoals hierboven al werd aangegeven dringt er zich een gefaseerde oplevering op waarbij de doorlooptijd verschillend is voor de verschillende niveaus van resultaten (leerlingen-, school-, of systeemniveau).

Tabel 3.3.

Beleidskeuzes voor de deelscenario's rond tijdsscenario's van de gecentraliseerde toetsen.

Deelscenario	Keuzeopties
<i>Planning afname(s)</i>	
Tijdstip	15 april - 15 mei; 15 mei - 15 juni; eind juni;
Inhaalmoment	neen; ja (wanneer?)
Herkansingsmoment	neen; ja (wanneer?)
<i>Afnamevenster</i>	1 dag; 1 week; 2 weken;
<i>Doorlooptijd</i>	Dag erna; Binnen de week; Einde schooljaar; Begin volgend schooljaar; 6 tot 12 maanden na afname

De beslissing rond het afnamevenster is complexer omdat bij de keuze een afweging gemaakt moet worden tussen kosten met betrekking tot de ict-infrastructuur (cf. Perceel 3) versus met betrekking tot de toetsontwikkeling (cf. Perceel 1). Een breder afnamevenster impliceert een minder sterke piekbelasting van de afnameserver, is daardoor goedkoper, en is mogelijk voor sommige scholen meer haalbaar omdat de leerlingen gespreid de toetsen kunnen afleggen. Daar staat tegenover dat een breder afnamevenster vereist dat er meerdere toetsversies ontwikkeld worden, wat een hogere kostprijs betekent voor toetsontwikkeling en wat ook leidt tot meer complexe analyses. Dit laatste heeft mogelijk implicaties voor de doorlooptijd.

Ten slotte moet de onderwijsminister ook beslissen of er voor afwezige leerlingen een inhaalmoment moet worden voorzien en of leerlingen kunnen herkansen op de toets. Zoals bij heel wat andere beleidskeuzes die nog genomen moeten worden over de gecentraliseerde toetsen hangen beide beslissingen deels samen met de vraag of de gecentraliseerde toetsen op leerling- en schoolniveau 'low-' of 'high-stakes' zijn. Als scholen er belang bij kunnen hebben dat leerlingen niet zouden opdagen op de dag van de toetsafnames, is het misschien nodig om een inhaalmoment te voorzien om zo een beeld te krijgen van alle leerlingen op school. Als de schoolloopbaan van leerlingen deels beïnvloed kan worden door hun uitslagen op de gecentraliseerde toetsen, dan kan het nodig zijn om een herkansingsmoment te voorzien. Beide extra afnamemomenten hebben duidelijk budgettaire implicaties, zowel naar de organisatiekosten (o.a. servercapaciteit, planlast voor scholen) als naar toetsontwikkeling en data-analyse voor het garanderen van equivalente resultaten met andere toetsversies.

4. UITVOERING AFNAME

4.1. INLEIDING

Dit hoofdstuk behandelt eerst onderzoeksvragen 19 en 20 die gaan over de uitvoering van de afname van de gecentraliseerde toetsen. Achtereenvolgens wordt gefocust op de locatie en infrastructuur om de toetsafname te laten plaatsvinden, de betrokken actoren met hun rollen en taken, de organisatie van het eigenlijke toetsmoment en de kwaliteitsmonitoring. We staan ook stil bij de naverwerking van de afgenomen toetsen. Tot slot wordt stilgestaan bij de nood aan toetsverversing (Vraag 17).

4.2. TOETSLOCATIE EN INFRASTRUCTUUR

Om de toetsafname plaats te laten vinden is er nood aan een locatie waar leerlingen in de best mogelijke omstandigheden de gecentraliseerde toetsen kunnen afleggen. Het gaat echter niet enkel om het voorzien in examenlokalen. Vermits het gaat om digitale toetsen is het bovendien noodzakelijk dat de nodige ICT-infrastructuur voorzien wordt. In een eerste scenario worden de gecentraliseerde toetsen op buitenschoolse toetslocaties afgenomen. Daarna volgt een bespreking waarbij de betrokken leerlingen de gecentraliseerde toetsen afleggen in hun eigen school.

4.2.1. BUITENSCHOOLSE TOETSLOCATIES

In een scenario waar de toetsen in buitenschoolse toetslocaties worden afgenomen, worden leerlingen opgeroepen om zich te begeven naar centrale locaties in hun regio om hun gecentraliseerde toetsen af te leggen. Dit is een scenario dat analoog is aan de toelatingsexamens voor arts/tandarts in Vlaanderen (Departement Onderwijs en Vorming, 2020), maar ook de nationale examens ter toelating tot het hoger onderwijs in Japan kent een aanpak waarbij voor buitenschoolse locaties worden gekozen (National Center for University Entrance Examinations, 2017). Op deze buitenschoolse locaties worden speciaal ingerichte examenlokalen voorzien om de toetsen af te leggen. Dit kunnen bijvoorbeeld grote evenementenhallen zijn. Het laten plaatsvinden van de toetsen op dergelijke centrale locaties bevordert de standaardisatie van de afname. Er is in een dergelijk scenario een grote controle over de omstandigheden waarin de leerlingen hun toetsen afleggen.

Indien alle leerlingen op hetzelfde moment hun toetsen dienen af te leggen moet er voor ongeveer 70000 leerlingen per cohorte een plek voorzien worden verspreid over Vlaanderen. Om te voorzien in deze locaties zal niet enkel een reservatie van deze gebouwen moeten gebeuren voor de periode van de toetsafname als dusdanig. Vermits deze locaties allicht niet standaard ingericht zijn om toetsen te laten afnemen, zal ook een periode voor en na de toetsafname moeten voorzien worden waarin de opbouw en afbraak van de werkplekken voor leerlingen moet gebeuren. De huurkosten van deze locaties zal dus betrekking hebben op een bredere periode dan die van de toetsafname.

Naast de kosten die samengaan met het huren van dergelijke buitenschoolse toetslocaties moet ook ingecalculeerd worden dat een dergelijk scenario een mobiliteitskwestie met zich meebrengt. Per cohorte

betekent dit immers dat zo'n 70000 leerlingen op de centrale toetslocaties in hun regio dienen te geraken. Deze verplaatsingen zullen een mobiliteitsstroom met zich brengen. Bovendien kan niet verwacht worden dat leerlingen, zeker de leerlingen uit het lager onderwijs, zich zelfstandig verplaatsen naar een centrale toetslocatie die mogelijk kilometers verwijderd ligt van hun school of woonplaats.

Het inrichten van deze buitenschoolse toetslocaties betekent niet enkel dat er een tafel en stoel wordt voorzien waaraan de leerlingen de gecentraliseerde toetsen kan maken. Vermits de gecentraliseerde toetsen digitaal afgenomen worden, moet er naast het inrichten van de werkplek als dusdanig ook voorzien worden in het installeren van de nodige ICT-infrastructuur. Dit betekent dat er hardware en software moet voorbereid en geïnstalleerd worden zodat de systemen klaar zijn voor gebruik. In het geval dat er beslist wordt om alle leerlingen op eenzelfde moment te toetsen zullen er ongeveer 70000 computers of apparaten beschikbaar moeten zijn om de toetsafname te laten plaatsvinden. Een alternatief daartoe is dat er gewerkt wordt onder het principe 'bring your own device', zoals dat in Finland het geval is voor het eindexamen secundair onderwijs (Ylioppilastutkintolautakunta, 2020). In een dergelijke situatie moet het toetsplatform erop voorzien zijn dat de toets op verschillende types van toestellen moet kunnen lopen. Bovendien wordt er in een dergelijk geval van uit gegaan dat elke leerling een apparaat ter beschikking heeft dat kan dienen voor het afleggen van de toetsen, en dus voldoet aan de technische vereisten.

Het organiseren van de gecentraliseerde toetsen in de context van buitenschoolse locaties betekent ook dat er een inzet van personeel moet gebeuren op die locaties. Het is immers zo dat leerlingen zullen moeten geregistreerd worden en er een identiteitscontrole zal moeten gebeuren alvorens gestart kan worden met het afleggen van de toets. Het is belangrijk dat elke leerling zijn/haar individuele toets maakt om een correcte inschatting van de prestaties te maken en latere leerwinstmeting mogelijk te maken. Hiervoor moeten op elke buitenschoolse toetslocatie voldoende personeelsleden ingezet worden die de toetsafname ondersteunen. Bovendien moeten er personeelsleden aangeworven en opgeleid worden die op vlak van ICT de toetsafname ondersteunen zodat er in het geval van technische problemen ingegrepen kan worden. Tot slot kunnen buitenschoolse locaties ook een nadelige impact hebben op de prestaties van leerlingen, bijvoorbeeld omdat een aantal van hen meer stress heeft om de toets op een niet-vertrouwde locatie af te leggen.

4.2.2. DE EIGEN SCHOOL ALS TOETSLOCATIE

Een ander scenario bestaat er uit dat de betrokken leerlingen de gecentraliseerde toetsen afleggen op hun eigen school. In vergelijking met het scenario met buitenschoolse toetslocaties staan hier geen huurkosten tegenover. In alle buitenlandse voorbeelden (Tabel 3.1 en 3.2) gebeuren grootschalige toetsafnames in het basisonderwijs stevast op school. Voor het secundair onderwijs gaat het in die voorbeelden telkens om 'high-stakes' toetsen op leerlingniveau. De toetsafname vindt zowel plaats op school (bv. CED1, CESS, Centrale eindexamens), in de sporthal van de school (Finland) als in aparte examencentra (NCEA).

Scholen beschikken doorgaans over lokalen die gebruikt kunnen worden om grotere groepen leerlingen toetsen te laten afleggen wanneer ze zouden werken met hun eigen laptop of tablet, indien in die ruimte een voldoende sterk wifi-signaal is (zie Perceel 3). In veel gevallen beschikken scholen bovendien over lokalen die uitgerust zijn met ICT-materiaal waarmee leerlingen de digitale toetsen kunnen afleggen. Hierin is allicht een verschil tussen het basisonderwijs en het secundair onderwijs. De MICTIVO3-studie toonde aan dat er wezenlijke verschillen bestaan tussen deze onderwijsniveaus in termen van het ter beschikking hebben van hardware (Heymans, Godaert, Elen, Van Braak & Goeman, 2018). Voor een verdere analyse van dit aspect verwijzen we naar Perceel 3.

Indien nodig zou het steunpunt eventueel bijkomende ICT-capaciteit kunnen voorzien wanneer dat nodig blijkt. Scholen kunnen voorafgaand aan de toetsafname bevraagd worden over de bestaande situatie op vlak van ICT zodat eventuele noden voor bijkomende ICT-infrastructuur kunnen ingelost worden. Alleszins kan in dit scenario wel maximaal ingezet worden op het gebruik van de reeds aanwezige middelen om de toetsafname te laten plaatsvinden zonder extra en ongebruikelijke mobiliteitsstromen naar centrale toetslocaties.

De registratie van leerlingen is minder complex binnen dit scenario. De leerlingen gaan immers gewoon naar school en kunnen door leraren waarmee ze vertrouwd zijn geïdentificeerd worden. Toch zullen er ook binnen dit scenario toetsassistenten (cf. infra) moeten worden ingezet die de toetsafname praktisch ondersteunen op de school. Hierin verschilt het scenario met de school als toetslocatie niet met die van buitenschoolse toetslocaties. Er dient immers gewaakt te worden over het gestandaardiseerd afnemen van de toetsen.

Een ander aspect dat verbonden is aan het organiseren van de gecentraliseerde toetsen op de eigen school situeert zich op het niveau van de leerling. Binnen dit scenario leggen de leerlingen de gecentraliseerde toetsen af in hun vertrouwde schoolomgeving begeven. Dit kan belangrijk zijn als geruuststellende factor voor leerlingen die stress ervaren bij het afleggen van de gecentraliseerde toetsen, wat ook de betrouwbaarheid en validiteit van de toetsresultaten ten goede komen (Bonnaccio & Reeve, 2010).

4.3. TAKEN EN ROLLEN

Bij een afname van de gecentraliseerde toetsen in de scholen, zijn er verschillende taken te onderscheiden die aan bepaalde rollen kunnen worden toegeschreven. Dit geldt zowel op het niveau van de school als op het centrale niveau (cf. het steunpunt). Daarbij kunnen deze taken toegewezen worden aan drie verschillende fasen binnen het implementatieproces: voorbereiding, toetsafname en toetsverwerking.

4.3.1. SCHOOLNIVEAU

Tabel 4.1 geeft een overzicht van de organisatorische taken die de afname van de gecentraliseerde toetsen met zich meebrengt voor de scholen. Deze planlast wordt opgedeeld in taken en rollen op het niveau van de school in de verschillende fasen van de afname van de gecentraliseerde toetsen. Deze rollen

en taken worden hieronder verder toegelicht. Het is daarbij belangrijk op te merken dat in deze beschrijving wordt uitgegaan van de taken die per rol vervuld moeten worden. Binnen een school kan eenzelfde persoon echter verscheidene van deze rollen opnemen.

TOETSCOÖRDINATOR

Vorbereidingsfase

Een belangrijke rol binnen de implementatie van de gecentraliseerde toetsen is die van toetscoördinator. In de voorbereidingsfase van de toetsafname heeft deze rol verschillende taken en verantwoordelijkheden te vervullen. Als cruciale taak dient de toetscoördinator te voorzien in de nodige infrastructuur zodat het toetsmoment zoals gepland kan plaatsvinden. Dit betekent dat er zowel voorzien moet worden in passende lokalen zodat elke leerling een werkplek (werktafel en stoel) heeft waarin deze de gecentraliseerde toetsen kan afleggen als in een goed uitgewerkte planning waarin elke klas aan een lokaal en eventueel tijdstip (cf. afnamevenster) wordt toegewezen. Vermits de gecentraliseerde toetsen digitaal zullen afgenomen worden, betekent dit ook dat de toetscoördinator de verantwoordelijkheid draagt over het ervoor zorgen dat het nodige ICT-materiaal nodig aanwezig is en dat desgevallend een ict-verantwoordelijke wordt aangesteld.

De toetscoördinator zorgt er in de voorbereidingsfase ook voor dat leerlingen de gelegenheid hebben om (deel)vrijstellingen of toetsaanpassingen aan te vragen (zie het hoofdstuk over brede afname). Afhankelijk van de manier waarop de procedures voor deze aanvragen verder wordt uitgewerkt, is het mogelijk dat de toetscoördinator ervoor zorgt dat de betrokken actoren worden samengebracht om het aanvraagdossier rond (deel)vrijstellingen of toetsaanpassingen te beoordelen.

Verder vormt de toetscoördinator de contactpersoon binnen de school in relatie tot het steunpunt. Communicatie omtrent richtlijnen, het signaleren van moeilijkheden in de organisatie van de gecentraliseerde toetsen of de inzet van begeleiders bij de uiteindelijke toetsafname tussen de school en het steunpunt gebeurt via de toetscoördinator. Een goede, vlotte communicatie is belangrijk, en het aanwijzen van één enkele contactpersoon is daarbij aangewezen.

Naast de communicatie tussen het steunpunt en de school, is de toetscoördinator ook het aanspreekpunt binnen de school voor alle collega's en leerlingen die vragen hebben bij de voorbereiding en de daadwerkelijke afname van de gecentraliseerde toetsen op school. Deze persoon is op schoolniveau de spin in het web met betrekking tot de organisatie van de gecentraliseerde toetsen.

Tijdens de toetsafname

Tijdens de fase van toetsafname is de toetscoördinator verantwoordelijk voor het goede verloop van de eigenlijke toetsafname. De toetscoördinator garandeert ook dat er tijdens het afleggen van de toetsen voldoende begeleiders aanwezig zijn. Bij het ontstaan van eventuele problemen is de toetscoördinator

diegene die werkzaamheden overziet en tracht op te lossen, al dan niet in samenspraak met het steunpunt.

Na de toetsafname

Na de toetsafname zal de toetscoördinator zorgen dat de school- en leerlingenfeedback bij de juiste collega's terecht komt en dat, zeker als de gecentraliseerde toetsen *high-stakes* (cf. infra) zijn voor leerlingen, desgevallend het recht op inzage voor leerlingen en hun ouders in de school georganiseerd wordt.

Tabel 4.1. *Rollen op schoolniveau met verantwoordelijkheden in relatie tot fases van het implementatieproces van de gecentraliseerde toetsen*

Rollen op schoolniveau	Fases implementieproces		
	Vorbereiding	Tijdens toetsafname	Na toetsafname
<i>Toetscoördinator</i>	<ul style="list-style-type: none"> - Voorzien in nodige infrastructuur - Registreren van aanvragen voor (deel)vrijstellingen en toetsaanpassingen - Contactpersoon zijn voor steunpunt op school - Aanspreekpunt zijn voor collega's op school m.b.t. organisatie gecentraliseerde toetsen 	<ul style="list-style-type: none"> - Waken over het goed verloop van de toetsafname - Garanderen inzet van toetsassistenten en begeleidende leerkrachten 	<ul style="list-style-type: none"> - Terugkoppeling school- en leerlingenfeedback aan betrokken actoren. - Eventueel: organisatie van het inzagerecht.
<i>Toetsassistenten en toetsbegeleiders</i>		<ul style="list-style-type: none"> - Begeleiden gestandaardiseerde toetsafname - Controle identiteit leerlingen 	
<i>ICT-bijstand</i>	<ul style="list-style-type: none"> - Voorzien in nodige ICT-infrastructuur - Aanspreekpunt voor technische bijstand door steunpunt 	<ul style="list-style-type: none"> - Bijstand bij eventuele ICT-problemen - Aanspreekpunt voor technische bijstand door steunpunt 	
<i>Toetsinformatie-makelaar</i>	<ul style="list-style-type: none"> - Schoolteam informeren over de gecentraliseerde toetsen 		<ul style="list-style-type: none"> - Schoolteam mee engageren bij het aan de slag gaan met de schoolfeedback

TOETSASSISTENTEN EN TOETSBEGELEIDERS

Een essentieel aspect van de gecentraliseerde toetsen is de standaardisatie van de afname (zie ook het hoofdstuk over brede afname). Een toetsassistent is diegene die de afname van de toetsen leidt en superviseert op de toetslocatie. Daarbij zijn er ook toetsbegeleiders die de toetsafnames mee begeleiden. Dit kunnen leerkrachten zijn (cf. infra), maar ook externen. Het is immers vanzelfsprekend dat de toetsassistenten niet bij elke klasgroep of in elke pc-klas tegelijk aanwezig kan zijn. De toetsassistent ziet erop toe dat het afleggen van de toetsen op het correcte tijdstip start, dat leerlingen onder de juiste omstandigheden de toetsen afleggen, al dan niet gebruik makend van de juiste (lokale) toetsaanpassingen. Deze persoon zorgt er ook voor dat de toetsen volgens de geijkte procedure worden afgerond. Mogelijk moet de toetsassistent ook een logboek bijhouden, waarin het verloop van de afname wordt beschreven. Dit logboek kan nadien gebruikt worden om de correctheid van de afnameprocedure te checken, maar zou ook de afwezigheid van leerlingen kunnen rapporteren. Bij high-stakes toetsen (cf. infra) wordt erin ook genoteerd welke personen allemaal aanwezig waren. Tot slot behoort het tot de taak van de toetsassistent om te garanderen dat de juiste leerling zijn/haar individuele toetsen aflegt.

Verderop in dit rapport zetten we verschillende scenario's over de recrutering van deze toetsassistenten naast elkaar. Daarbij gaan we ervan uit dat de toetsbegeleiders door de school zelf worden voorzien, zoals dit nu bijvoorbeeld ook het geval is bij de Vlaamse onderwijspeilingen of bij de toetsafnames in het kader van grootschalig internationaal onderzoek.

ICT-ONDERSTEUNING

In aanloop naar de toetsafname dient er ICT-ondersteuning voorzien te worden om de toetslocaties in te richten met het nodige materiaal. Dit betekent dat er zowel aan eisen op vlak van hardware als software voldaan moet worden om de toetsen te kunnen afleggen. ICT-ondersteuning voorziet in de technische competenties om de toetscoördinator bij te staan in het organiseren van de toetslokalen en werkplekken. Bovendien neemt deze rol ook de communicatie op over technische aspecten van hardware en software met het steunpunt.

Ook tijdens de toetsafname is technische bijstand voorzien cruciaal om de gestandaardiseerde toetsafname te ondersteunen. In geval van technische problemen tijdens de toetsafname moet er kunnen ingegrepen worden. Problemen kunnen zich situeren op vlak van netwerkverbindingen, hardware of software die geïnstalleerd staat op het toestel dat gebruikt wordt om de toets af te leggen of softwarematige problemen die zich zouden voordoen binnen het digitale toetsplatform. De persoon die deze rol opneemt kan ook in contact treden met het steunpunt bij eventuele problemen tijdens de toetsafname.

TOETSINFORMATIEMAKELAAR

Op basis van de toetsgegevens zullen er analyses dienen te gebeuren om resultaten van de toetsen terug te koppelen (zie ook Perceel 1). Deze terugkoppeling kan plaatsvinden op verschillende niveaus: aan de overheid, aan de school van de betrokken leerlingen, aan de leraar van de betrokken leerlingen, aan de leerling zelf, aan zijn/haar ouders en het brede publiek. Hierbij wordt de betrokken actor naar wie de resultaten worden teruggekoppeld een rol. Om de gegevens aan te wenden op schoolniveau lijkt het wenselijk om ook een rol te voorzien voor toetsinformatiemakelaars. Deze actor kan het informatieaanbod, zoals de informatie in schoolfeedbackrapporten, aan informatiebehoeften koppelen. Dit kan schoolintern (bv. de schoolleider zelf of een beleidsondersteuner), dan wel schoolextern (bv. pedagogisch begeleiders). De implementatie van samenwerk-vormen kan een bevorderend effect hebben op informatiemobiliteit, kennisontwikkeling, en uiteindelijk

zinvol informatiegebruik (Schildkamp et al., 2019; Van Gasse et al., 2017, 2019). Mogelijk kan deze rol van toetsinformatiemakelaar worden toegewezen aan de zorgcoördinator.

4.3.2. CENTRALE NIVEAU

Naast de rollen die de school opneemt, kan ook verwezen worden naar het centrale niveau dat rollen en taken dient op te nemen bij de uitvoering van de afname. Het gaat hier enerzijds over het nieuwe steunpunt gecentraliseerde toetsen en anderzijds over de aanbieder van het toetsplatform. Hoe deze twee partners bepaalde verantwoordelijkheden onder elkaar verdelen dient zorgvuldig afgestemd te worden.

Tabel 4.2 geeft een overzicht van de rollen en taken op het centrale niveau tijdens het implementatieproces. *Bijstand* omvat het informeren en aansturen van alle andere betrokken met het oog op het realiseren van de gewenste toetscontext voor leerlingen. Zowel in de voorbereiding als tijdens de toetsafname zal er daartoe een helpdeskfunctie dienen gecreëerd te worden. Hetzelfde geldt voor de *ICT-ondersteuning*. Men kan er voor kiezen deze rollen in één helpdesk te integreren. Met het oog op het verzamelen van betrouwbare en valide resultaten over het prestatieniveau van leerlingen is een kwaliteitsvolle afname van de gecentraliseerde toetsen van belang. Het steunpunt zal bij het ontwikkelen van de toets ook procedures moeten uitwerken die een gestandaardiseerde toetsafname beschrijven. Het steunpunt kan een *kwaliteitsmonitor* aanstellen die als taak heeft na te gaan of deze voorgeschreven procedures gevolgd worden bij de toetsafname. Het gaat hier niet om een louter controlerende taak, maar er kunnen ook suggesties uit voortvloeien die de procedures verder kunnen verfijnen of verbeteren. Hierop wordt in de volgende paragraaf verder ingegaan.

Tabel 4.2. Rollen op centraal niveau met verantwoordelijkheden in relatie tot fases van het implementatieproces van de gecentraliseerde toetsen

Rollen op centraal niveau	Fases implementieproces		
	Vorbereiding	Tijdens toetsafname	Na toetsafname
<i>Bijstand</i>	- Helpdeskfunctie toetscoördinatoren	- Helpdeskfunctie toetscoördinatoren	
<i>ICT-bijstand</i>	- Helpdeskfunctie ICT-ondersteuning	- Helpdeskfunctie ICT-ondersteuning	
<i>Kwaliteitsmonitor</i>		- Monitoring kwaliteitsvolle toetsafname	- Rapportering

4.4. VERDIEPING VAN SCENARIO'S M.B.T. TOETSASSISTENTEN EN KWALITEITSMONITORING

Standaardisatie is een voorwaarde voor een betrouwbare en valide meting van leerlingresultaten bij de gecentraliseerde toetsen. Het is aan het steunpunt om de procedures of protocollen uit te werken die op duidelijke wijze de afnameprocedure beschrijven. Daarbij zijn alleszins toetsassistenten nodig om de toetsafnames te begeleiden. Daarnaast kan overwogen worden om een expliciete kwaliteitsmonitoring van de afname te voorzien.

4.4.1. MOGELIJKE SCENARIO'S M.B.T. RECRUTERING VAN TOETSASSISTENTEN

De taak van een toetsassistent bestaat eruit om de eigenlijke toetsafname te begeleiden in samenwerking met toetsbegeleiders of leerkrachten van de school gezien het niet haalbaar is dat de toetsassistent alleen de afname van al de leerlingen van een bepaald leerjaar in een school begeleidt. De toetsassistent superviseert en ondersteunt de toetsbegeleiders bij de afname. Deze taak werd hierboven al in detail beschreven. Er zijn verschillende scenario's mogelijk met betrekking tot het aanstellen van deze toetsassistenten.

SCENARIO 1: EXTERN AANGEWORVEN TOETSASSISTENTEN

Een eerste scenario rond het begeleiden van de toetsafname is die van extern aangeworven toetsassistenten. De idee rond extern aangeworven toetsassistenten is analoog met de werkwijze binnen het huidige Vlaamse Steunpunt Paralleltoetsen en Peilingsonderzoek (STEP). Bij de peilingen worden er personen als vrijwillig medewerkers aangeworven (veelal gepensioneerde leraren) die tegen een forfaitaire vergoeding en onder een verzekering van het 'woon-werkverkeer' de peilingstoetsen afnemen bij de leerlingen. Deze werkwijze is ook analoog met eTIMSS (de digitale toetsing van TIMSS). In Nederland worden daarbij voornamelijk jobstudenten ingezet om de toetsafname te begeleiden op scholen. In Frankrijk wordt daarbij een beroep gedaan op oud-onderwijsinspecteurs (binnen een regio gelijk aan een

Departement). Een belangrijke kanttekening hierbij is dat zowel bij STEP als bij eTIMSS gewerkt met een steekproef van ongeveer 4000 à 5000 leerlingen, wat een beduidend verschil is met een populatiebrede toetsafname zoals die beoogd wordt met de gecentraliseerde toetsen bij ongeveer 70000 leerlingen.

Ook voor de implementatie van de gecentraliseerde toetsen in Vlaanderen kan een dergelijke werkwijze uitgedacht worden. Vanuit het steunpunt gecentraliseerde toetsen wordt personeel aangeworven dat uitgestuurd wordt naar de scholen waar de betrokken leerlingen de gecentraliseerde toetsen dienen af te leggen. De functie van deze extern aangeworven toetsassistenten is dan tweeledig. De onafhankelijkheid van de extern aangeworven toetsassistent kan bijdragen aan de controle over de standaardisatie waarmee de gecentraliseerde toetsen worden afgelegd door de leerlingen (cf. infra). Verder kan een extern aangeworven toetsassistent ook een blijk geven van ondersteuning voor de betrokken scholen.

Dit scenario impliceert dat extern personeel dient aangeworven te worden. Dit betekent dat er een profiel moet opgesteld worden dat als basis kan dienen om te werving van personeel te starten. Het gaat dan om een aantal generieke competenties waarover deze kandidaten dienen te beschikken, zoals administratieve competenties, maar ook probleemoplossend vermogen voor het geval zich onverwachte situaties voordoen. Uiteraard dienen zijn vooraf tot in detail gebrieft te worden over achtergronden en procedures. Daarnaast dienen kandidaten ook te beschikken over specifieke competenties zoals het passend benaderen en ondersteunen van de doelgroep leerlingen (lager en secundair onderwijs) in de toetsafname en het kunnen werken met ict-infrastructuur in het algemeen en met het toetsplatform in het bijzonder.

Indien de gecentraliseerde toetsen op één moment worden afgenomen dan dient er per afnamemoment in het basis- en secundair onderwijs respectievelijk ongeveer 2500 en 1000 toetsassistenten te worden voorzien, los van een aantal reserve-toetsassistenten per regio. Deze toetsassistenten moeten voorafgaand aan de toetsafnames een kortdurende opleiding volgen. Noch het inzetten van bepaalde personeelsgroepen van het departement onderwijs en vorming (bv. leden van de inspectie), noch het oproepen via bepaalde Vlaamse diensten (bv. VDAB), noch het zoeken van vrijwilligers, noch het recruterend via interimbureaus lijken realistische manieren te bieden om een succesvolle werving te garanderen. Na of tijdens de wervingsfase moet daarnaast ook een systeem worden voorzien om de toetsassistenten toe te wijzen aan de scholen en hun eventuele vestigingsplaatsen, waarbij vanzelfsprekend een koppeling gemaakt moet worden tussen hun woonplaats en de locatie van de scholen. Tot op de dag van afname moet het ook mogelijk worden gemaakt dat reserve-toetsassistenten kunnen toegewezen worden bij ziekte of andere problemen met de oorspronkelijk toegewezen toetsassistent. Indien er per toets een breder afnamemoment wordt voorzien (cf. het vorige hoofdstuk) dan zijn er misschien minder toetsassistenten nodig, maar dan wordt de complexiteit van het toewijzen van toetsassistenten aan afnamemomenten binnen scholen uit hun regio er niet eenvoudiger op. Gegeven de complexiteit van werving en toewijzing wordt voor dit scenario geen verdere budgetberekeningen gemaakt.

SCENARIO 2: LERAREN, EXTERN AAN DE SCHOOL

Een alternatief voor het werven van externe toetsassistenten, is het inzetten van leraren als externe toetsassistent. Zo zou bijvoorbeeld gewerkt kunnen worden met een doorschuifstelsel waar leraren uit een andere school of scholengemeenschap als externe toetsassistenten de gecentraliseerde toetsen afnemen. Dit vergt de medewerking van de scholen en leerkrachten bij het werven van deze externe toetsassistenten, die aangemeld zouden kunnen worden binnen een apart portaal van het digitale toetsenplatform. De complexiteit van de toewijzingsprocedure wordt niet verlicht, gezien de zo aangewezen leerkrachten in functie van hun beschikbaarheid nog moeten worden toegewezen aan afnamemomenten binnen naburige scholen. Mogelijk moet ook een verplaatsingsvergoeding worden betaald.

Het voordeel van deze procedure zou kunnen zijn dat schoolexterne leerkrachten minder belang hebben bij de toetsresultaten van de betreffende leerlingen en scholen (cf. infra) en zo op een meer betrouwbare manier de standaardisering van de toetsafnames kunnen bewaken en mee waarborgen. Het grote nadeel van dit scenario is dat de overheid daarmee impliciet het signaal geeft dat de eigen leerkrachten van een school niet vertrouwd kunnen worden voor het afnemen van de gecentraliseerde toetsen.

SCENARIO 3: LERAREN, INTERN AAN DE SCHOOL

Ten slotte bestaat een derde scenario erin dat de school zelf een toetsassistent aanduidt die samen met andere leerkrachten van de school de toetsafnames begeleidt. Zowel de werving als de toewijzing aan het toetsmoment wordt dan opgenomen worden door elke school apart. Daarbij zou men bijkomend als instructie kunnen geven dat men hiervoor leerkrachten aanduidt die niet de klasleerkrachten zijn (basisonderwijs) of die geen vakleerkrachten zijn in het getoetste domein (secundair onderwijs). De aangewezen toetsassistent moet bereid zijn om vooraf een opleiding te volgen. De taken van deze interne toetsassistent blijven gelijk aan aan wat eerder beschreven werd.

4.4.2. MOGELIJKE SCENARIO'S M.B.T. KWALITEITSMONITORING

Bij grootschalige internationale toetsafnames wordt ook kwaliteitsmonitoring van de toetsafnames voorzien. Dit kan ook overwogen worden bij de gecentraliseerde toetsen in Vlaanderen. Hieronder worden daartoe twee mogelijke scenario's, namelijk interne of steekproefsgewijze externe kwaliteitsmonitoring, besproken. Ook de combinatie van beide scenario's is mogelijk.

SCENARIO 1: INTERNE KWALITEITSMONITORING

Vermits het steunpunt een uitvoerige beschrijving zal maken van de te volgen procedures bij de toetsafname kan beslist worden dat er vertrouwen gelegd wordt in het Vlaamse lerarenkorps en aangeduide toetsassistenten om deze richtlijnen op een kwaliteitsvolle manier in de praktijk om te zetten. Er wordt in dit scenario een appèl gedaan op de beroepsernst van het bestaande lerarenkorps in

Vlaanderen. Dit sluit aan bij de huidige manier van evalueren binnen het Vlaamse onderwijs waar ook veel vertrouwen en verantwoordelijkheid gelegd wordt bij leraren en schoolteams. Zij zijn vertrouwd met het afnemen van toetsen bij leerlingen en hebben dus ook competenties ontwikkeld op vlak van leerlingenevaluatie.

Binnen dit scenario kan gevraagd worden aan de school om een kort rapport op te stellen waarbij uitgegaan wordt van een set vastgelegde topics met betrekking tot de wijze waarop de toetsafname is verlopen. Dit kan bijvoorbeeld gebeuren door de toetscoördinator, gezien deze persoon de volledige toetsafname op de school overschouwt. Hierbij kunnen eventuele problemen die zich hebben voorgedaan gemeld worden, of kan men bijvoorbeeld aangeven welke onduidelijkheden er zijn met betrekking tot de organisatie of procedures in verband met de toetsafname. Deze interne kwaliteitsmonitoring gebeurt dus naast de rapportage door de toetsassistent over de afname zelf. Deze schoolrapporten bieden het steunpunt de gelegenheid om de voorgeschreven procedures waar nodig bij te sturen of verder te optimaliseren.

SCENARIO 2: STEEKPROEFSGEWIJZE EXTERNE KWALITEITSMONITORING

Het steunpunt kan er ook voor opteren om een externe kwaliteitsmonitor aan te stellen, zoals bij het internationaal vergelijkend onderwijsonderzoek als bijvoorbeeld TIMSS, PIRLS, en TALIS. De kanttekening hierbij is dat het in deze gevallen gaat om studies die werken op basis van een steekproef van enkele duizenden leerlingen en niet op basis van populatiebrede toetsafnames.

Een dergelijke kwaliteitsmonitoring zou kunnen gebeuren bij een steekproef van deelnemende scholen aan de gecentraliseerde toetsen. Daarbij brengt de kwaliteitsmonitor of 'toetsinspecteur' een onaangekondigd bezoek aan de school tijdens de afname waarbij zowel de werkzaamheden van de toetsassistent als de begeleidende leerkrachten worden gecontroleerd. Van een dergelijk bezoek wordt een rapport opgesteld en bezorgd aan het steunpunt.

De organisatie van deze kwaliteitsmonitoring vergt de nodige voorbereiding. Allereerst dient er een werving te gebeuren van personen die tegen een vergoeding deze taak kunnen volbrengen. Hierbij wordt gezocht naar personen die de nodige competenties om scholen onafhankelijk en op gepaste wijze te benaderen. Verder dienen deze kandidaten te beschikken over administratieve competenties om een verslag te kunnen opstellen over hun werkzaamheden. De kandidaten moeten vervolgens een opleiding ontvangen over de procedures die de toetsafname moeten begeleiden en zij moeten geïnstrueerd worden over wat er van hen verwacht wordt met betrekking tot het verslag dat verwacht wordt. Tenzij er een beroep kan worden gedaan op leden van het departement onderwijs en vorming (bijvoorbeeld leden van de inspectie) is het niet duidelijk hoe deze recrutering op een goede manier kan gebeuren.

Naast de personeelskost die dient ingecalculerd te worden voor werving en opleiding van deze personen is het ook belangrijk om een goede inschatting te maken van het aantal toetsinspecteurs dat vereist is om

de externe kwaliteitsmonitoring te realiseren. Het aantal toetsinspecteurs dat nodig is hangt af van een aantal parameters. In eerste instantie dient bepaald te worden hoeveel toetslocaties bezocht dienen te worden (bv. 10% of 5% van de toetslocaties). Bovendien moet ook rekening gehouden met de geografische spreiding van de toetslocaties om verplaatsingen met daarbij horende kosten zoveel mogelijk te beperken. Tot slot moet ook rekening worden gehouden met het afnamevenster (zie hoofdstuk Tijdsscenario's). Indien de toetsen over meerdere dagen kunnen afgelegd worden, kunnen de toetsinspecteurs meerdere toetslocaties bezoeken en hoeven er minder kwaliteitsmonitoren te worden gerecruteerd.

4.5. BIJKOMENDE PROBLEMATIEK: OPEN VRAGEN

Een bijkomend gegeven dat onder de loep moet genomen worden is het scoren van toetsitems. Vanuit het uitgangspunt dat de toetsen op digitale wijze worden afgenomen, spreekt het voor zich dat vragen met korteantwoordvorm (waaronder meerkeuzevragen) automatisch gescoord worden. Het zou echter kunnen dat er ook open vragen gesteld zullen worden, bijvoorbeeld voor het evalueren van de schrijfvaardigheid in het Nederlands waarvoor schrijftaken nodig zijn. Deze problematiek werd hierboven al besproken m.b.t. de doorlooptijd van de gecentraliseerde toetsen. Zelfs als het gaat om slechts één schrijftaak, dan is de omvang van deze beoordelingsopdracht erg groot gegeven het aantal deelnemende leerlingen en daaraan corresponderend het grote aantal beoordelaars.

Men zou er voor kunnen opteren om deze beoordelingen centraal uit te voeren met extern te werven beoordelaars, die na een training en een beoordelaarsovereenstemmingsonderzoek zelfstandig aan de slag kunnen. Stel dat één beoordelaar daarbij 500 leerlingprotocollen kan beoordelen, dan zijn er per beoordelingstaak 140 beoordelaars nodig. Deze personen kunnen gerecruteerd worden onder studenten uit de lerarenopleiding, het lerarenkorps zelf indien men deze scoring kan laten uitvoeren tijdens de zomervakantie. Dat is echter te laat om nog resultaten op leerlingniveau te kunnen terugkopen.

Om toch te kunnen realiseren dat er voor het einde van het schooljaar ook leerlingenresultaten zijn bij open vragen, dan lijkt het nodig om de scoring ervan op decentrale wijze te organiseren. In de Franse gemeenschap worden voor het beoordelen van de CEB, CED1 en CESS bijvoorbeeld regionale verbetercentra opgericht, waar leerkrachten die door hun school worden afgevaardigd o.l.v. een lid van de inspectie de toetsen van de leerlingen verbeteren. Daarbij gaat het ook om het verbeteren van vragen met korteantwoordvormen. Een andere strategie bestaat erin om de leerkrachten van de betrokken scholen zelf de open vragen te laten beoordelen. In beide scenario's moeten de leerkrachten daarbij ondersteund worden door uitgebreide scoringsvoorschriften.

4.6. TOETSVERVERSING

In de volgende paragrafen situeren we het begrip toets- of itemverversing¹⁴ en het belang van confidentialiteit van de items. Vervolgens bespreken we in de volgende sectie hoe praktische aspecten van de afname de confidentialiteit van de items en de nood tot verversen van de items beïnvloeden.

4.6.1. CONCEPTEN M.B.T. TOETSVERVERSING

OMSCHRIJVING VAN HET BEGRIIP TOETSVERVERSING

Toetsverversing houdt in dat op verschillende momenten de items van de toetsen geheel of gedeeltelijke vervangen moeten worden (Stocking, 1994; Way, 1998; Zhang, 2014). Het voornaamste argument is dat wanneer de confidentialiteit of geheimhouding van de items onvoldoende gewaarborgd is, en er mogelijk leerlingen zijn die de items op voorhand kennen, deze leerlingen een voordeel zullen hebben bij de toetsafname. Het beter scoren van deze leerlingen is dan een gevolg van hun voorkennis van de items en niet van de vaardigheid die de toets dient te meten. Een ander argument voor toetsverversing is de mogelijke veroudering van de items, door bijvoorbeeld aanpassingen aan eindtermen of leerplannen of door een maatschappelijke context die niet meer relevant is. Ten slotte kan er ook uit itemanalyses na de toetsafname blijken dat bepaalde items niet de verwachte meetkwaliteiten hebben, en deze dienen dan ook vervangen te worden. Er zijn dus deze meerdere argumenten om items te verversen, maar het voornaamste argument voor toetsverversing is wanneer de geheimhouding van items niet meer gevrijwaard is. De mate waarin toetsverversing zal nodig zijn is dan ook in sterke mate afhankelijk van de mate waarin de confidentialiteit van de items bewaard kan worden.

CONFIDENTIALITEIT VAN ITEMS: EEN CATEGORISCH EN EEN CONTINU PERSPECTIEF

In de literatuur over itemverversing wordt de confidentialiteit van items op twee manieren beschreven (Stocking, & Lewis, 1998; Zhang, 2016). De eerste manier is categorisch en sluit aan op een eerder intuïtieve denkwijze waarbij de geheimhouding van een item al dan niet gewaarborgd is. De tweede manier is continu en sluit aan bij de situatie waarin items over verschillende afnamecyclussen bij grote groepen leerlingen worden afgenomen.

Categorische visie

Bij de categorische visie over de confidentialiteit van items, wordt de geheimhouding van een item ofwel als gewaarborgd beschouwd of als verbroken beschouwd. Als de confidentialiteit van een item verbroken is, dan kan het item niet meer opgenomen worden in een volgende afnamecyclus. Wanneer de confidentialiteit van een item gewaarborgd is, kan het wel opgenomen worden in een volgende

¹⁴ Itemverversing heeft betrekking op het vervangen van items in een toets, terwijl toetsverversing in principe betrekking heeft op het vervangen van een hele toets.

afnamecyclus. De redenering daarbij is dat de confidentialiteit van een item volgens één of meerdere eenvoudige regels bepaald kan worden. De volgende regels komen vaak voor:

Een item kan slechts een bepaald aantal afnamecyclussen worden opgenomen. Er kan bijvoorbeeld bepaald worden dat elk item slechts éénmaal onderdeel mag zijn van een afnamecyclus. Bij elke afnamecyclus moet er dan een nieuw geheel aan items ontwikkeld worden. Er kan ook bepaald worden dat items meerdere afnamecyclussen worden opgenomen. Als bijvoorbeeld elk item drie afnamecyclussen wordt opgenomen, dat moet bij elke afnamecyclus één derde van de items ververst worden.

Een item kan niet meer opgenomen worden wanneer het item bij (een gedeelte van) het bredere publiek bekend is, omdat daarmee de confidentialiteit van het item duidelijk verbroken is. Dit kan intentioneel zijn, wanneer men het bredere publiek wil inlichten over de toetsen en er ter illustratie ook enkele items getoond worden. Een ander voorbeeld is dat men kan kiezen om na elke afnamecyclus de toets volledig openbaar te maken voor het bredere publiek. Items kunnen ook onbedoeld bekend geraken, bijvoorbeeld door het lekken van items voor, tijdens of na de afnamecyclus. Ook in deze gevallen moeten de items ververst worden.

Het voornaamste voordeel van de categorische visie over de confidentialiteit van items is de duidelijkheid. Een item kan ofwel in een volgende afnamecyclus opgenomen worden, of het item kan niet meer opgenomen worden. Er moet echter wel bepaald worden hoeveel afnamecyclussen ieder item opgenomen zal worden. Deze keuze is sterk afhankelijk van zowel het doel van de toetsen als van de praktische aspecten van de toetsafname. Het voornaamste nadeel van deze categorische benadering is dat het alles of niets is. Dit is niet optimaal wanneer een adaptieve toets wordt gebruikt, want dan zijn er items die veel worden aangeboden en items die weinig worden aangeboden aan leerlingen. Met een categorische benadering van de confidentialiteit van items, zullen weinig gebruikte items even vaak ververst worden als veel gebruikte items. Een ander nadeel is dat het al dan niet verbroken zijn van de geheimhouding niet altijd opgemerkt zal worden. Items kunnen bijvoorbeeld vrij circuleren op bepaalde (internet)fora zonder dat de aanbieder van de toets daar weet van heeft.

Continue visie

De continue visie houdt in dat er voor elk item een risicoschatting gebeurt dat de confidentialiteit verbroken is. Wanneer de risicoschatting van een item laag is, kan het terug opgenomen worden in een volgende afnamecyclus. De redenering is dat de confidentialiteit van een item een relatief gegeven is, en er hiervoor verschillende indicatoren bekeken moeten worden. Er kunnen hiervoor bijvoorbeeld volgende indicatoren gebruikt worden.

Het risico op het verbroken zijn van de confidentialiteit wordt vaak bepaald aan de hand van de relatieve mate waarin een item aan leerlingen is aangeboden, de *exposure rate*. Wanneer deze *exposure rate* een bepaalde drempel overschrijdt na een afnamecyclus, wordt het risico op het verbreken van de confidentialiteit als te hoog beschouwd. Dit item kan dan niet meer opgenomen worden in de volgende

afnamecyclus. Afhankelijk van het doel van de toetsen en de praktische aspecten van de afname kan er een bepaalde drempel bepaald worden.

Van items die worden opgenomen in verschillende afnamecyclussen kan bijgehouden worden hoe hun psychometrische kwaliteiten veranderen over de afnamecyclussen heen. Als één item relatief eenvoudiger wordt over de afnamemomenten heen, dan kan dit betekenen dat dit item mogelijk is 'gelekt' en scholen/leerlingen specifiek oefenen op dit item. Afhankelijk van het al dan niet vinden van zo'n verschuiving, kan besloten worden het item te verwijderen. Soms wordt ook nagegaan of de responstijd van een item verandert, als deze gemiddeld korter is dan voorheen, dan kan dit als een indicatie beschouwd worden dat het item niet meer geheim is.

Het voornaamste voordeel van de continue visie is dat het al dan niet opnemen van een item in de volgende afnamecyclus beter is afgestemd op het eigenlijke risico dat de geheimhouding van een item is verbroken. Bijvoorbeeld, bij adaptieve toetsen zullen heel gemakkelijke en heel moeilijke items, die enkel aangeboden zijn bij een klein aantal leerlingen, minder vaak ververst moeten worden. Er moet echter wel bepaald worden welke *exposure rate* als te risicovol wordt beschouwd. Deze keuze is sterk afhankelijk van zowel het doel van de toetsen als van de praktische aspecten van de toetsafname. Een ander voordeel van de continue visie is dat als men niet weet of de geheimhouding van een item verbroken is, men dit toch analytisch kan vaststellen. Een nadeel van deze continue benadering is dat er na een afnamecyclus diepgaande analyses nodig zijn om vast te stellen of de confidentialiteit van de items bewaard is gebleven. Aansluitend zal ook pas nadat een afnamecyclus heeft plaatsgevonden opgemerkt worden dat de confidentialiteit van items verbroken is. Bij een korte doorlooptijd van de toetsen is er geen tijd om deze items te weren bij de vaardigheidsschattingen van de leerlingen. Het is de vraag of er snelle en heldere indicatoren kunnen ontwikkeld worden om items uit te sluiten bij net afgesloten toetsen, zodat de informatie uit deze posthoc analyses wel nog kan worden meegenomen bij het terugkoppelen van de resultaten.

Ten slotte merken we op dat de categorische en continue visie over de confidentialiteit van items elkaar natuurlijk niet uitsluiten en beiden gehanteerd kunnen worden. Zo kan een item bijvoorbeeld uitgesloten zijn van de volgende afnamecyclus omdat deze (bedoeld of onbedoeld) openbaar is gemaakt. De confidentialiteit van dit item is dan duidelijk verbroken. Tegelijkertijd kan het risico van elk item dat niet openbaar is gemaakt nog geëvalueerd worden via de *exposure rate* of andere statistieken.

HET BELANG VAN DE CONFIDENTIALITEIT VAN DE ITEMS

Hierboven beschreven we al dat de confidentialiteit van items belangrijk is omdat voorkennis over de items leerlingen een voordeel kan geven bij de toetsafname. Een toetsscore dient immers de vaardigheid te reflecteren die de toets dient te meten, niet de voorkennis die leerlingen hebben over items. Leerlingen met voorkennis over de items hebben dus een onterecht voordeel tegenover leerlingen zonder voorkennis van de items. Hoewel het onterecht bevoordelen van leerlingen met voorkennis over de items het vaakst vernoemd wordt om de nood aan toetsverversing te beargumenteren, is het een onvolledig

argument. We achten het dan ook noodzakelijk om ook andere redenen te bespreken waarom het verbreken van confidentialiteit van items zo schadelijk is en regelmatige toetsverversing noodzakelijk is.

Een tweede reden om de confidentialiteit van de items te bewaren is om een negatieve vorm van *'teaching to the test'* te voorkomen. Wanneer er in 2023 immers gestart wordt met de afname van de gecentraliseerde toetsen, zullen scholen, leerkrachten en leerlingen zich afvragen hoe men zich dient voor te bereiden op deze toetsen. Wanneer er al verschillende afnamecycli van de gecentraliseerde toetsen hebben plaatsgevonden, zullen leerlingen en scholen waarschijnlijk ook enige kennis hebben van de inhoud die in deze toetsen aan bod komen. Zeker wanneer scholen of leerkrachten een persoonlijk belang ervaren bij de resultaten van de leerlingen op de gecentraliseerde toetsen (cf. infra), zal men proberen het lesgeven af te stemmen op de gecentraliseerde toetsen. Dit heet *'teaching to the test'*.

Welsh, Eastwood en D'Agostino (2014) suggereren echter dat *'teaching to the test'* een problematische term is en dat *'item teaching'* onderscheiden moet worden van *'curriculum teaching'*. De auteurs argumenteren dat *'item teaching'* (Redden & Low, 2012) te vermijden is terwijl *'curriculum teaching'* wel wenselijk kan zijn, als de toetsen een sterke mate van inhoudsvaliditeit hebben (zie ook Popham, 2004). Een sterke mate van inhoudsvaliditeit binnen de Vlaamse context betekent dat de items van een gecentraliseerde toets samen alle facetten van het getoetste inhoudsdomein omvatten, zoals beschreven door de eindtermen. Op deze manier kunnen de toetsen dan ook dienen als een leidraad voor de te onderwijzen inhoud en de verhoopde resultaten (Goodman & Hambleton, 2005). Dit wordt ook wel *'teaching to the standard'* genoemd. Daartegenover staat *'item teaching'*, waarbij leerlingen onderwezen worden in items van vorige afnamecycli of items die hierop lijken. Door op deze manier les te geven wordt feitelijk minder nadruk gelegd op de lesinhoud, maar meer op oppervlakkige kenmerken van de items. *'Item teaching'* staat dus het verhoopde onderwijs in de weg en is dus een extra reden om de confidentialiteit van de items te beschermen.

Een derde reden waarom de confidentialiteit van de items best bewaakt wordt, is om de betrouwbaarheid en validiteit van de toetsafnames voor alle leerlingen te beschermen (Guo, Tay & Drasgow, 2009; Zhang, 2014). De toetsantwoorden van de leerlingen in een afnamecyclus zullen namelijk ook gebruikt worden om de items te kalibreren. Echter, hoe meer items er zijn waarvan de confidentialiteit verbroken is, hoe meer een kalibratie op basis van deze gegevens incorrect zal zijn. Dit leidt naar foute schattingen van itemparameters, wat bij adaptieve toetsen ertoe zal leiden dat leerlingen niet de voor hun gepaste items aangeboden krijgen. Zo kan het 'vals spelen' bij enkele items bij een minderheid van leerlingen of scholen ook een negatief effect hebben op de toetsafnames bij alle leerlingen. Ook lineaire toetsen zullen niet optimaal zijn als er bij het samenstellen van deze toetsen met foute itemparameters gewerkt wordt. Wanneer de itemparameters daarbij ook een bron van informatie zijn om te bepalen wanneer de eindtermen bereikt zijn (zoals bij de *'bookmark method'*, zie Buckendahl, Smith, Impara, & Plake, 2002), dan zullen ook hier de resultaten vertekend zijn. De betrouwbaarheid en validiteit van de toetsafnames voor alle leerlingen, de goede samenstelling van de toetsen en het correct bepalen van de mate waarin de eindtermen bereikt zijn, zijn dus ook afhankelijk van de confidentialiteit van de items.

4.6.2. PRAKTISCHE ASPECTEN VAN DE AFNAME M.B.T. TOETSVERVERSING

In deze sectie bespreken we hoe de praktische aspecten van de afname de confidentialiteit van de items beïnvloeden en daarmee ook de nood aan toetsverversing. We onderscheiden daarbij twee hoofdaspecten:

- Het toetsdoel: welke betrokkenen hebben belang bij de afname van de toets en zullen het meest geneigd zijn om de confidentialiteit van de items te verbreken?
- Mogelijke maatregelen voor het bewaken van de confidentialiteit.

TOETSDOEL

De nood aan de confidentialiteit van de items en de mate van itemverversing die daarvoor nodig is, hangt nauw samen met het precieze doel van de gecentraliseerde toetsen (International Test Commission, 2014; Wollack & Fremer, 2013). Het doel van de gecentraliseerde toetsen bepaalt immers of betrokkenen er baat bij hebben om de items te kennen vooraleer de toetsafname plaatsvindt. De betrokkenen die een persoonlijk belang kunnen hebben zijn ofwel de leerlingen, de leerkrachten, de scholen of een beleidsvoerende organisatie. Wanneer een betrokkene een groot belang toedicht aan de uitkomst van een toets, ervaren zij dit als *'high-stakes'*, terwijl een betrokkene die er geen belang aan hecht het als *'low-stakes'* zal ervaren. Doorgaans wordt een toets als *'high-stakes'* ervaren als men verwacht dat er persoonlijke consequenties verbonden zijn aan het toetsresultaat. Naargelang het toetsdoel zullen de *'stakes'* verschillen voor de betrokkenen en dus ook waar de voornaamste risico's liggen voor het doorbreken van de confidentialiteit van de items.

Als de gecentraliseerde toetsen gebruikt zullen worden om de (school)loopbanen van leerlingen te sturen, dan zijn de toetsen *'high-stakes'* voor de leerlingen. Dit zou bijvoorbeeld het geval zijn wanneer een toets op het einde van het zesde leerjaar lager onderwijs zou bepalen of een leerling toegang heeft tot de A-stroom. De toets is dan summatief van aard voor de leerlingen. Hieruit volgt dat het voornamelijk leerlingen (of hun verwanten) zullen zijn die de geheimhouding van de items zullen willen doorbreken. Leerlingen (of hun verwanten) zullen dan trachten om de items te kennen die in de voorgaande afnamecyclussen zijn afgenomen. Mogelijk kunnen echter ook leerkrachten en scholen geneigd zijn om de geheimhouding van de items te doorbreken, aangezien de meeste leerkrachten en scholen ook willen dat hun leerlingen slagen.

Als de gecentraliseerde toetsen gebruikt zullen worden om de prestatie van scholen te evalueren, en deze evaluatie gevolgen heeft voor de school, dan zijn de toetsen *'high-stakes'* voor de scholen. De toetsen worden hierbij voornamelijk afgenomen vanuit een zogenaamd verantwoordingsperspectief voor de scholen. De toets is dan summatief van aard voor de scholen, en de scholen moeten zich verantwoorden voor de prestaties van hun leerlingen. Hieruit volgt dat het voornamelijk scholen zullen zijn die de geheimhouding van de items zullen willen doorbreken. Scholen zullen dan trachten om de items te

kennen die in de voorgaande afnamecyclussen zijn afgenomen. In deze situatie zal toetsafname door leerlingen als eerder *'low-stakes'* ervaren worden.

Als de gecentraliseerde toetsen gebruikt zullen worden om de prestatie van leerkrachten te evalueren, en deze evaluatie gevolgen heeft voor de leerkracht, dan zijn de toetsen *'high-stakes'* voor de leerkrachten. De toets is dan summatief van aard voor de leerkrachten, en de leerkrachten moeten zich verantwoorden voor de prestaties van hun leerlingen. Hieruit volgt dat het voornamelijk leerkrachten zullen zijn die de geheimhouding van de items zullen willen doorbreken. Leerkrachten zullen dan trachten om de items te kennen die in de voorgaande afnamecyclussen zijn afgenomen. Het is daarbij waarschijnlijk dat een school er belang bij heeft dat de eigen leerkrachten een goede beoordeling krijgen, en de toetsen zullen dan ook door de school als *'high-stakes'* ervaren worden. Ook in deze situatie zal de toetsafname door leerlingen wellicht als eerder *'low-stakes'* ervaren worden.

Als de gecentraliseerde toetsen ingezet worden voor formatieve doeleinden, dan zal de toets als *'low-stakes'* ervaren worden door leerlingen, leerkrachten en scholen. Een formatief doeleind op leerlingniveau zou kunnen zijn dat de toetsen enkel dienen om leerlingen een advies te geven over hun studiekeuze, zonder dat het toetsresultaat een leerling echt dwingt tot een bepaalde studiekeuze. Een formatief doeleind op schoolniveau zou kunnen zijn dat de toetsen enkel dienen om scholen hun eigen prestaties te evalueren, zonder dat de school zich zou moeten verantwoorden over het verkregen resultaat. In deze situatie is er dus minder risico dat de confidentialiteit van de items verbroken wordt.

Voor hoe meer betrokkenen de gecentraliseerde toetsen echter als *'high-stakes'* worden ervaren, hoe meer maatregelen er genomen zullen moeten worden om de confidentialiteit van de items te beschermen. Welke betrokkenen de gecentraliseerde toetsen als *'high-stakes'* ervaren beïnvloedt ook welke maatregelen het meest nodig zullen zijn om de confidentialiteit van de items te beschermen. Aansluitend zal het tempo van de toetsverversing ook afhangen van hoe *'high-stakes'* de toetsen ervaren worden door de verschillende betrokkenen.

De Beleidsnota Onderwijs 2019-2024 stelt dat de belangrijkste doelen van het systeem van gecentraliseerd toetsen bestaan uit het beschrijven van de leerwinst op leerling en schoolniveau, en het beschrijven van hoeveel leerlingen de eindtermen behalen. Er is echter nog niet beslist of scholen zich zullen moeten verantwoorden voor hun gemiddelde leerwinst en of hun resultaten openbaar zullen worden. Aansluitend is het ook niet geweten of leerlingen beoordeeld zullen worden aan de hand van de gecentraliseerde toetsen. De mate waarin leerlingen, leerkrachten of scholen de gecentraliseerde toetsen als *'high-stakes'* zullen ervaren kan dus nog niet bepaald worden, en daarmee ook niet de mate waarin de confidentialiteit van de items beschermd moet worden.

MOGELIJKE BESCHERMINGSMATREGELEN

De International Test Commission (ITC, 2014) beschrijft hoe verschillende aspecten van de toetsafname de geheimhouding van de items kunnen bedreigen. Hieruit kunnen ook richtlijnen afgeleid worden die

specifiek relevant zijn om de confidentialiteit te beschermen. We bespreken deze deels verregaande richtlijnen hieronder om te illustreren waarop kan ingezet worden om de geheimhouding van items zoveel mogelijk te garanderen. Daarbij willen we wel vooraf opmerken dat de ITC deze richtlijnen formuleert in de veronderstelling dat de toetsafname *'high stakes'* is voor de scholen en/of voor de leerlingen.

De ITC benadrukt in de eerste instantie dat steeds geweten moet zijn tijdens de afname wie er toegang heeft tot de toetsen en wanneer zij deze toegang hebben. Gezien Vlaanderen ervoor kiest om digitale toetsen af te nemen, zal er enkel toegang zijn tot de toetsen wanneer de toetsafname van de leerlingen van een school gepland is. Vanzelfsprekend zullen leerlingen tijdens hun eigen toetsafname toegang hebben tot de items die hun aangeboden worden. Een zeker risico voor de confidentialiteit van de items is in die zin dan ook onvermijdelijk aanwezig. Echter, als de testlocatie de eigen school is, zullen leerkrachten (of andere medewerkers van de school) ook fysiek in staat zijn om de toetsafname mee op te volgen, zelfs wanneer dit niet de bedoeling zou zijn. Bij een toetsafname op school is er dus een groter risico op het verbreken van de confidentialiteit van de items, eenvoudigweg omdat er meer betrokkenen aanwezig zijn voor wie de toetsresultaten *'high stakes'* zouden kunnen zijn.

Om onterechte vrijgave van items te vermijden raadt de ITC aan om een systeem voor leerlingregistratie (zie ook het ICT-luik van Perceel 3) op te zetten om te voorkomen dat derden ongeoorloofd toegang hebben tot het toetsmateriaal. In Vlaanderen zou bijvoorbeeld gewerkt kunnen worden met een elektronische identiteitskaart, al is dat voor de leerlingen van de basisschool allicht minder evident. Zeker wanneer de toetsafname plaatsvindt in de eigen school zou er zo'n leerlingregistratiesysteem aanwezig moeten zijn, anders kunnen derden te gemakkelijk toegang krijgen tot het toetsmateriaal. Daarnaast moet er bij *'high stakes testing'* geregistreerd worden wie er aanwezig is in het testlokaal terwijl de toetsafnames bezig zijn, en waarom zij daar aanwezig zijn. In sommige gevallen zal dit een voorziene aanwezigheid zijn (bijvoorbeeld, een persoon die functioneert als toetsassistent, cf. infra) of een onvoorziene aanwezigheid zijn (bijvoorbeeld, een persoon die een onverwacht technisch probleem moet oplossen). In beide gevallen moet de aanwezigheid geregistreerd worden. Bij een testafname in een apart testcentrum is het risico op de aanwezigheid van derden kleiner dan wanneer de toetsafname in de eigen school plaatsvindt.

De ITC richtlijnen voor het bewaken van itemconfidentialiteit beschrijven ook dat er bij elke toetsafname toetsassistent(en) aanwezig moeten zijn, waarbij hun hoofdtak het begeleiden van de toetsafname is, waarbij ook de geheimhouding van de items bewaard blijft. Deze toetsassistenten moeten ook opgeleid zijn om de afname te begeleiden en kunnen in beperkte mate technische ondersteuning geven bij het leerlingregistratiesysteem en de toetsafname. Zij mogen echter geen inhoudelijke ondersteuning bieden tijdens de toetsafname en zijn best ook niet inhoudelijk vertrouwd met de toetsafname. Bijvoorbeeld, bij een gecentraliseerde toets voor wiskunde kunnen de toetsassistenten best geen wiskundige achtergrond hebben. Dit vermindert het risico dat zij inhoudelijke ondersteuning kunnen geven, inhoudelijke interesse zouden hebben in de toetsen of dat er sprake kan zijn van belangenverstrengeling in hun rol als toetsassistent.

De ITC richtlijnen benadrukken dat een leerkracht geen goede toetsassistent kan zijn omdat er vanzelfsprekend een belangenverstrengeling is tussen hun rol als leerkracht en hun rol als toetsassistent. Bekeken vanuit het standpunt van het zoveel mogelijk voorkomen van mogelijke toetsfraude, is het inzetten van externe toetsassistenten wanneer de afname op school plaatsvindt de beste keuze. De organisatorische haalbaarheid van deze optie werd hierboven echter als zeer miniem ingeschat, zeker als de afname tegelijkertijd plaatsvindt in de meeste scholen.

Daarnaast kunnen in principe ook nog de volgende preventieve maatregelen tijdens en na de afname genomen worden om het risico op het verbreken van de confidentialiteit van de items te verminderen.

- *Het inzetten van online toetsassistenten ('online proctors', Foster, 2009).* Een online toetsassistent houdt in dat er in het testlokaal op school een webcam wordt geïnstalleerd die een volledig overzicht geeft over alle deelnemende leerlingen. Op deze manier kan iemand op afstand (bijvoorbeeld, medewerkers of jobstudenten van het steunpunt) zien hoe de afname in het testlokaal verloopt. Dit vereist natuurlijk wel dat er voor ieder testlokaal minstens één webcam wordt voorzien, dat deze zo geplaatst kan worden zodat er een goed overzicht is over alle deelnemende leerlingen, en dat dit ook technisch haalbaar is voor een school om uit te voeren. Het systeem van online toetsleiders kan ook uitgebreid worden zodat iedere leerling een aparte webcam heeft, maar dit vereist niet alleen meer infrastructuur, maar verhoogt ook het aantal taken dat de online toetsassistent moet uitvoeren.
- *Automatische monitoring van de afname.* De software van de toetsafname kan de afname automatisch monitoren (Foster, 2009). Zo kan bijvoorbeeld gescreend worden op wanneer leerlingen 'inconsistent antwoorden'. Inconsistent antwoorden kan bijvoorbeeld zijn dat leerlingen moeilijke items juist beantwoorden, maar makkelijke items fout beantwoorden. Ook responstijden kunnen van deze screening een onderdeel zijn. Zo kan het erg snel juist beantwoorden van items als een indicatie beschouwd worden dat het item op voorhand gekend was. Daarbij kan ook worden nagegaan in welke scholen veel inconsistente patronen worden vastgesteld. Tegelijkertijd kan er tijdens de afname automatisch bijgehouden worden welke toetscombinaties er uitgevoerd worden (bijvoorbeeld, Alt-Tab, Prnt Scrn, ...). Deze automatische monitoring kan mogelijk ook zichtbaar zijn voor de hierboven vernoemde online toetsassistent. Dit verhoogt echter het aantal taken dat de online toetsassistent moet uitvoeren, en is mogelijk ook een technische uitdaging. Daarnaast is het mogelijk dat deze gegevens enkel gebruikt worden om na de toetsafname te analyseren, wat dan weer een extra taak voor de analisten inhoudt. Merk ook op dat deze screenings indicaties zijn van mogelijk inconsistent en frauduleus gedrag, maar daarom nog geen effectief bewijs daarvoor leveren.
- *Multiple-pool design bij het aanbieden van items.* Als er adaptieve toetsen gebruikt worden, kan het ingebouwd worden dat items die in een voorgaande afnamecyclus in de school gebruikt werden, niet meer in de huidige afnamecyclus opgenomen kunnen worden (Zhang & Chang,

2012). Dit heet een *'multiple-pool'* design, waarbij er in elke school per afnamecyclus slechts uit één *pool items* geselecteerd kan worden. Zo kan kennis van de items uit voorgaande afnamecyclussen niet hergebruikt worden, in de veronderstelling dat het delen van items enkel binnen de school zelf gebeurt. Een duidelijk nadeel van het *'multiple-pool'* design is dat er reeds vanaf de start een grote itemdatabank aanwezig moet zijn.

- *Vergelijking van oude items en nieuwe items door middel van posthoc analyses.* Na iedere afnamecyclus kan ook onderzocht worden of scholen beter presteren op oude items in vergelijking met nieuwe items (Wang & Liu, 2020). Zo kan vastgesteld worden of er scholen zijn waar naar grote waarschijnlijkheid geoefend is op de items van voorgaande afnamecyclussen. Dit veronderstelt echter dat er ieder jaar zowel voldoende nieuwe items worden afgenomen en voldoende oude items worden afgenomen om dit vergelijk mogelijk te maken.
- *Vergelijking van de feitelijke en verwachte prestaties van scholen.* Wanneer in het kader van leerwinstmetingen scholen longitudinaal opgevolgd worden, zou men aan de hand van statistische modellen ook kunnen onderzoeken of er scholen zijn die zeer onverwachte toetsresultaten behalen (zie bijvoorbeeld Christiansen, Janssen, & Bayes, 2020). Deze post-hoc screening is op zich geen bewijs voor toetsfraude maar kan dan worden ingezet om bijvoorbeeld bij bepaalde scholen het verloop van de toetsafname opnieuw te bekijken.

We merken ten slotte op dat de bovenstaande ITC richtlijnen geformuleerd zijn in de veronderstelling dat de toetsafname *'high stakes'* is voor de scholen en/of leerlingen. In het geval dat Vlaanderen zou kiezen voor een formatief *'low stakes'* gebruik van de toetsen, hebben scholen en leerlingen er zelf meer belang bij dat de confidentialiteit van de items bewaard blijft. In dit geval kan er meer ingezet worden op het verantwoordelijkheidsgevoel van de scholen en leerkrachten.

De keuze voor *"high stakes"* toetsen op school-of leerlingniveau vereist trouwens niet enkel beschermingsmaatregelen tijdens en na de afname, maar ook een veiligheidsplan tijdens de toetsontwikkeling zelf. Het risico om de confidentialiteit van de items te doorbreken is immers ook iets waarmee rekening gehouden moet worden in het nieuwe steunpunt zelf. Elke steunpunt-medewerker die toegang heeft tot de items vormt immers ook een risico om de confidentialiteit van de items te doorbreken. Dat aspect van confidentialiteit kan dus niet wachten tot 2022 of 2023, maar is iets waar eigenlijk aan de start rekening mee gehouden worden. Het nieuwe steunpunt zou dan een veiligheidsplan moeten uitwerken dat omschrijft wie toegang heeft tot de items (en die toegang moet voor iedereen duidelijk verantwoord zijn), hoe je toegang kan hebben tot het toetsmateriaal (enkel op locatie na een identificatie). Ook moet er allicht clausules worden opgenomen in de arbeidscontracten van de steunpuntmedewerkers waarbij er juridisch gespecificeerd staat wat bijvoorbeeld de gevolgen zijn van het lekken van toetsmateriaal). De keuze voor *"high stakes"* betekent daarbij ook dat het moeilijker zal zijn om met externe partners samen te werken rond itemontwikkeling, omdat in zo'n scenario elke externe partner immers een risico vormt.

BESLUIT

Er zijn verschillende manieren hoe praktische aspecten van de toetsafname de confidentialiteit van de item beïnvloeden. De mate waarin de confidentialiteit bewaard moet worden hangt dan ook sterk af van het toetsdoel en deels ook van de toetslocatie. Ook kan men bijkomende maatregelen nemen om de confidentialiteit te garanderen of zoveel mogelijk te bewaken. De eventuele keuze voor dergelijke maatregelen heeft grote budgettaire implicaties. Hoe meer zekerheid men wil hebben dat de confidentialiteit behouden blijft, hoe duurder de toetsafname wordt. Gegeven de hoge kostprijs is het wellicht niet haalbaar om een volledig systeem van bijkomende controles op de toetsafnames bij de gecentraliseerde toetsen te voorzien. Bovendien kan men ook de vraag stellen of dat wenselijk zou zijn, gegeven dat het bepalen van het toetsdoel van de gecentraliseerde toetsen de noodzaak van deze beveiligingen kan doen verminderen. Als leerlingen en scholen de baten inzien van het systeem van gecentraliseerde toetsen, dan zijn ze allicht minder geneigd om het vertrouwen dat in hen gesteld wordt om de confidentialiteit te beweren te verbreken.

4.6.3. SCENARIO'S M.B.T. DE CONFIDENTIALITEIT VAN ITEMS EN ITEMVERVERSING

De confidentialiteit van de items ligt aan de grond van een betrouwbare en valide afname van de Vlaamse gecentraliseerde toetsen. De resultaten van leerlingen op de toetsen mogen namelijk niet vertekend worden door het hebben van voorkennis over de items. Daarom moet de confidentialiteit van de items in de eerste plaats zoveel mogelijk gevrijwaard worden, maar moet er in de tweede plaats ook verversing plaatsvinden van items waarvan de confidentialiteit niet meer gewaarborgd is. Daarnaast zal itemverversing ook nodig zijn bij items die verouderd zijn of niet de beoogde kwaliteiten hebben. De mate waarin toetsverversing zal nodig zijn is echter voornamelijk afhankelijk van de mate waarin de confidentialiteit van de items bewaard kan worden.

Zoals in de vorige sectie werd besproken is het risico op het verbreken van de confidentialiteit van de items afhankelijk van het al dan niet *'high-stakes'* zijn van de toetsen en in welke omgeving de toetsafname plaatsvindt. We stellen hieronder drie scenario's voor die verschillen naargelang de *'stakes'* en toetsomgeving. Het derde scenario werd voor de volledigheid toegevoegd, maar is tot nu toe niet echt in beleidsdocumenten over de gecentraliseerde toetsen aan bod gekomen.

SCENARIO 1: 'LOW-STAKES' TOETSEN MET TOETSAFNAME OP SCHOOL

In een eerste scenario zijn de gecentraliseerde toetsen formatieve *'low-stakes'* toetsen. De toetsen worden bijvoorbeeld uitgevoerd zodat scholen hun eigen functioneren kunnen evalueren. Er zijn dan geen persoonlijke consequenties verbonden aan het toetsresultaat, zowel voor leerlingen, leerkrachten als scholen. Aangezien de toetsen niet *'high stakes'* zijn, is het minder waarschijnlijk dat scholen, leerkrachten of leerlingen de items zullen proberen te bemachtigen of als leidraad voor hun lesgeven gaan gebruiken. Er is dan ook geen uitgebreide controle nodig tijdens de toetsafname en leerlingen kunnen dan ook

deelnemen aan de toets in de eigen scholen zonder externe toetsassistenten. In dit scenario kunnen items dan ook herhaaldelijk gebruikt worden, ook in dezelfde scholen.

Desondanks dat het minder waarschijnlijk is dat scholen, leerkrachten of leerlingen de confidentialiteit van de items zullen proberen te verbreken, moeten scholen en leerkrachten wel gesensibiliseerd worden. Het moet voor scholen duidelijk zijn dat het onwenselijk is om de items als lesmateriaal te gebruiken, omdat dat de validiteit en het nut van de toetsen onderuit kan halen. Verder zal een beperkte mate van itemverversing nog steeds nodig zijn om verouderde items en items met technische problemen te vervangen.

SCENARIO 2: 'HIGH-STAKES' TOETSEN VOOR SCHOLEN MET DE AFNAME OP SCHOOL

In een tweede scenario zijn de gecentraliseerde toetsen summatieve 'high-stakes' toetsen voor scholen. De toetsen worden bijvoorbeeld uitgevoerd zodat scholen geëvalueerd kunnen worden voor hun bijdrage aan de leerwinst van hun leerlingen. Deze resultaten worden vervolgens openbaar gemaakt voor de ouders van leerlingen die zich wensen in te schrijven in de school. Verder zullen scholen die significant minder bijdragen tot de leerwinst van hun leerlingen verplicht worden deel te nemen aan een interventie. Er zijn voor de school dus merkbare consequenties.

Omdat de afname op school plaatsvindt is er uitgebreide controle nodig tijdens de toetsafname. Scholen hebben immers zelf belang bij hoge scores van leerlingen. Sommige scholen zullen dan mogelijk trachten om de items te kennen die in de voorgaande afnamecyclussen zijn afgenomen. Om dit te voorkomen zullen externe toetsleiders aangeworven moeten worden die de toetsafname in de school begeleiden. Dit vereist echter een grote inzet van personeel en is een complexe organisatorische puzzel. Een andere mogelijkheid is dat er toezicht wordt gehouden via online toetsleiders door middel van webcams en automatische monitoring van de toetssoftware. Dit vereist echter technisch voorbereidingswerk bij het nieuwe steunpunt en de school, en ook een uitgebreide en volledig betrouwbare internetverbinding. Ook voor de online toetsleiders zal er extra personeel nodig zijn.

Aangezien scholen er veel belang bij hebben dat hun leerlingen hoog scoren, is het risico groter dat de confidentialiteit van items verbroken wordt. Er zal dus jaarlijks een sterke mate van itemverversing nodig om items die enkele afnamecycli gebruikt zijn te vervangen. Daarbij zullen de itemstatistieken na elke afname ook goed gemonitord moeten worden door analisten, zodat bepaald kan worden welke items niet meer geheim zijn.

SCENARIO 3: 'HIGH-STAKES' TOETSEN VOOR LEERLINGEN MET DE AFNAME IN EEN TOETSCESTRUM

In een derde scenario zijn de gecentraliseerde toetsen summatieve 'high-stakes' toetsen voor leerlingen. De toetsen worden bijvoorbeeld uitgevoerd zodat leerlingen toegewezen kunnen worden aan een bepaalde stroom of onderwijsvorm. Deze resultaten worden enkel gedeeld met de leerling en diens

ouders. Leerlingen hebben hier een enorm belang bij hun toetsscore, aangezien het hun verdere (school)loopbaan bepaalt.

Omdat de afname in een toetscentrum plaatsvindt is het mogelijk om uitgebreide controle te voorzien tijdens de afname. Dit zal echter ook nodig zijn: het belang voor de leerlingen is immers zo groot dat velen zich graag op de items zouden kunnen voorbereiden. Het risico is dan toch relatief groot dat de confidentialiteit van items doorbroken wordt. Leerlingen zullen immers sommige items nog (gedeeltelijk) herinneren na hun afname en doorgeven aan derden. Men zou dit kunnen oplossen door met dergelijk grote itemdatabank te werken dat het zo goed als onmogelijk is voor derden om een volledig overzicht te verkrijgen. Dit zou echter een grote investering van personeel en middelen in itemontwikkeling vereisen. Daarbij zou echter ook het op voorhand kennen van een gedeelte van de items leerlingen al een merkbaar voordeel kunnen geven. Een andere mogelijkheid is dus om ieder jaar eenvoudigweg alle items te vervangen. Ook dit zou een grote investering van personeel en middelen in itemontwikkeling vereisen.

4.7. OVERZICHT DEELSCENARIO'S EN BELEIDSADVIEZEN

4.7.1. OVERZICHT DEELSCENARIO'S

In dit hoofdstuk werd ingegaan op hoe de afname van de gecentraliseerde toetsen concreet zou kunnen verlopen. Daarbij worden deelscenario's uitgewerkt over waar de toetsafnames plaatsvinden, wat de planlast is van de scholen, hoe de toetsassistenten gerecruteerd worden en hoe er kwaliteitsmonitoring wordt voorzien. Indien er voor bepaalde toetsen (bv. schrijfoopdrachten bij Nederlands) open vragen worden gesteld, dan wordt onderzocht wie deze opdrachten kan beoordelen. Tot slot wordt ingegaan op de problematiek van het bewaken van de confidentialiteit van de items en de eventuele noodzaak van itemverversing. In Tabel 4.3 wordt een overzicht geboden van de beleidskeuzes die moeten worden gemaakt en welke keuzeopties er zijn binnen elk deelscenario. De vraag naar of de toetsen 'low-' of 'high-stakes' zijn, staat daarbij centraal. Dit wordt in onderstaande paragraaf verder besproken.

4.7.2. BELEIDSADVIEZEN

Zoals in paragraaf 4.6 uitvoerig wordt toegelicht hangen de beleidskeuzes binnen deze deelscenario's grotendeels samen met de vraag of de gecentraliseerde toetsen 'low-stakes' of 'high-stakes' zijn voor de leerlingen en de scholen. Bij 'low-stakes' toetsen kunnen de toetsen worden afgenomen in de scholen o.l.v. een toetsassistent die door de school werd aangeduid en met eigen leerkrachten als bijkomende toetsbegeleiders. De kwaliteitsmonitoring kan worden opgenomen door de school of zoals bij de internationale grootschalige toetsafnames kan een kwaliteitsmonitor steekproefgewijs langsgaan bij een aantal scholen. Indien de toetsen 'high-stakes' zijn voor leerlingen of scholen, dan impliceert dit een heel andere afnameprocedure. Er kan dan overwogen worden om in externe toetscentra de toetsen af te nemen omdat hierdoor heel wat controles op mogelijke toetsfraude kunnen worden voorzien, zodat ook

Tabel 4.3.

Beleidskeuzes voor de deelscenario's rond de uitvoering van de afname van de gecentraliseerde toetsen.

Deelscenario	Keuzeopties
<i>Toetslocatie</i>	Toetscentra Scholen
<i>Toetsassistenten</i>	Externen Leerkrachten uit andere scholen Lid van de school
<i>Kwaliteitsmonitoring</i>	Externen Lid van de school
<i>Beoordeling open vragen</i>	Externe beoordelaars Beoordelingscentra Leerkrachten van de school
<i>Confidentialiteit en toetsverversing</i>	'low-stakes' toetsen met toetsafname op school 'high-stakes' toetsen voor scholen met de afname op school 'high-stakes' toetsen voor leerlingen met de afname in een toetscentrum

de toetsitems confidencieel blijven. Indien de 'high-stakes' afnames in scholen doorgaan, dan vraagt dit minstens om externe toetsassistenten, waarvan de recruitering en selectie uiterst complex is. Daarnaast zijn ook heel wat andere beschermingsmaatregelen mogelijk, die echter een duidelijk kostenplaatje hebben.

5. BELEIDSSAMENVATTING EN -ADVIEZEN

Ten behoeve van de opdrachtgever worden in dit hoofdstuk onderdelen uit de voorgaande tekst hernomen om zo een overzicht te geven van de scenario-ontwikkeling binnen het organisatorische perceel van de haalbaarheidsstudie. Daarbij worden ook de beleidsadviezen bij de verschillende inhoudelijke hoofdstukken hernomen.

5.1. VOORSTELLING PERCEEL 2

5.1.1. DRIE DEELDOMEINEN MET EIGEN ONDERZOEKSVRAGEN

In het praktisch-organisatorische luik van de haalbaarheidsstudie naar de invoering van gecentraliseerde proeven in Vlaanderen worden drie deeldomeinen van onderzoeksvragen onderscheiden. Ten eerste worden scenario's uitgewerkt inzake de *brede afname* van de toetsen: kunnen ook leerlingen met specifieke onderwijsbehoeften deelnemen aan de toetsafnames en hoe kan dat kwaliteitsvol gerealiseerd worden zonder dat de standaardisatie in het gedrang komt? Ten tweede worden verschillende *tijdsscenario's* uitgewerkt. Dit gaat over de vragen in welke periode in het schooljaar en met welk afnamevenster de gecentraliseerde toetsen worden afgenomen. Daarnaast wordt uitgewerkt welke scenario's er mogelijk zijn inzake de doorlooptijd tussen het afnemen van de toetsen en het bekomen van de resultaten. Ten derde worden scenario's ontwikkeld inzake de afnamemodaliteiten, i.e., de *uitvoering van de afname*. Daarbij worden de vragen beantwoord wie betrokken is bij uitvoering van de afname en welke verantwoordelijkheden deze actoren daarbij opnemen en hoe de vertrouwelijkheid van de toetsopgaven bewaakt kan worden.

5.1.2. AANPAK

De scenario's die in Perceel 2 worden ontwikkeld, resulteren uit een grondige, inhoudelijke analyse van de onderzoeksvragen uit de drie deeldomeinen. Daarbij worden ook buitenlandse praktijkvoorbeelden ter inspiratie aangehaald. Een probleem daarbij was dat de informatie daarover voornamelijk te vinden is in de zogenaamde 'grijze literatuur', waardoor niet alle organisatorische details altijd achterhaald kunnen worden.

Bij de verschillende deelscenario's worden indien mogelijk schattingen gemaakt van de bijkomende budgettaire kost ervan. De meeste kosten worden begroot in de benodigde voltijdsequivalenten (VTE) voor een ingeschatte periode. Voor de kosten van bijvoorbeeld extra toetsontwikkeling of de ontsluiting van een digitaal toetsplatform wordt verder doorverwezen naar de rapporten van Percelen 1 en 3. Voor erg complexe deelscenario's wordt afgezien van een financiële berekening omdat de omvang van het deelscenario op zich al voldoende suggereert dat het weinig haalbaar is (zie bijvoorbeeld het deel over externe toetsassistenten).

5.1.3. SAMENHANG MET DE ANDERE PERCELEN VAN DE HAALBAARHEIDSTUDIE

De beleidskeuzes die gemaakt moeten worden binnen het praktisch-organisatorische perceel hangen deels onderling samen (cf. infra), maar worden ook bepaald in relatie tot de bevindingen uit Perceel 1 (pedagogisch-psychometrisch luik) en Perceel 3 (ict en juridisch luik) en de beleidskeuzes die daar gemaakt moeten worden. Zo wordt in dit perceel voorbijgegaan aan de scenario-ontwikkeling rond de toetsontwikkeling (bijvoorbeeld omvang itembank) en toetsafnameprocedures (bijvoorbeeld semi-adaptiviteit) uit Perceel 1 en wordt er uitgegaan van één afnamemoment per leer- of vakgebied. Met betrekking tot het ict-luik van Perceel 3 wordt in dit perceel vertrouwd op (a) een goede ict-infrastructuur binnen de scholen en een afdoende servercapaciteit; (b) de mogelijkheid van digitale toetsaanpassingen en (c) de beschikbaarheid van een digitaal platform voor registratie van scholen en leerlingen, het gedeeltelijk individualiseren van de toetsafname en het terugkoppelen van leerling- en schoolresultaten. Tot slot raakt dit perceel ook aan het juridische luik van Perceel 3. Er dienen namelijk juridisch oplossingen gevonden te worden voor (a) het delen van privacy-gevoelige informatie bij het toekennen van (deel)vrijstellingen of toetsaanpassingen; (b) de leerlingregistratie binnen en over toetsafnames; (c) mogelijke vragen tot 'opt out' door leerlingen of hun ouders en (d) eventuele eisen tot het toegankelijk maken van de gecentraliseerde toetsen voor leerlingen met een beperking.

5.1.4. WHAT'S AT STAKE?

Een probleem waarmee we geconfronteerd werden bij de uitwerking van de deelscenario's over de organisatie van de gecentraliseerde toetsen is dat de deelscenario's minstens deels samenhangen met de vraag wat het precieze doel is van de gecentraliseerde toetsen. Om het met een boutade te stellen: het is moeilijk om met betrekking tot de organisatie van de gecentraliseerde toetsen in Vlaanderen wie-, waar-, wanneer- en hoe-vragen te beantwoorden los van de waarom-vraag. Ook al zijn de krijtlijnen van het doel van de gecentraliseerde toetsen omschreven, toch zijn er nog interpretaties mogelijk, zoals bijvoorbeeld of de gecentraliseerde toetsen overheersend ondersteunings- of verantwoordingsgericht zijn en in welke mate de toetsen voor leerlingen en leerkrachten 'high stakes' zijn. Ook in Percelen 1 en 3 wordt de vraag gesteld naar wat nu precies 'op het spel' zal staan bij de gecentraliseerde toetsen. Net als in Perceel 1 wijzen we er op dat het niet evident is om verschillende toetsdoelen te combineren en dat de perceptie van de toetsfinaliteit zal bepalen hoe verschillende actoren in het onderwijsveld met de gecentraliseerde toetsen omgaan.

5.2. BREDE AFNAME

5.2.1. OVERZICHT DEELSCENARIO'S

Vanuit juridisch, onderwijskundig en toetstechnisch perspectief worden argumenten aangehaald om een brede afname bij de gecentraliseerde toetsen te beogen, wat impliceert dat alle leerlingen kunnen deelnemen, ook die met specifieke onderwijsbehoeften of die gebruik maken van speciale onderwijsleermiddelen. Daarbij wordt eerst vanuit het perspectief van de deelnemende leerlingen en

vervolgens vanuit het perspectief van de deelnemende scholen in concreto uitgewerkt wat een brede afname op organisatorisch vlak met zich meebrengt. De verschillende beleidskeuzes m.b.t. de brede afname staan opgelijst staan in Tabel 5.1.

Tabel 5.1.

Beleidskeuzes voor de deelscenario's rond brede afname van de gecentraliseerde toetsen.

Deelscenario	Keuzeopties
<i>(Deel)vrijstellingen</i>	
Voor wie?	leerlingen met IAC (ja; neen) OKAN-leerlingen (ja; neen)
Op welke manier?	op alle toetsen (zonder inhoudelijke motivatie) met deelvrijstellingen - op toetsniveau - op itemniveau
Opt out?	ja; neen;
<i>Toetsaanpassingen</i>	
Voor wie?	alle leerlingen leerlinggroepen (individuele aanvraag)
Door wie bepaald?	oplijsting door experts (top-down) in samenspraak met onderwijsveld (bottom-up)
<i>Toekenningsprocedure</i>	
	school zelf CLB toekenningscommissie
<i>Toegankelijkheid</i>	
Voor leerlingen met een visuele beperking	ja; neen;
Voor leerlingen met een auditieve beperking	ja; neen;
Voor leerlingen met een motorische beperking	ja; neen;
<i>Onderwijssettings</i>	
Deelname buitengewoon onderwijs	ja; neen; aangepaste vorm van gecentraliseerde toetsen?
Deelname privé-scholen	ja; neen;
Deelname huisonderwijs	ja; neen;

5.2.2. BELEIDSADVIEZEN

De beleidskeuzes die moeten worden genomen rond het verlenen van (deel)vrijstellingen en toetsaanpassingen zijn in eerste instantie afhankelijk van de resultaten van het juridische luik van de haalbaarheidsstudie (cf. Perceel 3). In dat perceel moet worden uitgeklaard of leerlingen of hun ouders ervoor kunnen kiezen om niet deel te nemen aan de gecentraliseerde toetsen (cf. opt-out). Perceel 3 moet daarnaast ook verduidelijken of en op welke manier de privacy-gevoelige informatie rond de inhoud van verslagen of gemotiveerde verslagen kan worden gedeeld met de overheid of het steunpunt zodanig dat aan individuele leerlingen op inhoudelijke grond (deel)vrijstellingen of toetsaanpassingen kan worden verleend. Indien de inhoud van deze verslagen beschermd is door de wet op de privacy of door het

beroepsgeheim dan is het enerzijds mogelijk om louter algemene maatregelen te nemen die inhoudelijk niet gericht zijn op de concrete leerling, maar gelden voor alle leerlingen met een (gemotiveerd) verslag. Leerlingen met een IAC kunnen dan vrijstelling bekomen op alle gecentraliseerde toetsen. De toetsen worden dan zoveel mogelijk ontworpen binnen een 'universal design for assessments' en afgenomen met toetsaanpassingen (bv. voorleessoftware) die door alle leerlingen kunnen worden gehanteerd. Bij de analyses achteraf kan bij het opstellen van schoolfeedback of het rapporteren op systeemniveau wel worden rekening gehouden met het feit of een leerling al dan niet een gemotiveerd verslag had.

Anderzijds is het in die situatie mogelijk om de gunning van (deel)vrijstellingen en toetsaanpassing aan de autonomie van de school over te laten, eventueel in samenspraak met het CLB (op voorwaarde dat deze informatie op een GDPR-conforme wijze kan worden doorgegeven binnen het toetsplatform). Dit scenario is duidelijk minder kostelijk dan het installeren van een toekenningscommissie, maar heeft als nadeel dat er verschillen tussen scholen kunnen zijn in hun toekenningsbeleid, wat de noodzakelijke standaardisatie van de afname van de gecentraliseerde toetsen in het gedrang kan brengen. Zeer duidelijke instructies aan deelnemende scholen over wat wel en niet mogelijk is, kan hiervoor een oplossing bieden. Zo gebeurt dit ook binnen de Eindtoets basisonderwijs in Nederland.

Alleszins zal de overheid moeten bepalen welke toetsaanpassingen bij de afname van de gecentraliseerde toetsen worden toegelaten. Dit is vooral een inhoudelijke discussie die deels ook raakt aan didactische opvattingen. De overheid kan opteren om de lijst van toetsaanpassingen te laten opstellen door experts (top down), dan wel om op basis van bevraging van en in overleg met het onderwijsveld deze lijst op te stellen (bottom up).

Een aantal van de te nemen beleidsbeslissingen rond brede afname worden best genomen in overleg met psychometrische experts (cf. ook Perceel 1). Zo is het verre van duidelijk of het haalbaar is om leerlingen met een IAC deelvrijstellingen te geven op items van een toets zonder dat dat de meetkwaliteit in het gedrang brengt. Ook moet in principe worden nagegaan in welke mate de itemkenmerken van toegankelijk gemaakte items nog vergelijkbaar zijn met de oorspronkelijke items, al is die controle door de kleine leerlingenaantallen van leerlingen met een auditieve, visuele of motorische beperking allicht voor de Vlaamse gecentraliseerde toetsen in de praktijk niet mogelijk. Merk op dat het beperkt aantal leerlingen met een auditieve, visuele of motorische beperking leidt tot een relatief grote kostprijs per leerling voor het toegankelijk maken van items. De beslissing over het al dan niet toegankelijk maken van items is echter ook een principiële beslissing rond inclusie, die mogelijk door leerlingen of hun ouders langs juridische weg zou kunnen worden afgedwongen. Deze laatste problematiek werd in de looptijd van het project aan Perceel 3 gesignaleerd.

Tot slot is het aan de onderwijsminister om te beslissen over de deelname van scholen buiten het reguliere onderwijs aan de gecentraliseerde toetsen. Ook hier zou het beperkte aantal leerlingen waarover dit gaat in het privé-onderwijs of huisonderwijs mee als keuzepincipe gehanteerd kunnen worden, gegeven ook de budgettaire kost om de afnames te organiseren in de schoot van de examencommissie. Voor het buitengewoon onderwijs zijn er argumenten gegeven waarom het niet wenselijk is om de

gecentraliseerde toetsen daar op grote schaal af te nemen. Dit neemt niet weg dat de gecentraliseerde toetsen mogelijk zinvol kunnen zijn bij loopbaanbeslissingen rond individuele leerlingen. Daarnaast is er mogelijk ook een vraag vanuit scholen uit het buitengewoon onderwijs naar indicatoren om in te zetten in het kader van hun interne kwaliteitszorg.

5.3. TIJDSCENARIO'S

5.3.1. OVERZICHT DEELSCENARIO'S

In dit hoofdstuk werden verschillende aspecten van de tijdsplanning van de afname van de gecentraliseerde toetsen in detail besproken. Achtereenvolgens werd ingezoomd op verschillende deelscenario's rond de planning van de afname in het schooljaar, de breedte van het afnamevenster en de doorlooptijd. Daarbij werd telkens ook verwezen naar buitenlandse praktijkvoorbeelden. De beleidskeuzes die binnen dit deeldomein moeten worden gemaakt staan weergegeven in Tabel 5.2 en worden in de volgende paragraaf vanuit beleidsperspectief besproken. Daaruit zal blijken dat de beleidskeuzes die binnen dit deeldomein worden gemaakt, sterk afhankelijk zijn van elkaar en van andere keuzes (bv. rond toetsontwikkeling).

Tabel 5.2

Beleidskeuzes voor de deelscenario's rond tijdscenari'o's van de gecentraliseerde toetsen.

Deelscenario	Keuzeopties
<i>Planning afname(s)</i>	
Tijdstip	15 april - 15 mei; 15 mei - 15 juni; eind juni;
Inhaalmoment	neen; ja (wanneer?)
Herkansingsmoment	neen; ja (wanneer?)
<i>Afnamevenster</i>	
	1 dag; 1 week; 2 weken;
<i>Doorlooptijd</i>	
	Dag erna; Binnen de week; Einde schooljaar; Begin volgend schooljaar; 6 tot 12 maanden na afname

5.3.2. BELEIDSADVIEZEN

Een aantal van de beleidskeuzes die gemaakt moeten worden i.v.m. de timing van de afname van de gecentraliseerde toetsen zijn in overleg schijnbaar relatief eenvoudig te nemen. De beslissing over wanneer de afnames in het derde trimester van het schooljaar vallen kan best genomen worden in overleg met het onderwijsveld en hoeft niet gelijklopend te zijn voor het basis- en het secundair onderwijs. De

afspraken over de doorlooptijd dienen genomen te worden in overleg met het nieuwe steunpunt. Zoals hierboven al werd aangegeven dringt er zich een gefaseerde oplevering op waarbij de doorlooptijd verschillend is voor de verschillende niveaus van resultaten (leerlingen-, school-, of systeemniveau).

De beslissing rond het afnamevenster is complexer omdat bij de keuze een afweging gemaakt moet worden tussen kosten met betrekking tot de ict-infrastructuur (cf. Perceel 3) versus met betrekking tot de toetsontwikkeling (cf. Perceel 1). Een breder afnamevenster impliceert een minder sterke piekbelasting van de afnameserver, is daardoor goedkoper, en is mogelijk voor sommige scholen meer haalbaar omdat de leerlingen gespreid de toetsen kunnen afleggen. Daar staat tegenover dat een breder afnamevenster vereist dat er meerdere toetsversies ontwikkeld worden, wat een hogere kostprijs betekent voor toetsontwikkeling en wat ook leidt tot meer complexe analyses. Dit laatste heeft mogelijk implicaties voor de doorlooptijd.

Ten slotte moet de onderwijsminister ook beslissen of er voor afwezige leerlingen een inhaalmoment moet worden voorzien en of leerlingen kunnen herkansen op de toets. Zoals bij heel wat andere beleidskeuzes die nog genomen moeten worden over de gecentraliseerde toetsen hangen beide beslissingen deels samen met de vraag of de gecentraliseerde toetsen op leerling- en schoolniveau 'low-' of 'high-stakes' zijn. Als scholen er belang bij kunnen hebben dat leerlingen niet zouden opdagen op de dag van de toetsafnames, is het misschien nodig om een inhaalmoment te voorzien om zo een beeld te krijgen van alle leerlingen op school. Als de schoolloopbaan van leerlingen deels beïnvloed kan worden door hun uitslagen op de gecentraliseerde toetsen, dan kan het nodig zijn om een herkansingsmoment te voorzien. Beide extra afnamemomenten hebben duidelijk budgettaire implicaties, zowel naar de organisatiekosten (o.a. servercapaciteit, planlast voor scholen) als naar toetsontwikkeling en data-analyse voor het garanderen van equivalente resultaten met andere toetsversies.

5.4. UITVOERING AFNAME

5.4.1. OVERZICHT DEELSCENARIO'S

Tot slot wordt ingegaan op hoe de afname van de gecentraliseerde toetsen concreet zou kunnen verlopen. Daarbij worden deelscenario's uitgewerkt over waar de toetsafnames plaatsvinden, wat de planlast is van de scholen, hoe de toetsassistenten gerecruteerd worden en hoe er kwaliteitsmonitoring wordt voorzien. Indien er voor bepaalde toetsen (bv. schrijfofdrachten bij Nederlands) open vragen worden gesteld, dan wordt onderzocht wie deze opdrachten kan beoordelen. Tot slot wordt ingegaan op de problematiek van het bewaken van de confidentialiteit van de items en de eventuele noodzaak van itemverversing. In Tabel 5.3 wordt een overzicht geboden van de beleidskeuzes die moeten worden gemaakt en welke keuzeopties er zijn binnen elk deelscenario. De vraag naar of de toetsen 'low-' of 'high-stakes' zijn, staat daarbij centraal.

Tabel 5.3.

Beleidskeuzes voor de deelscenario's rond de uitvoering van de afname van de gecentraliseerde toetsen.

Deelscenario	Keuzeopties
<i>Toetslocatie</i>	Toetscentra Scholen
<i>Toetsassistenten</i>	Externen Leerkrachten uit andere scholen Lid van de school
<i>Kwaliteitsmonitoring</i>	Externen Lid van de school
<i>Beoordeling open vragen</i>	Externe beoordelaars Beoordelingscentra Leerkrachten van de school
<i>Confidentialiteit en toetsverversing</i>	'low-stakes' toetsen met toetsafname op school 'high-stakes' toetsen voor scholen met de afname op school 'high-stakes' toetsen voor leerlingen met de afname in een toetscentrum

5.4.2. BELEIDSADVIEZEN

Zoals in paragraaf 4.6 uitvoerig wordt toegelicht hangen de beleidskeuzes binnen deze deelscenario's grotendeels samen met de vraag of de gecentraliseerde toetsen 'low-stakes' of 'high-stakes' zijn voor de leerlingen en de scholen. Bij 'low-stakes' toetsen kunnen de toetsen worden afgenomen in de scholen o.l.v. een toetsassistent die door de school werd aangeduid en met eigen leerkrachten als bijkomende toetsbegeleiders. De kwaliteitsmonitoring kan worden opgenomen door de school of zoals bij de internationale grootschalige toetsafnames kan een kwaliteitsmonitor steekproefgewijs langsgaan bij een aantal scholen. Indien de toetsen 'high-stakes' zijn voor leerlingen of scholen, dan impliceert dit een heel andere afnameprocedure. Er kan dan overwogen worden om in externe toetscentra de toetsen af te nemen omdat hierdoor heel wat controles op mogelijke toetsfraude kunnen worden voorzien, zodat ook de toetsitems confidencieel blijven. Indien de 'high-stakes' afnames in scholen doorgaan, dan vraagt dit minstens om externe toetsassistenten, waarvan de recrutering en selectie uiterst complex is. Daarnaast zijn ook heel wat andere beschermingsmaatregelen mogelijk, die echter een duidelijk kostenplaatje hebben.

6. REFERENTIES

- ACARA (2016, 5 oktober 2020). *NAP National Assessment Program*. <https://www.nap.edu.au/>
- ACARA (2020a). *National protocols for test administration on paper - NAPLAN 2021*. <https://nap.edu.au/docs/default-source/resources/national-protocols-for-test-administration---naplan-paper-2021.pdf>
- ACARA (2020b). *National protocols for test administration online - NAPLAN 2021*. <https://nap.edu.au/docs/default-source/resources/national-protocols-for-test-administration---naplan-online-2021.pdf>
- ACARA (2020c). *Handbook for principals and NAPLAN coordinators - NAPLAN Online*. https://k10outline.scsa.wa.edu.au/data/assets/pdf_file/0006/592809/HFPNC-NAPLAN-Online-2020_Web.PDF
- AHOVOKS (augustus 2019). *Grootschalige toetsing op maat van de nieuwe eindtermen: literatuurstudie over competentiegericht evalueren met een kleinschalige praktijkstudie a.d.h.v. voorbeelden van toetsvragen* (niet gepubliceerd rapport). Brussel: AHOVOKS.
- AGODI (2018). *Huisonderwijs Jaarrapport 2015-2016 2016-2017*. Agentschap voor Onderwijsdiensten. https://www.agodi.be/sites/default/files/atoms/files/AGODI_Huisonderwijs_Jaarrapport_2015-2016_2016-2017.pdf
- AGODI (2019). *Speciale Onderwijsleermiddelen (SOL) 2017-2018*. Brussel: Agentschap voor Onderwijsdiensten. <https://www.agodi.be/sites/default/files/atoms/files/sol2018.pdf>
- AGODI (2020). *AGODI jaarverslag 2018*. Brussel: Agentschap voor Onderwijsdiensten. https://www.agodi.be/sites/default/files/atoms/files/AGODI%20JV%202018%20WEB%20V2_0.pdf
- Beuchert, L.V. & Nandrup, A.B. (2018). The Danish national tests at a glance. *Nationaløkonomisk Tidsskrift*, 2, https://childresearch.au.dk/fileadmin/childresearch/dokumenter/2018_1_2.pdf
- Blömeke, S., & Gustafsson, J.-E. (Eds.) (2017). *Standard setting in education: The Nordic countries in an international perspective*. Springer.
- Bolt, S. E., & Thurlow, M. L. (2004). Five of the most frequently allowed testing accommodations in state policy: Synthesis of research. *Remedial and Special Education*, 25(3), 141–152. <https://doi.org/10.1177/07419325040250030201>
- Bolt, S. E., & Ysseldyke, J. (2008). Accommodating students with disabilities in large-scale testing: A comparison of Differential Item Functioning (DIF) identified across disability types. *Journal of Psychoeducational Assessment*, 26(2), 121–138. <https://doi.org/10.1177/0734282907307703>
- Bonaccio, S., & Reeve, C. L. (2010). The nature and relative importance of students' perceptions of the sources of test anxiety. *Learning and Individual Differences*, 20(6), 617-625.
- Brossman, B. G., & Guille, R. A. (2014). A comparison of multi-stage and linear test designs for medium-size licensure and certification examinations. *Journal of Computerized Adaptive Testing*, 2(2), 18-36.
- Buckendahl, C. W., Smith, R. W., Impara, J. C., & Plake, B. S. (2002). A comparison of Angoff and Bookmark standard setting methods. *Journal of Educational measurement*, 39(3), 253-263.
- Buzick, H.M. & Jones, N.D. (2015). Using test scores from students with disabilities in teacher evaluation. *Educational Measurement: Issues and Practice*, 34, 28-38. <https://doi.org/10.1111/emip.12076>
- BVGT (2021, 5 januari 2021). *De toekomst van tolken VGT in Vlaanderen*. Geraadpleegd van <https://bvgt.be/sites/default/files/2019-10/ambitienota%20BVGT.pdf>
- Carmien, S., & Fischer, G. (2005). *Tools for living and tools for learning*. http://scarmien.com/papers/carmien_fischer_tools.pdf
- Christiansen, A., Janssen, R., & Bayes, Chr. (2020). Screening for aberrant school performances in high-stakes assessments using influential analysis. *International Journal of Quantitative Research in Education*, 5 (2), 173-193. doi: 10.1504/IJQRE.2020.111451
- College voor Toetsen en Examens (2020, 3 november). Brochure Centrale Eindtoets 2021. Geraadpleegd van <https://www.centraleeindtoetspo.nl/publicaties/publicaties/2020/11/03/brochure-centraleeindtoets-2021>
- Crooks, T. (2011). Assessment for learning in the accountability area: New Zealand. *Studies in Educational Evaluation*, 37, 71-77.
- Departement Kanselarij en Bestuur. (2019). *Vlaamse regering 2019-2024 Regeerakkoord*. Depotnummer: D/2019/3241/159 <https://publicaties.vlaanderen.be/download-file/31741>
- Departement Onderwijs en Vorming (2017). *Meta-evaluatie M-decreet: Synthese van evaluatieve publicaties verschenen sinds de inwerkingtreding van het M-decreet in 2015*.

- <https://onderwijs.vlaanderen.be/sites/default/files/atoms/files/Synthese%20tussentijdse%20evaluatie%20M-decreet.pdf>
- Departement Onderwijs en Vorming (2020, 6 januari). *Toelatingsexamen arts & tandarts*. <https://toelatingsexamenartstandarts.be/home>
- Departement Onderwijs en Vorming (2021, 6 januari). *Onthaalonderwijs voor anderstalige kinderen (OKAN)*. <https://www.vlaanderen.be/onderwijs-en-voorming/ondersteuning/onderwijs-voor-leerlingen-met-specifieke-noden/onthaalonderwijs-voor-anderstalige-kinderen-okan>.
- Diao, Y., & Sweller, J. (2007). Redundancy in foreign language reading comprehension instruction: Concurrent written and spoken presentations. *Learning and Instruction*, 17, 78-88.
- Dirkx, K. J. H., Kester, L., & Kirschner, P. A. (2014). The testing effect for learning principles and procedures from texts. *Journal of Educational Research*, 107, 1-8.
- Duncan, H. & Purcell, C. (2020). Consensus or contradiction? A review of the current research into the impact of granting extra time in exams to students with specific learning difficulties (SpLD). *Journal of Further and Higher Education*, 44(4), 439-453. <https://doi.org/10.1080/0309877X.2019.1578341>
- Ehren, M. (2013). *Startdossier toegevoegde waarde. Onderwijsraad (OR 20130127/1046)*. <https://www.onderwijsraad.nl/binaries/onderwijsraad/documenten/rapporten/2014/04/07/studie-startdossier-toegevoegde-waarde/Studie-Startdossier-toegevoegde-waarde-site-1.7.pdf>
- Ekstrom, R. B. and Smith D. K. (2002). Assessing individuals with disabilities in educational, employment, and counseling settings. Item #: 431785A.
- Enseignement.be (2021, 12 januari 2021). Évaluations externes certificatives. Geraadpleegd van <http://www.enseignement.be/index.php?page=26245&navi=2032>.
- Foster, D. (2009). Secure, online, high-stakes testing: Science fiction or business reality? *Industrial and Organizational Psychology*, 2(1), 31.
- Frey, A., & Seitz, N. (2009). Multidimensional adaptive testing in educational and psychological measurement: Current state and future challenges. *Studies in Educational Evaluation*, 35(2-3), 89-94. <http://dx.doi.org/10.1016/j.stueduc.2009.10.007>
- González-Sancho, C., & Vincent-Lancrin, S. (2016). Transforming education by using a new generation of information systems. *Policy Futures in Education*, 14(6), 741-758. <https://doi.org/10.1177/1478210316649287>
- Goodman, D., & Hambleton, R. K. (2005). Some misconceptions about large-scale educational assessments. In R. P. Phelps (Ed.), *Defending standardized testing* (pp. 91-110). Mahwah, NJ: Lawrence Erlbaum Associates.
- Guo, J., Tay, L., & Drasgow, F. (2009). Conspiracies and test compromise: An evaluation of the resistance of test systems to small-scale cheating. *International Journal of Testing*, 9(4), 283-309.
- Harju-Luukkainen, H., McElwany, N., Stang, J. (Eds.) (2020). *Monitoring student achievement in the 21st century: European policy perspectives and assessment strategies*. Springer.
- Heymans, P. J., Godaert, E., Elen, J., van Braak, J., & Goeman, K. (2018). MICTIVO2018. *Monitor voor ICT-integratie in het Vlaamse onderwijs*. Eindrapport van O&O-opdracht: Meting ICT-integratie in het Vlaamse onderwijs(MICTIVO). KU Leuven / Universiteit Gent
- International Test Commission. (2014). *International guidelines on the security of tests, examinations, and other assessments*. Verkregen op www.intestcom.org, https://www.intestcom.org/files/guideline_test_security.pdf.
- Johnstone, C. J. (2003). *Improving validity of large-scale tests: Universal design and student performance (Technical Report 37)*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. <https://nceo.info/Resources/publications/OnlinePubs/Technical37.htm>
- Kettler, R. J. (2012) Testing Accommodations: Theory and research to inform practice. *International Journal of Disability, Development and Education*, 59(1), 53-66. <https://doi.org/10.1080/1034912X.2012.654952>
- Kettler, R. J., Elliott, S. N. & Beddow, P. A. (2009). Modifying achievement test items: A theory-guided and data-based approach for better measurement of what students with disabilities know. *Peabody Journal of Education*, 84(4), 529-551. <https://doi.org/10.1080/01619560903240996>
- Kruispuntbank Wetgeving (2020, 28 september). *Verdrag inzake de rechten van personen met een handicap* (13 december 2006). http://www.ejustice.just.fgov.be/cgi/loi/change_lg.pl?language=nl&la=N&table_name=wet&cn=2006121354
- Lai, S. A., & Berkeley, S. (2012). High-stakes test accommodations: Research and practice. *Learning Disability Quarterly*, 35(3), 158-169. <https://doi.org/10.1177/0731948711433874>

- Laitusis, C., Buzick, H., Stone, E., Hansen, E., & Hakkinen, M. (2012). *Literature review of testing accommodations and accessibility tools for students with disabilities*. Smarter Balanced Assessment Consortium. Educational Testing Service: Princeton, NJ.
- Lindstrom, J. H. (2010). Mathematics assessment accommodations: Implications of differential boost for students with learning disabilities. *Intervention in School and Clinic*, 46(1), 5–12. <https://doi.org/10.1177/1053451210369517>
- Liu, C., Han, K. T., & Li, J. (2019). Compromised Item Detection for Computerized Adaptive Testing. *Frontiers in psychology*, 10, 1-16.
- Lovett, B. J., & Nelson, J. M. (2020). Systematic review: Educational accommodations for children and adolescents with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*. <https://doi.org/10.1016/j.jaac.2020.07.891>
- Madaus, G. F., & Clarke, M. (2001). The adverse impact of high-stakes testing on minority students: Evidence from one hundred years of test data. In G. Orfield & M. L. Kornhaber (Eds.), *Raising standards or raising barriers? Inequality and high-stakes testing in public education* (pp. 85–106). New York: Century Foundation Press.
- MICTIVO (2019). *Eindrapport MICTIVO-3*. http://www.mictivo.be/wp-content/uploads/2019/01/BOOK_Eindrapport_MICTIVO-3_01.2019_OK.pdf
- Ministry of Education (2020a, 7 oktober). Assessment for learning. <https://assessment.tki.org.nz/>
- Ministry of Education (2020b, 7 oktober). Education counts. <https://www.educationcounts.govt.nz/>
- National Center for University Entrance Examinations (2017). *National Center for University Entrance Examinations. Fiscal Year 2017*. <https://www.dnc.ac.jp/albums/abm00033004.pdf>
- NCES (2021, 11 januari). *NAEP: Inclusion of Students with Disabilities and English Learners*. <https://nces.ed.gov/nationsreportcard/about/inclusion.aspx>
- Nusche, D., et al. (2012). *OECD Reviews of Evaluation and Assessment in Education: New Zealand 2011*, OECD Publishing. <http://dx.doi.org/10.1787/9789264116917-en>
- Nusche, D., et al. (2014). *OECD Reviews of Evaluation and Assessment in Education: Netherlands 2014*, OECD Publishing. <https://doi.org/10.1787/9789264211940-en>
- NZQA (2020, December 14). *National Certificates of Educational Achievement (NCEA)*. New Zealand Qualification Authority. <https://www.nzqa.govt.nz/ncea/>
- OECD (2010). *Educational Information System for Innovation and Improvement. Workshop Summary Report*. Paris: Organisation for Economic Co-operation and Development. [http://www.oecd.org/education/ceri/EDU-CERI-CD\(2015\)10.pdf](http://www.oecd.org/education/ceri/EDU-CERI-CD(2015)10.pdf)
- OECD (2013). *Synergies for better learning: an international perspective on evaluation and assessment*. Series Reviews of Evaluation and Assessment in Education. <http://dx.doi.org/10.1787/9789264190658-en>
- Ouders voor Inclusie (2020, november 24). *Nota voor Ben Weyts*. <https://www.oudersvoorinclusie.be/wp-content/uploads/2020/03/Nota-voor-minister-Ben-Weyts-def.pdf>
- Overheid.nl (2020a, november 27). *Regeling toetsreglement van de centrale eindtoets po schooljaar 2019–2020*. <https://wetten.overheid.nl/BWBR0042400/2019-08-01#Bijlage>
- Overheid.nl (2020b, november 27). *Besluit eindexamens voortgezet onderwijs 2020*. <https://wetten.overheid.nl/BWBR0043534/2021-01-01>
- Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of Learning Sciences*, 13(3), 423-451. https://doi.org/10.1207/s15327809jls1303_6
- Penninckx, M., Vanhoof, J., Quintelier, A., De Maeyer, S., & Van Petegem, P. (2017). *Zicht op leerwinst: Scenario's voor gestandaardiseerd toetsen*. Leuven: Acco.
- Pitoniak, M., & Royer, J. (2001, October 18). Testing accommodations for examinees with disabilities: A review of psychometric, legal, and social policy issues. *Review of Educational Research*, 71(1), 53-104. <http://www.jstor.org/stable/3516067>
- Plass, J. L., Moreno, R., & Brunken, R. (2010). *Cognitive load theory*. Cambridge, MA: Cambridge University Press.
- Popham, W. J. (2004). Why Assessment Illiteracy Is Professional Suicide. *Educational leadership*, 62(1), 82.
- Prodia (2020, 29 december). *Algemeen Diagnostisch Protocol*. <http://www.prodiagnostiek.be/?q=algemeen-diagnostisch-protocol>

- Quinn, D. M. (2020). Experimental evidence on teachers' racial bias in student evaluation: The role of grading scales. *Educational Evaluation and Policy Analysis*, 42(3), 375–392. <https://doi.org/10.3102/0162373720932188>
- Redden, G., & Low, R. (2012). My school, education, and cultures of rating and ranking. *Review of Education, Pedagogy, and Cultural Studies*, 34(1-2), 35-48.
- Rutkowski, L., von Davier, M., Rutkowski, D. (Eds.) (2014). *Handbook of international large-scale assessment: background, technical issues, and methods of data analysis*. CRC Press, Taylor & Francis, Chapman & Hall.
- Sahlberg, P. (2011). The professional educator: Lessons from Finland. *American Educator*, 35(2), 34-38.
- Santiago, P., Donaldson, G., Herman, J., & Shewbridge, C. (2011). *OECD Reviews of Evaluation and Assessment in Education: Australia 2011*. OECD Publishing. <https://doi.org/10.1787/9789264116672-en>.
- Scheerens, J., Brouwer, A., Sanders, P., Veldkamp, B., & van der Vegt, A.L. (2019) *Fundamentele vragen over examens en toetsing. Eindrapportage*. Oberon B.V. NRO-projectnummer: 40.5.18461.001
- Schildkamp, K., Poortman, C. L., Ebbeler, J., & Pieters, J. M. (2019). How school leaders can build effective data teams: Five building blocks for a new wave of data-informed decision making. *Journal of Educational Change*, 20(3), 283– 325. <https://doi.org/10.1007/s10833-019-09345-3>
- Shewbridge, C., et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: Denmark 2011*. OECD Publishing. <https://doi.org/10.1787/9789264116597-en>.
- Smarter Balanced Assessment Consortium (2014, October 13). Accessibility and accommodations framework. <https://portal.smarterbalanced.org/library/en/accessibility-and-accommodations-framework.pdf>
- Stocking, M. L. (1994). *Three practical issues for modern adaptive testing item pools* (ETS Research Report No. 94-5). Princeton, NJ: Educational Testing Service
- Stocking, M. L., & Lewis, C. (1998). Controlling item exposure conditional on ability in computerized adaptive testing. *Journal of Educational and Behavioral Statistics*, 23, 57-75.
- Thurlow, M. L., Lazarus, S. S., Thompson, S. J., Morse, A. B. (2005). State policies on assessment participation and accommodations for students with disabilities. *Journal of Special Education*, 38, 232–240. <https://doi.org/10.1177/00224669050380040401>
- Thurlow, M. L., Warren, S. H., & Chia, M. (2020). *Guidebook to including students with disabilities and English learners in assessments (NCEO Report 420)*. University of Minnesota, National Center on Educational Outcomes.
- Timperley, H. (2014). Promoting and supporting improvement in schools through external review. Prepared for the Education Review Office
- UNIA (2016). *Inclusief onderwijs in Vlaanderen. Een tussentijdse evaluatie*. https://www.unia.be/files/Documenten/Tussentijdse_analyse_inclusief_onderwijs-CRPD.pdf
- Van Gasse, R., Vanhoof, J., & Van Petegem, P. (2019). Brokerage for data use in schools: Potential, occurrence and facilitators. In D. Godfrey & C. Brown (Eds.), *An ecosystem for research-engaged schools: Reforming education through research* (pp. 108–122). Routledge.
- Van Gasse, R., Vanlommel, K., Vanhoof, J., & Van Petegem, P. (2017). The impact of collaboration on teachers' individual data use. *School Effectiveness and School Improvement*, 28(3), 489–504. <https://doi.org/10.1080/09243453.2017.1321555>
- Vandecandelaere, M., De Fraine, B., Verschuere, K., Colpin, H., Ghesquière, P., Flamez, E., De Jaeger, K., Gevaert, G., Thys, A., Van de Goor, M., & Wissels, S. (2020). Module 7. Verschillen tussen leerlingen in onderwijs. In J. Elen & A. Thys (Eds.), *Leren in maatschappelijk betrokken onderwijs. Basisinzichten voor leraren nu en in de toekomst (2de editie)* (pp. 261-306). Universitaire Pers Leuven.
- Vanhoof, J., De Maeyer, S., Van Petegem, P., Penninckx, M., Quintelier, A. (2015). *Scenario's voor leer(winst)monitoring in Vlaanderen: Een ontwerponderzoek naar haalbaarheid en wenselijkheid*. Eindrapport OBPWO 13.03
- Verhaeghe, G., Vanhoof, J., Van Petegem, P., Verhaeghe, J.P., Van Damme, J. (2010). *Het gebruik van outputgegevens in basisscholen: concretisering en illustraties uit het Schoolfeedbackproject*. https://www.researchgate.net/publication/263962959_Het_gebruik_van_outputgegevens_in_basisscholen_concretisering_en_illustraties_uit_het_schoolfeedbackproject
- Vlaams Ministerie van Onderwijs en Vorming (2020). *Het ondersteuningsmodel in het basis- en secundair onderwijs en in het hoger onderwijs*. <https://data-onderwijs.vlaanderen.be/edulex/document.aspx?docid=15071#2>
- Vlaams Ministerie van Onderwijs en Vorming (2020, oktober 15) <https://onderwijs.vlaanderen.be>

- Vlaamse Overheid (2020, 9 januari). *Resultaten en deliberatie bij de examencommissie secundair onderwijs*. <https://www.vlaanderen.be/een-diploma-of-getuigschrift-secundair-onderwijs-behalen-via-de-examencommissie-secundair-onderwijs/resultaten-en-deliberatie-bij-de-examencommissie-secundair-onderwijs>
- Vlaamse regering (2019). *Vlaamse regering 2019-2024. Regeerakkoord*. Brussel: Vlaamse regering. <https://www.vlaanderen.be/publicaties/regeerakkoord-van-de-vlaamse-regering-2019-2024>
- Wagemaker, H. (Ed.) (2020). *Reliability and validity of international large-scale assessment: Understanding IEA's comparative studies of student achievement*. International Association for the Evaluation of Educational Achievement (IEA)/ Springer.
- Wang, X., & Liu, Y. (2020). Detecting Compromised Items Using Information From Secure Items. *Journal of Educational and Behavioral Statistics*, 45(6), 667–689.
- Way, W. D. (1998). Protecting the integrity of computerized testing item pools. *Educational Measurement: Issues and Practice*, 17(4), 17-27.
- Welsh, M. E., Eastwood, M., & D'Agostino, J. V. (2014). Conceptualizing teaching to the test under standards-based reform. *Applied Measurement in Education*, 27(2), 98-114.
- Willem, L., Aesaert, K., Dierick, S., Van Renterghem, K., Heyvaert, M., & Janssen, R. (2019, december). *Implicaties van de nieuwe eindtermen voor de eerste graad secundair onderwijs voor toekomstige peilingen* (niet gepubliceerd rapport). Leuven: Steunpunt voor Toetsontwikkeling en Peilingen (STEP).
- Wollack, J. A., & Fremer, J. J. (2013). Introduction: The test security threat. In A. Wollack & J. J. Fremer (Eds.), *Handbook of test security* (pp. 1–13). New York, NY: Routledge
- Wollack, J. A., & Fremer, J. J. (2013). Introduction: The test security threat. In A. Wollack & J. J. Fremer (Eds.), *Handbook of test security* (pp. 1–13). New York, NY: Routledge
- Woods, K., McCaldin, T., Hipkiss, A., Tyrrell, B., & Dawes, R. (2019). Linking rights, needs, and fairness in high-stakes assessments and examinations, *Oxford Review of Education*, 45(1), 86-101. <https://doi.org/10.1080/03054985.2018.1494555>
- Ylioppilastutkintolautakunta (2020, 15 oktober). *Matriculation Examination*. <https://www.ylioppilastutkinto.fi/>
- Zhang, J. (2014). A sequential procedure for detecting compromised items in the item pool of a CAT system. *Applied Psychological Measurement*, 38(2), 87-104.
- Zhang, J., & Li, J. (2016). Monitoring items in real time to enhance CAT security. *Journal of Educational Measurement*, 53(2), 131-151.
- Zhang, J., Chang, H. H., & Yi, Q. (2012). Comparing single-pool and multiple-pool designs regarding test security in computerized testing. *Behavior research methods*, 44(3), 742-752.

7. BIJLAGE 1: ONDERZOEKSVRAGEN PERCEEL 2 VOLGENS DE OPROEP VOOR DE HAALBAARHEIDSTUDIE

Toetsverversing

17) Op welke wijze beïnvloeden de aspecten van praktische afname de nood aan toetsverversing? Denk hierbij zeker aan de confidentialiteit van de items.

Brede afname

18) Hoe kan er voor gezorgd worden dat alle leerlingen de toets kunnen afleggen en tegelijkertijd de standaardisatie van de afname bewaakt wordt? Ingeval de toets niet voor alle leerlingen haalbaar is, hoe gaan we dan om met het verlenen van vrijstellingen?

Denk hierbij onder meer aan:

- Aanpassingen voor leerlingen met functie-uitval (bijvoorbeeld blinde of dove leerlingen)
- Strategie met betrekking tot redelijke aanpassingen voor leerlingen met of zonder attest:
 - o Wie bepaalt wie redelijke aanpassingen krijgt?
 - o Hoe wordt bepaald wie redelijke aanpassingen krijgt?
 - o Welke redelijke aanpassingen zijn mogelijk en wenselijk?
- Strategie met betrekking tot het toetsen van leerlingen die huisonderwijs krijgen.
- Strategie met betrekking tot het toetsen van leerlingen die gewettigd of ongewettigd afwezig zijn op het geplande afnamemoment.
- Strategie met betrekking tot het verlenen van vrijstelling van de toets.

Verantwoordelijke/uitvoerder toetsafname

19) Welke actoren dienen betrokken te worden bij de afname van de toetsen? Welke rol dienen zij daarin op te nemen? Wat kan de mogelijke relatie zijn tot de instantie die verantwoordelijk is voor de ontwikkeling van de toets?

20) Indien de scholen de verantwoordelijkheid krijgen om de toetsen af te nemen, wat betekent dit dan naar het gebruik van de toetsresultaten (op schoolniveau, op individueel leerlingniveau), mogelijke toetsfraude en handavingsbeleid toe? Welke graad van autonomie inzake de toetsafname kan worden gegeven aan de scholen en wat zijn hier de consequenties van?

Afnametijd

21) Wat zijn de voor- en nadelen en de randvoorwaarden om een toets Vlaanderenbreed op één dag af te nemen of gespreid over een bepaalde termijn? Welke andere mogelijkheden zijn er?

Afnamemoment in het schooljaar

22) Wat zijn mogelijke momenten/periodes om de toetsen af te nemen binnen een schooljaar?

Hou daarbij zeker rekening met:

- de impact op de praktische organisatie van het schooljaar;
- de effecten op de toetsresultaten van het samengaan met andere evaluatiepraktijken;
- de tijdige beschikbaarheid van de resultaten (bijvoorbeeld voor de klassenraden in functie van studie-oriëntering) gelet op de doorlooptijd tussen het afnemen van de toets en het bekomen van de resultaten.

Doorlooptijd

23) Op welke manier kan de doorlooptijd tussen het afnemen van de toets en het bekomen van de resultaten in tijd beperkt worden? Wat zijn hiervoor de randvoorwaarden? Wat is een realistische doorlooptijd?

8. BIJLAGE 2: VERSLAG FOCUSGESPREK STEP

VERSLAG



Focusgesprek met partners uit het buitengewoon onderwijs:

Samen nadenken over paralleltoetsen als toetsinstrument binnen interne kwaliteitszorg in het buitengewoon onderwijs.

DATUM: dinsdag 11 februari 2020 van 13u30 tot 15.30u

LOCATIE: Campus Clenardus UCLL, Weerstandsplein 2, Diest

AANWEZIG: Atzeni Paola (BSBO Mikado), Claes Veerle (BLO De Olm), Creemers Valerie (MPI-Helix), De Keyzer Dominik (Katholieke PBD), Dreesen Karen (MPI Helix), Goris Steven (MPI Helix), Huysmans Petra (POV), Loomans Elke (MPI Helix), Luc (POV), Segers Laura (Sint-Gerardus), Snijders Daisy (BSBO Mikado), Sys Els (OVSG), Thys Kristof (MPI Helix), Tyskens Liesbet (WICO campus TIO/BBO), Vanmarsenille Carolien (Sint-Gerardus), Wastiels Tine (Sint-Gerardus), Dierick Sabine (STEP), Laenen Isabel (STEP) en Willem Lien (STEP).

VERONTSCHULDIGD: Degline Daniëlle (De Bloesem), Lambrechts Dominique (Ten Desselaer), Nagels Francis (BSBOGO De Zonnestraal), Peeters Sabine (STEBO Dilsen), Verhoeven Bieke (VIBO De Brem, BKLO), Werbroeck Soetkin (POV).

Agenda

1. **Presentatie** als inleiding en kadering voor gesprek:
 - Doel van peilings- en paralleltoetsen
 - Paralleltoetsen in de context van het buitengewoon onderwijs (BuO)
 - Schoolfeedback bij paralleltoetsen

2. Focusgesprek

Bespreking

1. Vragen en reacties naar aanleiding van de presentatie:

- De vraag wordt gesteld of alle eindtermen getoetst worden bij een peiling. Reactie van STEP: Ja, tenzij ze niet schriftelijk, gestandaardiseerd te toetsen zijn. Er gebeurt wel telkens een selectie van ET op voorhand.
- In het BuO wordt er gewerkt in erg heterogene klasgroepen op verschillende niveaus (eerder individueel). Een paralleltoets afnemen binnen één klasgroep lijkt niet haalbaar.
- De paralleltoets dient afgenomen te worden op het **einde van het 6e leerjaar**, maar weinig leerlingen in het BuO bereiken dan de eindtermen. Het ritme van een schooljaar wordt niet gevolgd in het BuO.
- De **afnamecondities** kunnen ook sterk verschillen tussen leerlingen in het BuO.- Voor deze leerlingen is het nodig om ook te **differentiëren in de instructie** (instructieboekje).

- De vraag wordt gesteld naar de vergelijking van scholen binnen het BuO zelf. Reactie van STEP: Op dit moment is **vergelijking tussen scholen in het BuO** niet mogelijk. Een vergelijking maken tussen scholen vraagt gestandaardiseerd toetsen op een grootschalige manier. Dit is moeilijk in het BuO omdat de afnamecondities overal anders zijn en de groep scholen BuO vrij klein is. In het gewoon onderwijs zijn er 1500 leerlingen per toets nodig om hierover een betrouwbare uitspraak te kunnen doen.
- De vraag wordt gesteld of **vergelijken** überhaupt zinvol is binnen de context van het BuO. In het gewone onderwijs zijn de noden ook erg verschillend, maar de standaard die iedereen moet bereiken is dezelfde. Hoe kan je in het BuO kinderen op één lijn testen? In het BuO hebben kinderen allemaal verschillende (andere) behoeften. **De lat is niet dezelfde voor elk kind**. Ook een vergelijking binnen het BuO zou weinig zinvol zijn door de grote verschillen tussen de scholen.
- De vraag wordt gesteld voor **welke doelgroep van leerlingen** de paralleltoets wel een **meerwaarde** kan zijn. Reactie van STEP: De paralleltoets kan afgenomen worden bij die leerlingen waarbij de eindtermen of ontwikkelingsdoelen wel worden nagestreefd.
- Een leerkracht geeft aan dat de insteek bij afname kan zijn: *Hoe zou die leerling het in het gewoon onderwijs doen?*
- STEP: Een paralleltoets kan **niet gebruikt** worden voor **oriëntering** voor leerlingen. De toets is hier niet voor bedoeld. Feedback wordt pas opgeleverd in oktober. De resultaten van een paralleltoets kunnen **wel** gebruikt worden als **evaluatie van de eigen werking op leerling-, klas- en schoolniveau**.
- PBD OMSG: Paralleltoetsen kunnen interessant zijn omdat ze een breed beeld geven, bijvoorbeeld voor wiskunde basisonderwijs. Ook als je ze jaar na jaar afneemt kan het wel een beeld geven. Scholen hebben nood aan voldoende tools om er op het niveau van kwaliteitszorg mee aan de slag te gaan.
- PBD POV: Het is belangrijk voor een school om een focus te kiezen. Je moet niet alle paralleltoetsen afnemen. Binnen POV bijvoorbeeld gaan we focussen op PAV.

2. Focusgesprek

Rondvraag: Waarom zijn jullie hier? Welke informatiebehoefte is er, en in hoeverre kunnen paralleltoetsen daar volgens jullie een antwoord op bieden?

Kadering vanuit de PBD van KOV: Wat is de meerwaarde van de paralleltoetsen? Waarom zou je ze afnemen? Er zijn 2 niveaus:

- uitspraken doen over leerlingen op **individueel niveau**: In hoeverre bereikt deze leerling de eindtermen?
- als toetsinstrument in het kader van **kwaliteitszorg** in de school (evaluatie van de eigen werking)
- Leerkracht BuBaO: staat in een autiklas van leerlingen niveau 4de, 5de en 6de leerjaar met een normale begaafdheid. De toetsen van de methode en het leerlingvolgsysteem worden gebruikt als toetsinstrument. Puur praktisch zou afname van paralleltoetsen erg moeilijk zijn (één instrument voor alle leerlingen) omdat de **diversiteit in de groep zo groot** is (verschillende niveaugroepen). Er is interesse voor paralleltoetsen om na te gaan of de leerlingen die **doorstromen** de ET bereikt hebben (oriëntering), maar de **feedback** komt daarvoor **te laat** (de kinderen zijn al vertrokken).
- Leerkracht BuBaO: staat in een leerklas. Dit jaar zijn er twee leerlingen die doorstromen naar het gewone onderwijs. Er is interesse voor afname paralleltoetsen om te weten of **verwachtingen nog juist gesteld** worden in functie van doorstroming. *“We passen ons zoveel aan, aan de leerlingen. We raken soms de voeling kwijt met wat verwacht wordt in het gewone onderwijs.”*
- Directie van een school BuO: “We vergelijken ons nooit met andere scholen – je weet niet hoe we nu als school bezig zijn. Hiervoor kunnen paralleltoetsen interessant zijn, in functie van **interne kwaliteitszorg (IKZ)**” Het gaat echter maar over een klein stukje van de puzzel: enkel het cognitieve stukje, en eigenlijk zijn de prestaties van leerlingen in het BuO ondergeschikt. Bij buitengewoon onderwijs zijn er heel veel factoren die meespelen – niet (alleen) het behalen van ET (in OV4 bijvoorbeeld gedrag). Kwaliteitsindicatoren voor een BuO-school zijn bijvoorbeeld: de begeleiding, het welbevinden van de leerlingen en de mate waarin leerlingen aangezet worden tot zelfredzaamheid.

- Directie van een BuO-school: *“Gestandaardiseerde toetsen zoals AVI nemen wij niet af, maar inspectie vraagt wel wat we doen met de **output**.”* Er wordt opgeworpen dat er **weinig instroomdata** zijn in het BuO om aan kwaliteitsbeleid te doen, zeker in vergelijking met het gewoon onderwijs, waar dataloop erg veel gegevens bevat.
- In het gewone onderwijs worden er variabelen ingevoerd om scholen met een vergelijkbaar doelpubliek te kunnen vergelijken. In het BuO wordt er überhaupt weinig vergeleken met andere scholen. Er zijn zoveel variabelen die meespelen dat zo’n **vergelijking weinig zinvol** zou zijn voor interne kwaliteitszorg.
- Paralleltoetsen geven feedback ten opzichte van een standaard - in het BuO is er geen standaard. Eén instrument dat geschikt is voor een hele klas kan al bijna niet in het BuO. De **verwachtingen zijn niet standaard**.
- Het **aantal leerlingen** dat in aanmerking komt voor afname van de paralleltoets in het BuO is **te beperkt**, waardoor het gebruik van de paralleltoets(en) voor de evaluatie van de eigen werking op klas- en schoolniveau weinig zinvol is. De focus op schoolbeleidsvlak ligt bovendien meestal elders in het BuO (*“Er zijn zoveel dingen om te evalueren – is dat het waard?”*).
- De vraag wordt gesteld of het niet zinvoller is in het BuO om een gestandaardiseerde test af te nemen op een bepaald niveau. Er is **vraag naar/nood aan een gestandaardiseerde test** om te kunnen beslissen of een kind klaar is **voor de A-stroom of B-stroom**. Reactie vanuit de PBD van het KOV: Wettelijk gezien is er geen niveaubepaling voor toegang tot de B-stroom. Het niveau van een B-klas is ook sterk verschillend, afhankelijk van de regio en zelfs van school tot school.
- **Opleidingsvorm 4** in het BuSO is wel gericht op het bereiken van de eindtermen. Vanuit bevraging door POV blijkt dat de **huidige paralleltoetsen** dan **qua vormgeving en vraagstelling** wel **aangepast** zouden moeten worden: eenvoudige vragen, beperkt in tijd, geen figuurlijk of meervoudig taalgebruik. Reactie vanuit STEP: Indien de paralleltoetsen worden afgestemd op het doelpubliek van BuO zou dit binnen een apart project moeten gebeuren, want paralleltoetsen werden ontwikkeld voor het gewoon onderwijs.
- Er wordt opgeworpen dat BuO een te kleine afzetmarkt is. Is er een andere instrument dan paralleltoetsen? Wij moeten alles zelf ontwikkelen...
- De vraag ‘zijn we goed bezig’ houdt veel scholen bezig. De pedagogisch begeleider van KOV werpt op dat vergelijking niet per se nodig is om op die vraag te antwoorden en geeft mee dat veel leerkrachten en scholen worstelen met de zware verantwoordelijkheid die bij **oriëntering** komt kijken. Je bepaalt de toekomst van die leerling in grote mate. Leerkrachten zoeken naar middelen die hun mening staven om die verantwoordelijkheid kleiner te maken, liefst met harde cijfers. Hiervoor is individuele feedback vanuit paralleltoetsen niet geschikt en het is er niet voor bedoeld. De vraag wordt gesteld of er een ander instrument moet komen. De IDP’s hebben het geprobeerd en weer stopgezet. Volgens POV mogen we het niet doen. **Interdisciplinair overleg is de toetssteen**.

Wat wordt meegenomen uit dit gesprek?

- Leerkrachten willen hun oordeel gestaafd zien, liefst met harde cijfers, om niet zelf de hele verantwoordelijkheid te moeten dragen. Door het gesprek is helder geworden dat de paralleltoets niet het geschikte toetsinstrument is voor oriëntering van leerlingen. Nu kunnen we dit loslaten.
- Bij gebruik van de paralleltoets als (toets)instrument binnen OV 4 BuSO is de vormgeving en vraagvorm niet steeds geschikt voor dat doelpubliek. Ook dan dient de paralleltoets enkel ingezet te worden op schoolniveau en niet als oriëntering.
- Er blijft wel vraag naar een instrument om te weten of je als BuO-school goed bezig bent.
- Anderzijds kan je niet zeggen of je goed bezig bent door te vergelijken met anderen (vraagt mindshift).

Verdere opvolging:

- Dit verslag wordt doorgestuurd naar het Departement Onderwijs en Vorming (Afdeling Strategische beleidsondersteuning). Zij beslissen mee over verdere stappen die gezet worden.
- Eventuele verdere stappen: verkennen of (en hoe) de paralleltoets een geschikt instrument kan zijn binnen OV4 BuSO. Dit willen we eerst verder afstemmen met de PBD.