STEUNPUNT
CENTRALE TOETSEN
IN ONDERWIJS

FEEDBACKGEBRUIK UIT CENTRALE TOETSEN

Survey-onderzoek bij schoolleiders en leraren

januari 2023

Glen MOLENBERGHS
Roos VAN GASSE
Sven DE MAEYER
Jan VANHOOF

Rapport Steunpunt Centrale Toetsen in Onderwijs
Werkdomein H: Schoolfeedback
Promotor: Jan Vanhoof

Vlaanderen
is onderwijs & vorming

© 2023 Steunpunt Centrale Toetsen in Onderwijs
Henri Dunantlaan 2
9000 Gent

Referentienummer: SCTO/2023.H/1/5

Gelieve als volgt naar deze publicatie te verwijzen:

Molenberghs, G., Van Gasse, R., De Maeyer, S., Vanhoof, J. (2023). *Feedbackgebruik uit centrale toetsen. Survey-onderzoek bij schoolleiders en leraren.* Steunpunt Centrale Toetsen in Onderwijs.

Deze publicatie kwam tot stand met steun van de Vlaamse Gemeenschap, Ministerie voor Onderwijs en Vorming. Deze publicatie is ook beschikbaar via www.steunpunttoetsen.be

INHOUD

1 Situering 6

2 (Waarom) Gebruikersonderzoek 9

3 Theoretische uitgangspunten 11

3.1 Centrale toets 12

3.2 Feedback 13

3.2.1 De relevantie van feedback 13
3.2.2 De betekenisvolheid van feedback 14
3.2.3 De perceptie van validiteit en betrouwbaarheid van feedback 15

3.3 Feedbackgebruik 16

3.3.1 Een systematisch en cyclisch proces 16
3.3.2 Een doelgericht proces 17

3.4 Stimulerende en hinderende factoren voor feedbackgebruik 21

3.4.1 Schoolfactoren 21
3.4.2 Contextuele factoren 22
3.4.3 Gebruikersfactoren 23

3.5 Effecten van feedbackgebruik 28

4 Methodologie 30

4.1 Instrument(ontwikkeling) 30

4.2 Steekproeftrekking en Dataverzameling 32

4.3 Data-analyse 35

4.3.1 Beschrijvende statistiek 36
4.3.2 Verschillen naar functie, onderwijsniveau en onderwijsnet 38

4.4 Kenmerken van respondenten en hun scholen 39

4.4.1 Steekproef lager onderwijs 39
4.4.2 Steekproef secundair onderwijs 43

5 Resultaten 48

5.1 Perceptie van relevantie en betekenisvolheid van feedback uit centrale toetsen en van faire en valide vergelijkingen 48

5.1.1 Perceptie van relevantie en betekenisvolheid van feedback uit centrale toetsen 49
5.1.2 Perceptie van faire en valide vergelijkingen 53

5.2 Doelen voor feedbackgebruik uit centrale toetsen 57

5.2.1 Doelen voor feedbackgebruik op leerlingniveau 57
5.2.2 Doelen voor feedbackgebruik op leraarniveau 59
5.2.3 Doelen voor feedbackgebruik op schoolniveau 60

5.3 Randvoorwaarden voor feedbackgebruik uit centrale toetsen 63

5.3.1 Toegang tot de feedback 63
5.3.2 Schoolfactoren met betrekking tot feedbackgebruik 71
5.3.3 Schoolexterne verwachtingen met betrekking tot feedbackgebruik 80
5.3.4 Gebruikersfactoren met betrekking tot feedbackgebruik 81

5.4 Verwachte onwenselijke effecten van centrale toetsen 94

5.4.1 Verwachte onwenselijke effecten van centrale toetsen op leerlingniveau 94
5.4.2 Verwachte onwenselijke effecten van centrale toetsen op leraarniveau 95
5.4.3 Verwachte onwenselijke effecten van centrale toetsen op schoolniveau 97
5.4.4 Verwachte onwenselijke effecten van centrale toetsen op onderwijssysteemniveau 99

5.5 Specifieke actuele thema's 100

6 Conclusie en aandachtspunten voor beleid en Steunpunt 106

6.1 Toetsresultaten op leerling-, klas- en schoolniveau maken de feedback relevant 106

6.2 Het aftoetsen van toetsresultaten aan eindtermen, het rapporteren van leerwinst en 'faire (school)vergelijkingen' maken de feedback betekenisvol 107

6.3 Vooral leraren verwachten feedbackgebruik op leerlingniveau en vooral schoolleiders verwachten feedbackgebruik op schoolniveau 109

6.4 Toegang voor specifieke gebruikersgroepen tot gerichte feedback 110

6.5 (Slechts) een beperkt actueel (intensief) gebruik van (de resultaten van) gestandaardiseerde toetsen 112

6.6 Schoolinterne verwachtingen inzake feedbackgebruik zijn eerder afwezig 113

6.7 Beleid inzake feedbackgebruik in scholen wordt eerder niet gevoerd 113

6.8 Kenmerken van transformationeel leiderschap inzake feedbackgebruik blijken slechts in beperkte mate in scholen aanwezig 114

6.9 Schoolexterne verwachtingen inzake feedbackgebruik zijn eerder afwezig 115

6.10 Schoolleiders en leraren staan globaal neutraal tegenover (feedbackgebruik uit) centrale toetsen 115

6.11 Schoolleiders en leraren tonen zich (in zekere mate) welwillendheid inzake feedbackgebruik 116

6.12 Schoolleiders en leraren voelen zich (eerder wel) bekwaam inzake feedbackgebruik 117

6.13 Schoolleiders en leraren zijn bekommerd over het systeem van centrale toetsen en vragen naar informatie over feedbackgebruik 117

6.14 Ook verwachte onwenselijke effecten op leraar-, school- en systeemniveau 118

6.15 Afrondend: belang van informatie, ondersteuning en professionalisering 120

Bibliografie 121

1 SITUERING

Scholen dienen – autonoom – hun onderwijskwaliteit te bewaken en te verbeteren. Hierbij wordt van scholen verwacht in hun kwaliteitsbeleid meer en meer expliciet – ook – aandacht te besteden aan hun leer- en onderwijsprestaties (Verhaeghe, Vanhoof, Van Petegem, Verhaeghe, & Van Damme, 2010). Tevens wordt van scholen in toenemende mate verwacht zich op basis van objectieve parameters te verantwoorden voor de eigen onderwijskwaliteit (Marsh, Pane, & Hamilton, 2006; Schildkamp, Ehren, & Lai, 2012). Hierbij is er de tendens naar het meer gericht aanspreken van onderwijsaanbieders (leraren, scholen en intermediairen) op de wijze waarop ze de autonomie die ze genieten succesvol weten aan te wenden en hoe deze autonomie zich vertaalt naar leerprestaties en leerwinst van leerlingen (Hamilton, Stecher, & Klein, 2002).

Hoewel er in de (inter)nationale context (nog) aanhoudende discussies woeden over de implementatie van systemen die leerprestaties en leerwinst van leerlingen gestandaardiseerd in kaart brengen, is het aantal landen waar deze systemen – de voorbije decennia – gemeengoed is geworden fors toegenomen (OECD, 2013). Voorstanders van gestandaardiseerde toetsen zien immers niet alleen een meerwaarde in het objectiveren van de leerprestaties van leerlingen (Marlow et al., 2014) en in het verminderen van sociale ongelijkheid door outputstandaardisatie door het hanteren van dezelfde minimumvereisten voor alle leerlingen (Bishop & Wößmann, 2004; van Ackeren, Block, Klein, & Kühn, 2012), ze zien ook beloftevolle mogelijkheden in dergelijke systemen inzake het versterken van zowel interne (Marsh, Pane, & Hamilton, 2005) als externe kwaliteitszorg (Parkes & Stevens, 2003; Verger & Parcerisa, 2018b). Tegenstanders van gestandaardiseerde toetsen halen veelal *teaching (to) the test* (Au, 2007; Janssens et al., 2014), het ontstaan van *rankings* (Slone et al., 2013) en de daaraan gekoppelde '*naming-and-blaming-cultuur*' (Elstad, 2009) aan als onwenselijke effecten.

Ook in Vlaanderen kan men niet om de vaststelling heen dat steeds meer aandacht wordt geschonken aan objectieve resultaten van onderwijsprocessen op micro-, meso- en macroniveau. Het is vanuit dit oogpunt dat scholen sinds het schooljaar 2018-2019 verplicht zijn gevalideerde toetsen af te nemen bij alle leerlingen

op het einde van de basisschool en dit voor tenminste drie leergebieden[1]. Typisch voor het Vlaamse onderwijs zijn daarnaast de verscheidene signalen uit wetenschappelijke monitoring die aangeven dat de kwaliteit van het Vlaamse onderwijs tanende is. Zo tonen de resultaten van recente peilingsonderzoeken en internationaal vergelijkende studies (o.a. PIRLS, PISA) aan dat de leerprestaties van de Vlaamse leerlingen van het lager en secundair onderwijs zich in een dalende lijn bevinden en dit zowel voor Nederlands (lezen) als voor wiskunde (De Meyer et al., 2020; Dockx, Van Landeghem, Aesaert, Van Damme, & De Fraine, 2019; Faddar et al., 2020). Daarenboven blijkt uit de meest recente vaststellingen van de onderwijsinspectie dat voor de meeste scholen (zowel in het basisonderwijs als het secundair onderwijs) evaluatie een uitdaging is, waarbij scholen kansen onbenut laten tot een transparante en betrouwbare evaluatie die is afgestemd op het gevalideerde doelenkader[2].

Vanuit de bezorgdheid over "dalende tendensen van een aantal indicatoren die iets zeggen over de kwaliteit van ons onderwijs" voorziet de Vlaamse Overheid daarom de implementatie van net- en koepeloverschrijdende gevalideerde, gestandaardiseerde en genormeerde proeven (hierna 'centrale toetsen'), waarbij het meten van leerwinst beoogd wordt[3]. Er werd beslist te starten met de ontwikkeling van – op eindtermen gebaseerde – digitale centrale toetsen voor Nederlands en wiskunde, terwijl andere leergebieden in de toekomst nog overwogen zullen worden. De toetsen zullen worden afgenomen in 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 van het secundair onderwijs. Feedback wordt wellicht teruggekoppeld op zowel leerling-, klas-, als schoolniveau.

Feedback uit centrale toetsen kan belangrijke informatie voor scholen opleveren. Door feedback vanuit vooropgestelde doelen systematisch te interpreteren, te analyseren en verbeterpunten voorop te stellen en

---

[1] Omzendbrief BaO/2017/02 betreffende 'Decreet basisonderwijs van 25 februari 1997, artikel 44ter, zoals ingevoegd bij Onderwijsdecreet XXVII'. Online beschikbaar via https://data-onderwijs.vlaanderen.be/edulex/document.aspx?docid=15086

[2] Onderwijsspiegel 2021. Online beschikbaar via https://www.onderwijsinspectie.be/sites/default/files/atoms/files/OS2021-web.pdf

[3] Beleidsnota Onderwijs 2019-2024. Online beschikbaar via https://www.vlaanderen.be/publicaties/beleidsnota-2019-2024-onderwijs

aan te pakken, worden kernprocessen binnen de school in kaart gebracht en geoptimaliseerd (Barrezeele, 2012). Op deze manier kan het gebruik van feedback dan ook gezien worden als een belangrijke stap om zicht te krijgen op zowel het functioneren en presteren van de school en de leraren in het algemeen, als van de leerlingen binnen de school in het bijzonder. Vervolgens kan de feedback mee aan de basis kan liggen voor het nemen van geïnformeerde beslissingen (Levin & Datnow, 2012). Vlaams onderzoek toont echter aan dat Vlaamse scholen – op microniveau – slechts gering gebruikmaken van outputinformatie met betrekking tot externe monitoring en voor het vormgeven van beleid en praktijk slechts beperkt steunen op benchmarks als de feedback niet tegemoetkomt aan de noden van de gebruikers (Van Gasse, Vanhoof, Mahieu, & Van Petegem, 2015). Voor de implementatie van (het gebruik van feedback uit) centrale toetsen is het daarom van belang zowel inzicht te verwerven in hoe de feedback in Vlaamse scholen en klassen zo optimaal mogelijk in kwaliteitszorgactiviteiten kan worden binnengebracht (Schildkamp & Poortman, 2015; Van Gasse et al., 2015), als zicht te krijgen op de informatiebehoeften (Van Petegem & Vanhoof, 2007; Vanhoof, Verhaeghe, & Van Petegem, 2009) en verwachtingen van stakeholders (Datnow, Park, & Wohlstetter, 2007) met betrekking tot het gebruik van centrale toetsen. Op deze manier kunnen immers randvoorwaarden voor een succesvolle implementatie en gebruik van centrale toetsen in het Vlaamse onderwijs in kaart worden gebracht (Campbell & Levin, 2009).

2 (WAAROM) GEBRUIKERSONDERZOEK

Op basis van een brede literatuurstudie blijkt dat feedbackgebruik beïnvloed kan worden door talrijke randvoorwaarden. Zowel feedback-, als feedbacksysteem-, school-, context- en gebruikersfactoren kunnen het gebruik van feedback uit centrale toetsen bevorderen of belemmeren. Op basis van deze factoren werden indicatoren en een instrumentarium ontwikkeld voor een kwantitatief gebruikersonderzoek.

Om een effectief gebruik van feedback uit centrale toetsen in het Vlaams onderwijs doordacht vorm te geven, wil dit gebruikersonderzoek inzicht verwerven in hoe toekomstige schoolleiders en leraren deze randvoorwaarden invullen. Hiertoe formuleerden we vier onderzoeksvragen:

1. Wat maakt feedback uit centrale toetsen relevant, (gepercipieerd) betrouwbaar en valide en betekenisvol voor Vlaamse schoolleiders en leraren?
2. Vanuit welke doelen verwachten Vlaamse schoolleiders en leraren met feedback uit centrale toetsen aan de slag te gaan?
3. Welke feedbacksysteem-, school-, context- en gebruikersfactoren beïnvloeden Vlaamse schoolleiders en leraren in hun feedbackgebruik uit centrale toetsen?
4. Welke onwenselijke effecten verwachten Vlaamse schoolleiders en leraren van feedbackgebruik uit centrale toetsen?

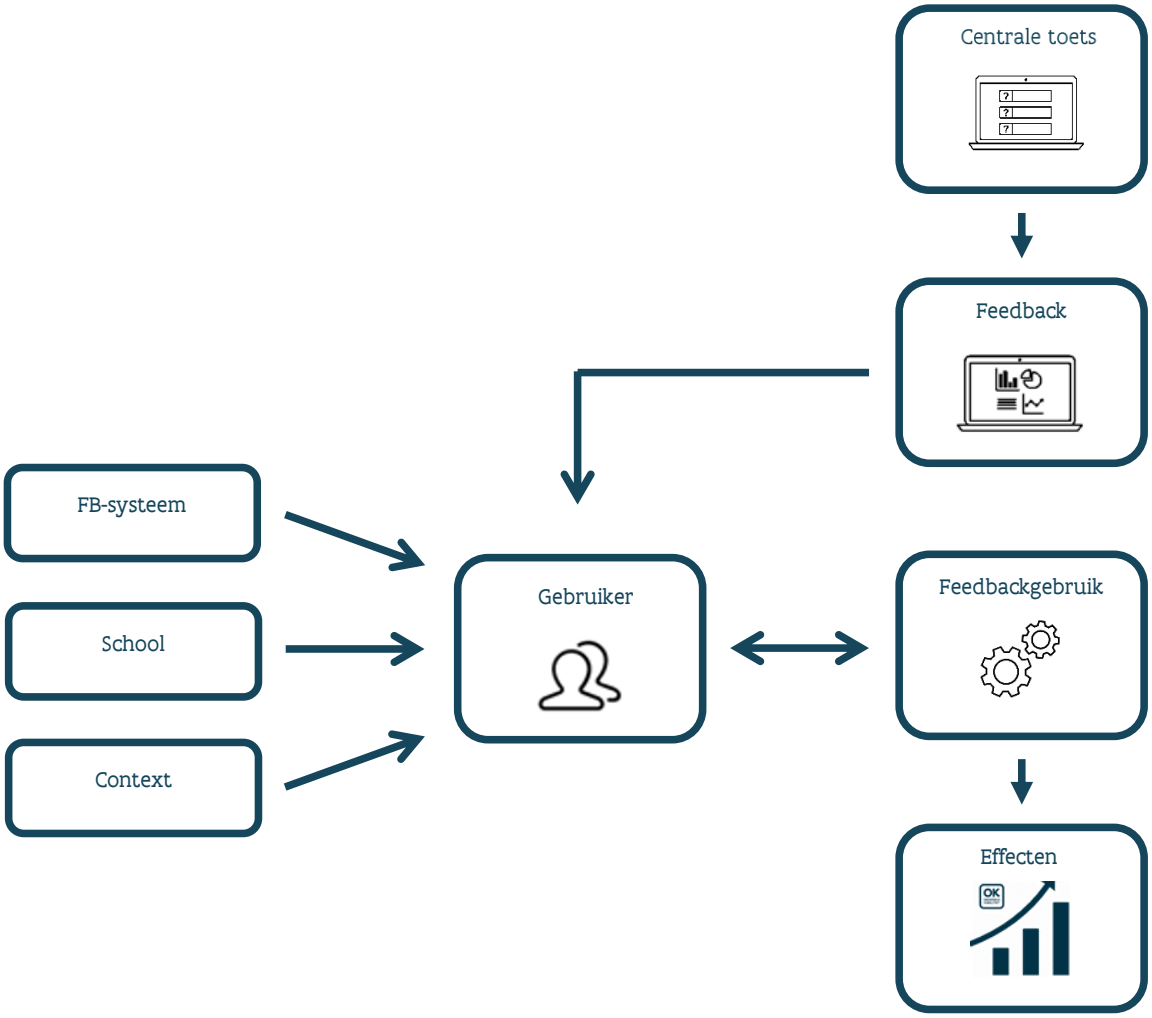
Een zicht op de (gepercipieerde) verwachtingen en randvoorwaarden van schoolleiders en leraren ten aanzien van het gebruik van feedback uit centrale toetsen zal de verschillende samenwerkingspartners (zoals overheid, Steunpunt Centrale Toetsen in Onderwijs, pedagogische begeleidingsdiensten) in staat stellen kort op de bal te spelen, voorliggende scenario's te herzien en de nodige optimalisaties door te voeren. Tevens willen we met dit gebruikersonderzoek ook verder onderzoek voeden en bijdragen aan wat nationaal en internationaal geweten is over feedbackgebruik in het onderwijs.

Aangezien dit onderzoek gevoerd werd in een periode waarin informatie over het opzet van het Vlaams systeem van centrale toetsen eerder onbestaande was, is het van belang deze gepercipieerde verwachtingen en randvoorwaarden binnen deze context te interpreteren en hierbij de nodige voorzichtigheid aan de dag te leggen.

In wat volgt, brengen we verslag uit over surveyonderzoek gericht op het beantwoorden van de voorliggende onderzoeksvragen. We starten met het weergeven van een conceptueel figuur als gids door de theoretische uitgangspunten (paragraaf 3). Deze beschrijven een set aan indicatoren met betrekking tot het gebruik van feedback uit centrale toetsen. In paragraaf 4 gaan we dieper in op de gehanteerde methodologie. De resultaten van het kwantitatief onderzoek worden in paragraaf 5 rond de vier onderzoeksvragen uitgewerkt en in de concluderende samenvatting (paragraaf 6) formuleren we aandachtspunten voor beleid en Steunpunt Centrale Toetsen in Onderwijs (verderop: Steunpunt).

3 THEORETISCHE UITGANGSPUNTEN

Om inzicht te krijgen in reeds bestaand onderzoek naar het gebruik van feedback uit centrale toetsen en randvoorwaarden bij feedbackgebruik werd een brede literatuurstudie uitgevoerd. Op basis van deze brede literatuurstudie ontwikkelden we onderstaande conceptuele figuur (Figuur 1), die de rode draad vormt voor de beschrijving van de conceptuele uitgangspunten.



Figuur 1
Conceptuele figuur van feedbackgebruik uit centrale toetsen

Leerlingen leggen een centrale toets af (paragraaf 3.1). De toetsresultaten op deze centrale toetsen kunnen relevante feedback opleveren voor scholen. In paragraaf 3.2 besteden we daarom aandacht aan essentiële concepten en aspecten van feedback. Feedbackgebruik (paragraaf 3.3) kan gezien worden als een waardevolle stap om een zicht te krijgen op zowel het functioneren van de school en de leraren in het algemeen, als van de leerlingen binnen de school in het bijzonder. We focussen ons op doelen die hieruit kunnen voortvloeien. Verschillen in feedbackgebruik worden aan een variatie van beïnvloedende factoren toegeschreven. In paragraaf 3.4 beschrijven we feedbacksysteem-, school-, context- en gebruikersfactoren als randvoorwaarden voor een optimaal feedbackgebruik. Tot slot staan we in paragraaf 3.5 stil bij mogelijke effecten van het gebruik van feedback uit centrale toetsen.

3.1 CENTRALE TOETS

Een gestandaardiseerde toets is een toets waarbij '*the test content is equivalent across administrations and the conditions under which the test is administered are the same for all the test takers*' (Sireci, 2005, p. 35). Deze omschrijving impliceert dat zowel de toetsinhoud, als de afnamecondities overheen leerlingen en scholen bij elke toetsafname identiek zijn en kent geen vooropgezette bepalingen met betrekking tot verplichting of vrijwilligheid van afname, het doel van de toets, mogelijke gevolgen, frequentie van afname, getoetste leerdomeinen, leeftijd van leerlingen of toetsvorm.

Met het oog op consistentie en vergelijkbaarheid worden gestandaardiseerde toetsen veelal extern – centraal – ontworpen en beheerd (OECD, 2013; Popham, 1999). In dit kader spreekt men dan ook over 'centrale' toetsen. Centrale toetsen baseren zich – doorgaans – op een betreffend leerplan en spiegelen de leerprestaties van leerlingen aan een externe standaard (Woessmann, 2002).

De Vlaamse centrale toetsen zijn gestandaardiseerde, genormeerde en gevalideerde net- en koepeloverschrijdende toetsen, met als doel de interne kwaliteitszorg van scholen te ondersteunen en de onderwijskwaliteit te versterken. De toetsen focussen in eerste instantie op een selectie van de eindtermen van Nederlands (begrijpend lezen, schrijven, grammatica) en wiskunde. Het is de bedoeling om zoveel mogelijk leerlingen uit het lager en secundair onderwijs te laten deelnemen aan de centrale toetsen. In principe legt elke leerling in het leerplichtonderwijs vier keer een centrale toets af: aan het einde van het vierde en het zesde leerjaar van het lager onderwijs en aan het einde van de eerste en derde graad van het secundair onderwijs. Er wordt zo inclusief mogelijk gewerkt. De centrale toetsen worden digitaal

afgenomen. De toetsafname wordt op school georganiseerd volgens een procedure die gelijk is voor alle scholen. De centrale toetsen worden centraal opgesteld, verwerkt en geanalyseerd[4].

3.2 FEEDBACK

Nadat leerlingen de centrale toetsen hebben afgelegd, ontvangt de school hierover feedback. Feedback kan gedefinieerd worden als informatie over het functioneren en presteren van een school, die extern en vertrouwelijke aan de school wordt aangeleverd (Visscher, 2002; Visscher & Coe, 2003). Feedback uit centrale toetsen betreft bijgevolg alle informatie (op school-, klas- en leerlingenniveau) die naar aanleiding van centrale toetsen afkomstig is uit het feedbacksysteem. Om een effectief gebruik van feedback te stimuleren toont onderzoek dat het belangrijk is dat feedback als relevant en betekenisvol wordt beschouwd en als valide en betrouwbaar gepercipieerd wordt.

3.2.1 De relevantie van feedback

Evidentie toont aan dat feedbackprocessen niet op gang komen wanneer de informatie die aan scholen teruggekoppeld wordt niet tegemoet komt aan de behoeften van de gebruikers (Van Gasse et al., 2015). Om optimaal feedbackgebruik te stimuleren, is het bijgevolg van belang dat feedback als relevant ervaren wordt. Hierbij heeft relevantie betrekking op de mate waarin de feedbackgebruiker de informatie als zinvol ervaart en is dus sterk context- en persoonsgebonden en niet zomaar te objectiveren (Van Gasse et al., 2015). Het is met andere woorden de perceptie van de gebruiker hieromtrent die uiteindelijk van belang is en niet alleen deze van de ontwikkelaars van de feedback (Visscher & Coe, 2003). Om effectief feedbackgebruik te bewerkstellingen, is het dus cruciaal dat feedback tegemoetkomt aan de informatiebehoeften van de gebruikers. Deze informatiebehoeften vloeien in sterke mate voort uit de verwachtingen omtrent feedback (Schildkamp & Teddlie, 2008; Van Petegem & Vanhoof, 2007; Vanhoof, Verhaeghe, et al., 2009; Visscher, 2002). Bovendien leert Vlaams onderzoek dat schoolleiders en leraren behoeften ervaren aan informatie die ze meteen kunnen relateren aan hun dagelijkse praktijk. In het kader van centrale toetsen start dit bij de relevantie van getoetste leergebieden en leerinhouden (Van Gasse et al., 2015).

[4] https://onderwijs.vlaanderen.be/nl/beleid/onderwijsregelgeving-in-ontwikkeling/vlaamse-toetsen

3.2.2 De betekenisvolheid van feedback

'Sterktes' en 'zwaktes' in onderwijskwaliteit van een school zijn niet zonder meer contextloos te bepalen. Toetsresultaten op zich zijn immers betekenisloos en niet meer dan een verzameling gegevens. Door zulke gegevens – binnen een bepaald referentiekader of criterium – af te toetsen, worden ze omgezet in informatie, krijgen ze betekenis en kunnen ze een basis vormen voor het nemen van beslissingen of het ondernemen van concrete acties (Neumann, Trautwein, & Nagy, 2011; Schildkamp, Rekers-Mombarg, & Harms, 2012; Vanhoof, Mahieu, & Van Petegem, 2009). Uit de literatuur blijken er drie criteria gangbaar: een standaardcriterium, een groepsnormcriterium en een zelfcriterium (Lysne, 1984). Bij een standaardcriterium (of criteriumgerichte scoring of absolute of standaardgebaseerde scoring) worden leerprestaties vergeleken met een absoluut doel of een standaard (Hamilton et al., 2002; Neumann et al., 2011), zoals een eindterm of een ontwikkelingsdoel. Een groepsnormcriterium (of normgerefereerd criterium of normatieve scoringsstrategie of relatieve scoring) spiegelt bepaalde toetsresultaten af tegen de betreffende toetsresultaten van anderen. Als het groepsnormcriterium op mesoniveau gebruikt wordt, worden de gemiddelde toetsresultaten in een school vergeleken met de gemiddelde toetsresultaten in andere scholen. Dergelijke vergelijkingsbasis wordt in de literatuur als '*benchmarking*' omschreven (Kelly, 2004). Bij een zelfcriterium (of leerlinggerichte of ipsatieve scoringsstrategie) worden de toetsresultaten van een leerling, klas, school of onderwijssysteem vergeleken met eerder behaalde toetsresultaten (Van Petegem & Vanhoof, 2002). Bij de interpretatie van centrale toetsen in het kader van leer(winst)monitoring worden deze criteria veelal gecombineerd (McMillan, Myran, & Workman, 2002; Willingham, Pollack, & Lewis, 2002).

Bij de invoering van centrale toetsen binnen het Vlaamse onderwijs staat op termijn 'leerwinst' centraal. Om leerwinst in kaart te brengen, wordt er gebruik gemaakt van een zelfcriterium. Leerwinst wordt immers beschouwd als de voortgang of toename in kennis, vaardigheden en competenties die een leerling doorheen een bepaalde periode van zijn schoolloopbaan maakt (Harris, 2011). Het is een maat voor de groei van individuele leerlingen en situeert zich in de eerste plaats op het niveau van de individuele leerling (Janssens, Rekers-Mombarg, & Lacor, 2014). Leerwinst kan tevens op niveau van de school en het niveau van het onderwijssysteem beschreven worden. Zo is de leerwinst van een school gelijk aan de gemiddelde leerwinst van leerlingen in de school en de leerwinst van een onderwijssysteem gelijk aan de gemiddelde leerwinst over alle leerlingen uit alle scholen heen (Castellano & Ho, 2015). Leerwinst kan op verschillende manieren geconceptualiseerd worden. Veelal wordt de leerwinst van een leerling bepaald op basis van een vergelijking van minstens twee leerprestaties van een leerling op twee verschillende toetsen, waarbij een positief verschil gezien wordt als leerwinst (Gong, Perie, & Dunn, 2006). Deze conceptualisering van leerwinst maakt het tevens mogelijk correctiefactoren zoals eerdere leerprestaties, geslacht, etniciteit, armoede, speciale

onderwijsbehoeften en beheersing moedertaal in rekening te brengen (Ray, 2006). Er zijn evenwel vele mogelijke correctiefactoren. Of en welke correctiefactoren precies in rekening moeten worden gebracht, blijft voer voor discussie (Wang, Beckett, & Brown, 2006).

Het aanverwante begrip 'toegevoegde waarde' richt zich op de prestatie van de school (Steunpunt, 2022a). Het is een maat voor de bijdrage van de school aan het leren (of de leerwinst) van leerlingen (Janssens et al., 2014). Een school waarin leerlingen meer leerwinst maken in vergelijking met een vooraf bepaalde verwachting of in vergelijking met andere scholen, heeft een positieve toegevoegde waarde. Een school waarin leerlingen daarentegen minder leerwinst maken dan vooraf verwacht heeft een negatieve toegevoegde waarde.

3.2.3 De perceptie van validiteit en betrouwbaarheid van feedback

Daarenboven is het om een zicht te krijgen op sterktes en zwaktes binnen de school van belang een zo accuraat en eerlijk mogelijk beeld van de onderwijs- en leerprestaties van de leerlingen en de school te schetsen. Om de objectiviteit van feedback te waarborgen, is het daarom zaak voldoende oog te hebben voor de validiteit en betrouwbaarheid van de feedback (Van Gasse et al., 2015).

Een moeilijkheid inzake de validiteit en betrouwbaarheid van feedback is de discrepantie tussen de perspectieven van waaruit men deze validiteit en betrouwbaarheid kan benaderen. Het blijkt daarom aangewezen dat feedback zowel aan de wetenschappelijke vereisten als aan het gebruikersperspectief van validiteit en betrouwbaarheid voldoet (Van Gasse et al., 2015). Opdat feedback waardevolle informatie bevat, moet zowel de totstandkoming van de feedback valide en betrouwbaar zijn, als ook zo door de gebruikers gepercipieerd worden. Gebruikers zijn immers pas geneigd met feedback aan de slag te gaan wanneer deze ook aan hun eigen vereisten inzake validiteit en betrouwbaarheid voldoet (Pierce & Chick, 2011; Schildkamp & Teddlie, 2008).

Zo blijkt uit Vlaams onderzoek dat schoolleiders en leraren selectief zijn op basis van hun perceptie met betrekking tot het correct in beeld brengen van onderwijs- en leerprestaties van de school. Scholen blijken sneller een beroep te doen op feedback die voldoende rekening houdt met achtergrondkenmerken van leerlingen en compositiekenmerken van scholen en zijn eerder niet geneigd feedback te gebruiken wanneer die naar hun gevoel niet 'fair' is (Vanhoof, Mahieu, et al., 2009).

3.3 FEEDBACKGEBRUIK

Feedbackgebruik uit centrale toetsen kan niet zonder meer letterlijk geïnterpreteerd worden en verengd worden tot het 'gebruik van feedback'. Achter het concept 'feedbackgebruik' gaat een complex proces met verschillende (f)actoren schuil, waarbij alle fasen gelieerd zijn aan vooropgestelde perspectieven en doelen en met elkaar interageren (Coburn, Touré, & Yamashita, 2009; Coburn & Turner, 2011; Mandinach, Honey, Light, & Brunner, 2008; Schildkamp, 2019; Schildkamp, Karbautzki, & Vanhoof, 2014).

3.3.1 Een systematisch en cyclisch proces

Effectief feedbackgebruik kenmerkt zich als een systematisch en cyclisch proces. Door verschillende fasen in het proces te doorlopen wordt op een systematische manier betekenis gegeven aan feedback uit centrale toetsen. Dat gebeurt stapsgewijs, door stil te staan bij en te bediscussiëren hoe de feedback geïnterpreteerd kan worden en wat deze betekent voor de school. Feedbackgebruik gaat echter ook verder, door te onderzoeken wat mogelijke verklaringen voor de feedback kunnen zijn en welke acties binnen de school gepast zijn (Schildkamp, Poortman, & Handelzalts, 2016).

Feedbackgebruik start bij het raadplegen van de feedback. Dit wil zeggen dat gebruikers een actieve rol spelen in het consulteren van de feedback uit centrale toetsen. Deze actieve rol kan voortkomen uit doelen (zie 3.3.2) die voorop gesteld worden in het kader van kwaliteitszorg (Schildkamp et al., 2016). Na het raadplegen van feedback start een proces van verkennen en overleggen. Het gaat erom te verkennen welke informatie de feedback naar de oppervlakte (kan) brengen en welke niet (Verhaeghe, Vanhoof, Valcke, & Van Petegem, 2010b). Vervolgens is het zaak om de feedback correct te interpreteren (Verhaeghe, Vanhoof, et al., 2010b). Dit wil zeggen dat er betekenis gegeven wordt aan de informatie uit het feedbacksysteem. Interpretatie van feedback vertrekt dus vanuit de vraag wat de resultaten precies betekenen voor de school. Na het interpreteren van de feedback kan men binnen de school of de klas gaan nadenken over mogelijke oorzaken en verklaringen voor de feedback. Het doel is met andere woorden een correcte diagnose te stellen (Verhaeghe, Vanhoof, et al., 2010b), waarvoor eventueel aanvullende informatie nodig kan zijn (Schildkamp et al., 2016). Wanneer oorzaken voor bepaald presteren gevonden zijn, is het zaak om gepaste acties te plannen en te implementeren. Deze acties kunnen heel divers zijn en kunnen zich situeren op verschillende lagen in de school: op het niveau van (bepaalde) leerlingen of groepen, van leraren of ondersteuning van leraren, of op het beleidsniveau (Johnson, 1998; Verhaeghe, Vanhoof, Valcke, & Van Petegem, 2010a). Bij het invoeren van zulke veranderingen is het tevens van belang te gepasten tijde

informatie te verzamelen over de effectiviteit van deze acties, door deze acties met andere woorden tijdig te evalueren.

Het belang van feedbackgebruik voor schoolontwikkeling zit echter niet alleen in concrete resultaten of uitkomsten van het proces. Het systematisch doorlopen van de verschillende activiteiten en terugkeren naar bepaalde fases wanneer het nodig is, kan op zichzelf ook waardevol zijn. Het kan een cultuur van reflectie introduceren, wat de groei van ideeën doorheen de tijd stimuleert en op deze manier beleidsvoering en besluitvorming kan beïnvloeden (Johnson, 1998).

3.3.2 Een doelgericht proces

Tot een proces van feedbackgebruik komen, kan pas wanneer er voldoende duidelijkheid is over wat men wil bereiken (Mandinach et al., 2008; Schildkamp & Poortman, 2015). Deze doelen krijgen in het algemeen vorm vanuit twee perspectieven: een ontwikkelingsperspief en een verantwoordingsperspectief (Coburn & Talbert, 2006; Diamond & Spillane, 2004; Mandinach, 2012; Schildkamp, Ehren, et al., 2012; Schildkamp et al., 2014; Schildkamp & Kuiper, 2010). De centrale gedachte bij feedbackgebruik uit centrale toetsen is dat de beschikbare feedback aanleiding kan zijn om beleid en praktijk onder de loep te nemen en op die manier mede een basis kan vormen voor het nemen van geïnformeerde beslissingen over de ontwikkeling van beleid en praktijk.

Vanuit een ontwikkelingsperspectief wendt men de aanwezige kennis en expertise van scholen aan om voornamelijk interne processen autonoom vorm te geven en bij te sturen (Steunpunt, 2022b; Vanhoof & Van Petegem, 2007), zodat een evolutie van de gangbare naar de wenselijke situatie mogelijk wordt (Visscher & Ehren, 2011). Scholen wenden met andere woorden de beschikbare feedback aan om het functioneren van de school in kaart te brengen, inzicht te verwerven in de kwaliteit van haar onderwijs(processen) en tot maatregelen te komen die de onderwijskwaliteit in het algemeen en het onderwijzen en leren in het bijzonder optimaliseren (Schildkamp, Visscher, & Luyten, 2009). Dit perspectief gaat hierbij uit van een gedeelde visie op de gewenste onderwijskwaliteit en richt zich op de door de school geformuleerde doelen met betrekking tot zowel output als processen (Vanhoof & Van Petegem, 2007). Schoolontwikkeling moet dus zichtbaar worden in bijvoorbeeld verbeterde leerprestaties van leerlingen, verbeterde onderwijsprocessen en een verbeterd functioneren van de school.

Verantwoording kan gedefinieerd worden als '*a relationship between an actor and a forum, in which the actor has an obligation to explain and to justify his or her conduct, the forum can pose questions and pass judgement, and the actor may face consequences*' (Bovens, 2007, p. 450). Vanuit een verantwoordingsperspectief zullen schoolleiders en leraren feedback met andere woorden aanwenden om het functioneren van de school en de schoolkwaliteit te verantwoorden. In dit kader kunnen we een onderscheid maken tussen interne en externe verantwoording (Abelmann, Elmore, Even, Kenyon, & Marshall, 1999; Elmore, 2005a; Pollock & Winton, 2016). Externe verantwoording impliceert een op prestaties gebaseerde verantwoording of – met andere woorden – een beleid dat scholen evalueert, beloont en sanctioneert op basis van gemeten leerprestaties van leerlingen (Elmore, 2005b). Van scholen wordt bijgevolg verwacht dat zij vooraf bepaalde doelstellingen verwezenlijken. Interne verantwoording kenmerkt zich in een samenhang en afstemming tussen individuele verantwoordelijkheden, collectieve verwachtingen (op schoolniveau) en de processen waarmee actoren binnen de school verantwoording afleggen over wat ze doen. Hierbij gaat interne verantwoording vooraf aan en bepaalt het het reflectief vermogen van de school. Interne verantwoording bepaalt bijgevolg mee de wijze waarop de implementatie van centrale toetsen gepercipieerd en toegepast wordt (Elmore, 2005b).

De balans tussen enerzijds een ontwikkelingsperspectief en anderzijds een verantwoordingsperspectief inzake feedbackgebruik kan beschreven worden aan de hand van vier factoren. Ten eerste wordt het perspectief bepaald door de mate van autonomie waarover scholen beschikken in het prioriteren van doelen en standaarden voor de leerprestaties van leerlingen. In onderwijssystemen met een sterkere nadruk op verantwoording expliciteert de overheid duidelijke doelstellingen en normen met betrekking tot (voornamelijk cognitieve) leerprestaties van leerlingen, terwijl scholen in een eerder ontwikkelingsgericht onderwijssysteem meer autonomie krijgen in het voorop stellen van doelen (Coghlan & Desurmont, 2007; Figlio & Page, 2003; O'Day, 2002). Ten tweede wordt het perspectief bepaald door de mate en de vorm waarin er controle aanwezig is op het bereiken van expliciet geformuleerde doelen en standaarden. De nadruk ligt sterker op verantwoording bij de aanwezigheid van een schoolexterne meting van de doelen en standaarden (Figlio & Page, 2003; O'Day, 2002) en sterker op ontwikkeling wanneer scholen meer autonomie krijgen in de methoden waarop men doelen evalueert (Coghlan & Desurmont, 2007). Ten derde hebben ook de gevolgen van de controle op het bereiken van doelen en standaarden een belangrijke invloed op de balans tussen verantwoording en schoolontwikkeling. Doelen in een verantwoordingsperspectief zijn gericht op het maken van een summatieve beoordeling waardoor de leerprestaties vanuit dit perspectief automatisch tot consequenties leiden (Vanhoof & Van Petegem, 2007; Wang et al., 2006; Woessmann, 2007). Ze kunnen leiden tot een blijvend of aangetast vertrouwen enerzijds en blijvende of verminderde

beschikbaarheid tot bestaande (overheids)middelen anderzijds (Geijsel, Krüger, & Sleegers, 2010). Vanuit een ontwikkelingsperspectief dient een gestandaardiseerde toets een formatief doel (Figlio & Page, 2003; O'Day, 2002). Ten slotte speelt de openbaarheid van informatie over de toetsresultaten van leerlingen en dientengevolge over de onderwijskwaliteit van de school een bepalende rol in het perspectief. In een verantwoordingsperspectief wordt deze informatie – via publieke databanken of *rankings* – doorgaans openbaar gemaakt. In een ontwikkelingsperspectief focust het evaluatiesysteem zich op haar formatieve karakter en wordt deze informatie – doorgaans – niet openbaar gemaakt (Figlio & Page, 2003; O'Day, 2002; Pollock & Winton, 2016; Van Petegem, Vanhoof, Daems, & Mahieu, 2005).

Binnen het samenspel van deze twee perspectieven kunnen vervolgens verschillende doelen, door verschillende stakeholders, op zowel micro-, meso- als macroniveau tot stand komen. Op basis van de literatuur worden in Tabel 1 voor elk niveau zowel vanuit een ontwikkelings- als vanuit een verantwoordingsperspectief mogelijke doelen geformuleerd.

Het microniveau betreft zowel feedbackgebruik op niveau van de klas en de individuele leraar als op niveau van de individuele leerling. Zo kan de leraar – vanuit een ontwikkelingsperspectief – op klasniveau feedback over zijn/haar professionele effectiviteit met betrekking tot de getoetste leerdoelen ontvangen. Op deze manier kan feedback de leraar helpen over de eigen onderwijspraktijk te reflecteren (Breiter & Light, 2006; Brunner et al., 2005; Schildkamp & Kuiper, 2010) en om het gebruik van instructie- en interventiedoelen te plannen (Schildkamp & Kuiper, 2010), waardoor het voor de leraar mogelijk wordt op een doelgerichte(re) en efficiënte(re) manier (vormen van) differentiatie in te zetten (Collins, Reiss, & Stobart, 2010). Tevens kan de feedback uit centrale toetsen leerlingen inzicht helpen verwerven in de eigen voortgang (Kerr, Marsh, Ikemoto, Darilek, & Barney, 2006). Vanuit een verantwoordingsperspectief kan de feedback, naast andere informatiebronnen (zoals observaties van de leraar, portfolio's en andere toetsresultaten) bijvoorbeeld worden ingezet om gesprekken met leerlingen en ouders over heroriënteringen in de loop van het schooljaar voor te bereiden (Breiter & Light, 2006; Brunner et al., 2005; Young, 2006) of om beslissingen te nemen met betrekking tot het al dan niet toe laten van leerlingen tot bepaalde vormen van vervolgonderwijs (Acquah, 2013; Atkinson & Geiser, 2009). Daarenboven kan de feedback een element zijn in de beoordeling van lerarenprestaties (Fryer, 2013; Harris, 2011). Dit gebruik lokt in de literatuur echter kritisch reacties op.

Uit onderzoek blijkt dat schoolleiders zich voornamelijk richten naar informatie op mesoniveau die hen kan ondersteunen in beleidsvoorbereiding en beleidsvoering (Schildkamp & Kuiper, 2010). Centrale toetsen

kunnen schoolleiders feedback aanreiken over de leerprestaties en voortgang van leerlingen over de klassen heen of over standaarden voor leerprestaties (Coburn & Talbert, 2006). In het kader van leer(winst)monitoring kunnen de resultaten op centrale toetsen daarenboven op schoolniveau geaggregeerd worden, waardoor kenmerken over kwaliteiten en werkpunten met betrekking tot onderwijzen en leren op schoolniveau in kaart kunnen worden gebracht (Schildkamp & Kuiper, 2010). Vanuit een ontwikkelingsperspectief kan dit de school ondersteunen bij haar processen van interne kwaliteitszorg en bij het vormgeven van effectieve professionaliseringstrajecten voor het schoolteam (Loeb, 2013; Schildkamp & Kuiper, 2010). Daarnaast kan de feedback ingezet worden voor verantwoordingsdoeleinden. Scholen kunnen de feedback bijvoorbeeld gebruiken in externe rapportages, zoals rapporten voor de inspectie (Coburn & Talbert, 2006; Diamond & Spillane, 2004; Lai & Schildkamp, 2013; Schildkamp & Kuiper, 2010; Wohlstetter, Datnow, & Park, 2008; Young, 2006). Tevens blijkt dat de resultaten van centrale toetsen openbaar worden gemaakt om marktmechanismen toe te laten waarbij ouders de scholen kiezen op basis van – onder meer – outputinformatie (Bishop & Wößmann, 2004). Scholen wenden op deze manier de resultaten van de centrale toetsen in hun *public relations* aan (Schildkamp & Kuiper, 2010) om zich – op een strategische manier – te verantwoorden over de eigen onderwijskwaliteit tegenover ouders (Saunders, 1999).

Ook op macroniveau kan de feedback uit centrale toetsen een informatierijke omgeving creëren voor de overheid in haar onderwijsbeleid en een belangrijke basis leggen voor de systematische evaluatie van haar effectiviteit (Loeb, 2013). Daarnaast kan het onderwijsbeleid zich aan de hand van deze objectieve informatie ook verantwoorden voor beleidskeuzes of het gevoerde beleid: "Metingen en cijfers maken een wezenlijk en zelfs erg bepalend deel uit van het ontwerpen, plannen, bijsturen en beheren van een onderwijssysteem" (Standaert, 2014, p. 16).

Tabel 1
Mogelijke doelen van feedbackgebruik uit centrale toetsen

	Ontwikkelingsperspectief	Verantwoordingsperspectief
Micro	- Zelfevaluatie leraren: professionele effectiviteit onderwijsprestaties - Differentiatie - Zelfevaluatie leerlingen: zicht op voortgang	- Beoordeling onderwijsprestaties leraren - Beslissingen vervolgonderwijs leerlingen
Meso	- Zelfevaluatie scholen: interne kwaliteitszorg - Basis voor professionaliseringstrajecten	- Verantwoording onderwijskwaliteit naar inspectie - Outputinformatie ten dienste van schoolkeuze

Macro	-	Beleidsontwikkeling	-	Beleidsevaluatie

3.4 STIMULERENDE EN HINDERENDE FACTOREN VOOR FEEDBACKGEBRUIK

Het louter implementeren van centrale toetsen zal niet voldoende zijn om de onderwijskwaliteit op te krikken. Schoolleiders en leraren moeten doelgericht aan de slag met de resulterende feedback. Uit de literatuur blijken naast feedbackfactoren (3.2) onder andere ook schoolfactoren (3.4.1), contextfactoren (3.4.2) en gebruikersfactoren (3.4.3) een stimulerende of hinderende rol te spelen bij feedbackgebruik door schoolleiders en leraren.

3.4.1 Schoolfactoren

De mate waarin feedbackgebruik uit centrale toetsen tot positieve effecten of onwenselijke neveneffecten leidt, vertoont tevens samenhang met hoe feedbackgebruik binnen de school vorm krijgt (Firestone et al., 2002).

3.4.1.1 Schoolinterne verwachtingen

Schoolinterne verwachtingen betreffen verwachtingen die scholen ten aanzien van feedbackgebruik naar voren schuiven. Uit onderzoek blijkt dat als leraren zich gedwongen voelen om feedback te gebruiken vanwege verantwoording, ze feedbackgebruik niet koppelen aan schoolontwikkeling (Hubbard, Datnow, & Pruyn, 2014; Schildkamp & Teddlie, 2008; Schildkamp & Visscher, 2010). Dit kan feedbackgebruik belemmeren, daar feedbackgebruik hierbij ervaren wordt als een verplichting (Sutherland, 2004). Daarnaast concludeerde Young (2006) dat schoolinterne verwachtingen feedbackgebruik kunnen belemmeren als leraren geen verband ervaren tussen feedbackgebruik en hun klaspraktijk. Aan de andere kant blijkt dat schoolinterne verwachtingen – bijvoorbeeld door middel van aanmoediging door de schoolleider – feedbackgebruik positief kunnen beïnvloeden (Schildkamp & Kuiper, 2010; Schildkamp & Teddlie, 2008; Schildkamp & Visscher, 2010). Daarenboven stelden Van Gasse en collega's (2015) vast dat hogere verwachtingen ten aanzien van feedbackgebruik, hetzij vanuit een ontwikkelings-, hetzij vanuit een verantwoordingsperspectief, een bevorderende invloed blijken te hebben op effectief feedbackgebruik.

3.4.1.2 Kenmerken van beleidsvoerend vermogen

Tevens blijken kenmerken van het beleidsvoorbereidend vermogen (Vanhoof & Van Petegem, 2017) van scholen samen te hangen met effectief feedbackgebruik. In de context van geïnformeerde besluitvorming blijkt uit de literatuur dat gezamenlijke doelgerichtheid (Schildkamp et al., 2014; Van Gasse et al., 2015; Vanhoof, Verhaeghe, Van Petegem, & Valcke, 2012) en ondersteunende relaties en samenwerking (Van Gasse et al., 2015) tot wezenlijke verschillen in feedbackgebruik kunnen leiden.

Hierbij spelen schoolleiders een sleutelrol. Ze hebben niet enkel een cruciale rol bij het formuleren van interne verwachtingen inzake feedbackgebruik (Park, Daly, & Guerra, 2013), ze kunnen tevens als transformationeel leider belangrijke 'betekenis-gevers' zijn om met feedback aan de slag te gaan (Cosner, 2011) en kunnen in die hoedanigheid ook de katalysator zijn in het opbouwen van een cultuur van feedbackgebruik (Vanhoof et al., 2012).

3.4.2 Contextuele factoren

Feedbackgebruik wordt niet opgezet in een beleidsmatig vacuüm. Kenmerken van het macro onderwijsbeleid kunnen impact hebben op het gebruik van feedback uit centrale toetsen (Vanhoof, Verhaeghe, et al., 2009).

3.4.2.1 Schoolexterne verwachtingen

Scholen worden verantwoordelijk geacht voor het bewaken van de eigen onderwijskwaliteit in het algemeen en van het monitoren en waarborgen van de leerprestaties van leerlingen in de richting van betreffende standaarden in het bijzonder (Vanhoof, Vanlommel, Thijs, & Vanderlocht, 2014). Dit maakt scholen enerzijds autonoom in de organisatie van haar onderwijs, maar impliceert anderzijds dat zij zich op basis van objectieve standaarden moeten verantwoorden over haar onderwijskwaliteit tegenover verschillende stakeholders in het onderwijsveld (Hofman, Dijkstra, & Hofman, 2009; Schildkamp, Ehren, et al., 2012; Van Gasse et al., 2015; Vanhoof et al., 2014). Bijgevolg zijn er twee perspectieven van waaruit schoolexterne verwachtingen rond feedbackgebruik binnen scholen vorm kunnen krijgen: vanuit een ontwikkelingsperspectief en vanuit een verantwoordingsperspectief. Binnen onderwijscontexten met een sterk verantwoordingsperspectief, wordt van scholen verwacht dat zij feedback gebruiken ter verantwoording van hun onderwijskwaliteit. Vanuit een ontwikkelingsperspectief wordt van scholen verwacht dat zij zich onder meer baseren op feedback om onderwijskwaliteit te verbeteren. Beide perspectieven hoeven elkaar niet uit te sluiten, maar de aanwezigheid van één van beide perspectieven beïnvloedt feedbackgebruik in scholen danig (Vanhoof, Verhaeghe, et al., 2009).

Daarnaast wordt de '*stakes'* of inzet van de centrale toets in de literatuur doorgaans bepaald door twee criteria. Ten eerste bepaalt het perspectief van waaruit de centrale toets wordt ingezet mee de inzet van de toets. *High-stakes* centrale toetsen worden vaak gebruikt in een verantwoordingsperspectief, terwijl *low-stakes* centrale toetsen eerder worden ingezet voor een ontwikkelingsgericht doel. Ten tweede is een centrale toets *high-stakes* als er rechtstreekse gevolgen verbonden zijn aan de toetsresultaten of wanneer het gebruik van de toetsresultaten – bedoeld of onbedoeld – kan leiden tot beslissingen met ernstige gevolgen voor leerlingen, leraren of scholen. Een systeem heeft daarentegen een lage inzet – of is met andere woorden *low-stakes* – als er nauwelijks consequenties aan de resultaten zijn verbonden (Au, 2007; Cizek, 2005; Verger & Parcerisa, 2018a).

3.4.3 Gebruikersfactoren

Finaal valt of staat feedbackgebruik bij het gebruik ervan door de betrokken gebruikers. Ongeacht het belang van andere randvoorwaarden, is de gebruiker zelf een cruciale schakel in het proces van feedbackgebruik. Zowel de attitude, de motivatie, de persoonlijke doelmatigheidsbeleving en bekommernissen van de individuele gebruiker ten aanzien van feedbackgebruik kunnen effectief feedbackgebruik stimuleren of hinderen (o.a. Winke, 2011).

3.4.3.1 Attitude ten aanzien van feedbackgebruik

Feedbackgebruik is mede afhankelijk van de attitude van de schoolleiders en leraren die geacht worden met de feedback aan de slag te gaan (Ikemoto & Marsh, 2007; Schildkamp & Kuiper, 2010). Elke schoolleider en elke leraar heeft immers een beeld over feedbackgebruik dat omschreven kan worden als de attitude hieromtrent. Een attitude is met andere woorden een geheel van normen, waarden, gevoelens, ideeën en meningen die bepalen hoe een persoon zich in een bepaalde situatie (i.c. feedbackgebruik) gedraagt (Krathwohl, Bloom, & Masia, 1971). Een negatieve attitude van leraren en schoolleiders wordt daarbij als een van de voornaamste hinderpalen gezien voor het gebruik van feedback (Datnow, Park, & Kennedy-Lewis, 2012; Schildkamp, Rekers-Mombarg, et al., 2012; Verhaeghe, Vanhoof, et al., 2010b). Een positieve houding daarentegen kan feedbackgebruik stimuleren (Sach, 2013).

In de attitude ten aanzien van feedbackgebruik kunnen zowel een cognitieve als een affectieve component onderscheiden worden (Sanbonmatsu & Fazio, 1990). De cognitieve component omvat overtuigingen die bepalen hoe een schoolleider of een leraar over feedbackgebruik denkt. Deze component betreft de mate

waarin schoolleiders en leerkrachten het feedbackgebruik waardevol vinden en ervan overtuigd zijn dat feedbackgebruik een meerwaarde vormt om het schoolbeleid en de -praktijk vorm te geven (Sanbonmatsu & Fazio, 1990). De affectieve component omvat de gevoelens van schoolleiders en leraren die een invloed hebben op keuzes die ze maken om met feedback aan de slag te gaan. Deze component bepaalt de mate waarin schoolleiders en leraren zich comfortabel voelen in een feedbackgebruikactiviteiten, enthousiast zijn over het gebruik van feedback op school- en klasniveau, of net gevoelens van angst of onzekerheid hieromtrent ervaren (Sanbonmatsu & Fazio, 1990).

Onderzoek toont aan dat mensen de eigen cognitieve attitude over een onderwerp vaak negeren en hun gedrag eerder laten leiden door hun affectieve attitude. In de context van feedbackgebruik wil dit mogelijk zeggen dat schoolleiders en leraren zich in hun handelen niet alleen laten leiden door de cognitieve kennis die ze over feedbackgebruik hebben, maar dat ook – of zelfs vooral – het gevoelsaspect een rol zal spelen (Sanbonmatsu & Fazio, 1990; Vanhoof et al., 2014).

3.4.3.2 Motivatie ten aanzien van feedbackgebruik

Feedbackgebruik in ontwikkelingsgerichte onderwijssystemen is sterk afhankelijk van de kwaliteit van motivatie van stakeholders (Sutherland, 2004; Vanlommel, Vanhoof, & Van Petegem, 2016). Bij het bestuderen van motivatie, onderscheidt de zelfdeterminatietheorie (Deci & Ryan, 2000) zich van andere motivatietheorieën doordat deze theorie de nadruk legt op deze kwaliteit van de motivatie in plaats van op de kwantiteit van de motivatie (Vansteenkiste, Lens, & Deci, 2007). In de zelfdeterminatietheorie (Tabel 2) hangen verschillen in de kwaliteit van motivatie samen met de mate waarin (extrinsiek gemotiveerd) gedrag (i.c. feedbackgebruik) autonoom dan wel gecontroleerd wordt gereguleerd. Gedragsregulatie geeft aan waarom mensen dingen doen. Het onderscheid tussen autonome en gecontroleerde motivatie is in grote lijnen terug te brengen tot het verschil tussen ‘willen’ en ‘moeten’ aan de slag gaan met de feedback (Vansteenkiste et al., 2007). In beide gevallen is er sprake van motivatie, al is de bron van motivatie sterk verschillend.

Tabel 2
Schematisch overzicht van de verschillende types van motivatie volgens de zelfdeterminatietheorie van Deci & Ryan (2000)

	Autonome motivatie		Gecontroleerde motivatie		Amotivatie
Hoeveelheid motivatie	Hoog	Hoog	Hoog	Hoog	Laag
Type motivatie	Intrinsiek	Extrinsiek	Extrinsiek	Extrinsiek	Amotivatie
Type regulatie	Intrinsiek	geïdentificeerd	Geïntrojecteerd	Extern gereguleerd	Niet gereguleerd
Onderliggende emoties	Welwillendheid, vrijheid	Welwillendheid, vrijheid	Stress, druk	Stress, druk	Apathie, hulpeloosheid
Kwaliteit motivatie	++ ← --				→

Bij autonome motivatie gaat het om schoolleiders en leraren die vrijwillig aan de slag gaan met feedback. Dit sluit in eerste instantie aan bij intrinsieke regulatie. Gaan schoolleiders of leraren aan de slag met feedback omdat ze ervan overtuigd zijn dat het een persoonlijk relevant doel kan dienen, is er sprak van geïdentificeerde motivatie.

Bij gecontroleerde motivatie voor feedbackgebruik ervaren schoolleiders en leraren druk, waardoor er eerder sprake is van 'moeten' (Deci & Ryan, 2000). Geïntrojecteerde regulatie kenmerkt zich door het ervaren van controle en druk vanuit de schoolleider of leraar zelf, waardoor de schoolleider of leraar zijn/haar eigenwaarde associeert met het al dan niet gebruiken van feedback. Externe regulatie is een vorm van extrinsieke motivatie waarbij schoolleiders en leraren feedback gebruiken om beloningen te verkrijgen of om straffen of kritiek te vermijden en het gevoel ervaren dat hun gedrag volledig door factoren van buitenaf wordt gestuurd (Deci & Ryan, 2000).

Amotivatie is te omschrijven als een gebrek aan motivatie. Schoolleiders en leraren ervaren in dit geval doorgaans gevoelens van hulpeloosheid en/of gedragen zich apathisch. Hun verwachtingen zijn zeer laag, gaan ervan uit dat ze toch geen controle hebben over het bevorderen van de onderwijskwaliteit en hebben het gevoel dat ze hun gedrag niet kunnen reguleren. Het gevolg hiervan is een totaal gebrek aan motivatie, zowel intrinsieke als extrinsieke (Deci & Ryan, 2000).

De zelfdeterminatietheorie stelt dat autonome motivatie kwaliteitsvoller is dan gecontroleerde motivatie. Autonome motivatie stimuleert immers tot optimaal functioneren, terwijl gecontroleerde motivatie ervoor zorgt dat mensen minder goed presteren (Vansteenkiste et al., 2007). Als schoolleiders en leraren hun motivatieoriëntaties daarenboven kunnen verschuiven van extrinsiek naar intrinsiek, zal feedbackgebruik bijkomend gestimuleerd worden (Sutherland, 2004).

3.4.3.3 Persoonlijke doelmatigheidsbeleving ten aanzien van feedbackgebruik

Naast de attitude en motivatie nemen we ook de persoonlijke doelmatigheidsbeleving op als factor die een invloed kan uitoefenen op het feedbackgebruik van schoolleiders en leraren. Het uitgangspunt van persoonlijke doelmatigheidsbeleving of '*self-efficacy'* is dat wat personen denken, geloven en voelen hun handelen beïnvloedt. Het draait met andere woorden om het geloof in eigen kunnen.

Bandura (1997) definieerde het concept 'doelmatigheidsbeleving' als een zelfreflectieve, toekomstgerichte overtuiging dat men over de nodige competenties beschikt om een taak met succes te volbrengen. *Self-efficacy* beschrijft dus in welke mate iemand zichzelf voldoende competent acht binnen een bepaald domein om activiteiten zo te organiseren en te ordenen dat ze resulteren in gewenste resultaten. Schoolleiders en leerkrachten met een sterke persoonlijke doelmatigheidsbeleving zullen ervan overtuigd zijn dat zij succesvol feedback kunnen gebruiken om hun doel te bereiken en zullen daardoor dan ook meer ambitieuze doelen voor feedbackgebruik vooropstellen (Deci & Ryan, 2000; Vanhoof et al., 2014; Woolfolk, 2008).

Eigen aan de persoonlijke doelmatigheidsbeleving is dat deze kan verschillen in algemeenheid. Personen kunnen zich doelmatig achten voor een bepaald scala aan activiteiten dan wel ten aanzien van bepaalde handelingen. Dit maakt dat de persoonlijke doelmatigheid van schoolleiders en leraren kan variëren naargelang de soort activiteit die voorligt of de situatie waarin men tot feedbackgebruik wil komen. Uit de wetenschappelijke literatuur komt naar voor dat schoolleiders en leraren in vele gevallen twijfelen aan hun eigen bekwaamheden om met feedback aan de slag te gaan binnen de school (Kowalski & Lasley, 2009). De reden voor deze lagere doelmatigheid is niet altijd eenduidig. In sommige gevallen acht men zichzelf niet voldoende datageletterd. Schoolleiders en leraren zien zich dan niet in staat om feedback bijvoorbeeld te analyseren of te interpreteren, waardoor ze onvoldoende tot feedbackgebruik komen (Pierce & Chick, 2011). Het is echter ook mogelijk dat schoolleiders en leraren twijfelen aan de eigen bekwaamheid om beslissingen en concrete acties aan de feedback te koppelen (Kowalski & Lasley, 2009). Binnen het proces van

feedbackgebruik (zie 3.3) kan de persoonlijke doelmatigheidsbeleving van schoolleiders en leraren dus verschillende fasen beïnvloeden.

3.4.3.4 Bekommernissen bij het implementatieproces

Onderzoek leert dat schoolleiders en leraren bij innovatieprocessen – i.c. de implementatie van feedbackgebruik uit centrale toetsen – specifieke en individualistische bekommernissen hebben over de verandering en hun betrokkenheid erbij. Om deze bekommernissen het hoofd te kunnen bieden, werd het *Concerns-Based Adoption Model* (CBAM) ontwikkeld (Hall & Hord, 1979). Bekommernissen betreffen in dit model 'een samengestelde voorstelling van de gevoelens, zorgen, gedachten en overwegingen die gegeven worden aan een bepaalde kwestie of taak' (i.c. feedbackgebruik) (Hall, George, & Rutherford, 1979) en evolueren volgens een hiërarchisch patroon van zeven stadia, georganiseerd in drie ontwikkelingsniveaus (Hall & Hord, 1987, 2006; Hall & Hord, 2020) (Tabel 3).

In een eerste ontwikkelingsniveau zal een schoolleider of leraar zich gewoonlijk afvragen wat de innovatie voor hem of haar betekent. Bekommernissen hieromtrent houden verband met vragen en onzekerheden die wijzen op een sterke zelfbetrokkenheid en leiden veelal tot gevoelsmatige reacties. Als de zelfbetrokkenheidsbekommernissen in intensiteit verminderen of uitvoerig zijn besproken, blijkt de aandacht daarna te verschuiven naar de processen en taken in verband met het gebruik van een innovatie in de concrete praktische situatie. Hierbij ontstaan bekommernissen met meer taakgerichte vragen, waarbij men dan de aandacht richt op de consequenties van de vernieuwing voor de eigen processen en taken. Tenslotte volgt gewoonlijk de variant waarbij de aandacht vooral gericht zal zijn op anderen (George, Hall, & Stiegelbauer, 2006; Hall & Hord, 2020; Hord, Rutherford, Austin, & Hall, 1987; Vandenberghe, 1983).

De zeven stadia zijn hierbij indicatief voor iemands bereidheid om een nieuwe innovatie – i.c. feedbackgebruik – te adopteren (Charalambous & Philippou, 2010; Hall & Hord, 2006; Hall & Hord, 2020). Bekommernissen van een hoger niveau zullen bijgevolg duiden op adoptie van feedbackgebruik, terwijl bekommernissen van een lager niveau meer duiden op weerstand of gebrek aan voorbereiding om feedback te adopteren (George et al., 2006; Hall & Hord, 2020).

Tabel 3
Stadia en niveaus van bekommernissen, naar (George et al., 2006)

Ontwikkelingsniveaus	Ontwikkelingsstadia		
ANDERE-bekommernissen	6	Herziening	De gebruiker richt zich op het verkennen van manieren om meer universele voordelen uit de innovatie te halen, inclusief de mogelijkheid om er grote wijzigingen aan te brengen of te vervangen door een krachtiger alternatief.
	5	Samenwerking	De gebruiker richt zich op het coördineren en samenwerken met anderen met betrekking tot het gebruik van de innovatie.
	4	Gevolgen voor de leerling	De gebruiker richt zich op de impact van de innovatie op leerlingen in zijn of haar onmiddellijke invloedssfeer.
TAAK-bekommernissen	3	Beheersing	De gebruiker richt zich op de processen en taken bij het gebruik van de innovatie.
ZELF-bekommernissen	2	Persoonlijke betrokkenheid	De gebruiker is onzeker over de eisen van de vernieuwingen, over zijn/haar geschiktheid om aan die eisen te voldoen, en/of zijn/haar rol bij de innovatie.
	1	Informatie	De gebruiker geeft blijk van algemene bekendheid met de innovatie en interesse om er meer details over te leren. De gebruiker lijkt niet bezorgd over zichzelf in relatie tot de innovatie.
	0	Bewustwording	De gebruiker toont weinig bezorgdheid over of betrokkenheid bij de innovatie.

3.5 EFFECTEN VAN FEEDBACKGEBRUIK

De relatie tussen verantwoording en ontwikkeling kan verschillende configuraties hebben en in elk onderwijssysteem waarin centrale toetsen worden geïmplementeerd, gaat men ervan uit dat dit zal leiden tot een verbetering van het onderwijs. Toch blijkt er steeds een spanningsveld te bestaan tussen feedbackgebruik voor verantwoordingsdoelen en ontwikkelingsdoelen (Hofman et al., 2009; Maier, 2010; Vanhoof & Van Petegem, 2007). Hoewel het verantwoordingsperspectief voor scholen een stimulans kan zijn voor feedbackgebruik, laat men zich in een sterke op verantwoording gerichte context – waarbij leer- en onderwijsprestaties veelal geopenbaard worden – eerder leiden door *outcomes* dan door processen (Elmore, Abelmann, & Fuhrman, 1996), waardoor – in dit perspectief – de focus van feedbackgebruik voornamelijk op (het verhogen van) leerlingenprestaties ligt (O'Day, 2002). Te sterke gerichtheid naar verantwoording binnen een onderwijssysteem brengt bijgevolg een te grote resultaatsgerichtheid met zich mee, wat er mogelijks toe leidt dat het doel van feedbackgebruik – het verbeteren van de kwaliteit van

onderwijs binnen de school – uit het oog verloren wordt (Schildkamp, Ehren, et al., 2012). Onderzoek suggereert dan ook dat een te sterke nadruk op verantwoording soms tot onbedoelde en onwenselijke gevolgen kan leiden (Au, 2007; Ehren & Swanborn, 2012; Lee, Louis, & Anderson, 2012).

In Tabel 4 worden veronderstelde effecten van feedbackgebruik uit centrale toetsen op micro-, meso-, en macroniveau beschreven. Er wordt hierbij een onderscheid gemaakt tussen effecten die als gewenst beschouwd kunnen worden en neveneffecten die als onwenselijk geacht kunnen worden.

Tabel 4
Veronderstelde gewenste en onwenselijke effecten van (het gebruik van) centrale toetsen op micro-, meso- en macroniveau (Naar Penninckx, Vanhoof, Quintelier, De Maeyer, & Van Petegem, 2017)

	Gewenste effecten	Onwenselijke effecten
Micro	- Beter beeld van leerprestaties door een meer betrouwbare en valide evaluatie - Garantie op gelijke minimumdoelen voor alle leerlingen in alle scholen - Betere leerprestaties - Beter onderwijzen - Impact op doelmatigheidsbeleving en motivatie van leerlingen - Betere evaluatie door leraren zelf - Betere onderbouwing studiekeuze - Betere onderbouwing schoolkeuze	- Emotionele impact op schoolleiders, leraren en leerlingen - Impact op motivatie van leerlingen - Impact op motivatie, professionaliteit en status van leraren - *Teaching (to) the test* - Impact op studieloopbaan - Onwenselijk strategisch gedrag van leerlingen - Werkt ongelijkheid in de hand
Meso	- Betere interne kwaliteitszorg - Betere inzet van professionele ontwikkeling van schoolteams - Betere inzet van schoolinspectie	- *Cheating* - *Rankings*
Macro	- Beter inzicht in onderwijskwaliteit - Bijstellen van curriculum	- Grotere segregatie tussen 'sterke' en 'zwakke' scholen

4 METHODOLOGIE

Voor dit kwantitatief gebruikersonderzoek voerden we een *web-based* surveyonderzoek uit om verwachtingen bij en randvoorwaarden voor feedbackgebruik uit centrale toetsen Vlaanderen-breed in kaart te brengen.

In de beschrijving van dit luik beschrijven we achtereenvolgens de instrumentontwikkeling, de steekproeftrekking en de dataverzameling en data-analyse.

4.1 INSTRUMENT(ONTWIKKELING)

Voor dit kwantitatief gebruikersonderzoek werd een vragenlijst ontwikkeld. Er werd op toegezien dat zoveel mogelijk conceptuele uitgangspunten in de vragenlijst verwerkt werden. De inhoud van deze vragenlijst werd vooreerst samengesteld uit bestaande, gevalideerde schalen en geadapteerde items uit eerder onderzoek naar informatiegebruik in het algemeen en onderzoek naar gestandaardiseerde/centrale toetsen en feedbackgebruik in het bijzonder. Daarnaast werden bepaalde items zelf geconstrueerd. Hiervoor werd om de ecologische validiteit van de vragenlijst te versterken zowel inspiratie geput uit eerder afgenomen diepte-interviews, als uit de bestaande literatuur.

Tabel 5 geeft een overzicht van de relatie tussen de gemeten concepten en de onderzoeksvragen en geeft ook een indicatie van welke schalen al dan niet reeds gevalideerd werden in eerder wetenschappelijk onderzoek.

Tabel 5
Relatie onderzoeksvragen – conceptuele uitgangspunten en gebruik schalen

	Concepten	Schaal
OV1	Wat maakt feedback uit centrale toetsen relevant, gepercipieerd betrouwbaar en valide, en betekenisvol?	
	Relevantie van feedback	Nieuw geconstrueerd, o.b.v. kwalitatief gebruikersonderzoek en literatuur (Penninckx et al., 2017; Schildkamp, Rekers-Mombarg, et al., 2012; Van Gasse et al., 2015)

	Gepercipieerde betrouwbaarheid en validiteit van feedback	Nieuw geconstrueerd, o.b.v. kwalitatief gebruikersonderzoek en literatuur (Penninckx et al., 2017; Schildkamp, Rekers-Mombarg, et al., 2012; Van Gasse et al., 2015)
	Betekenisvolheid van feedback	Nieuw geconstrueerd, o.b.v. kwalitatief gebruikersonderzoek en literatuur (Penninckx et al., 2017; Schildkamp, Rekers-Mombarg, et al., 2012; Van Gasse et al., 2015)
OV2	Vanuit welke doelen verwachten Vlaamse schoolleiders en leraren feedback uit centrale toetsen aan de slag te gaan?	
	Doelen feedbackgebruik	Gebaseerd op gevalideerde schalen, vertaald naar context centrale toetsen (Penninckx et al., 2017; Van Gasse et al., 2015)
OV3	Welke feedbacksysteem-, school-, context- en gebruikersfactoren beïnvloeden Vlaamse schoolleiders en leraren in hun feedbackgebruik uit centrale toetsen?	
	Feedbacksysteemfactoren	Nieuw geconstrueerd, o.b.v. kwalitatief gebruikersonderzoek
	Schoolfactoren	Gebaseerd op gevalideerde schalen, vertaald naar context centrale toetsen (Jimerson, 2016; Schildkamp & Poortman, 2015; Van Gasse et al., 2015; Vanhoof & Van Petegem, 2017; Wayman, Wilkerson, Cho, Mandinach, & Supovitz, 2016)
	Contextfactoren	Nieuw geconstrueerd, o.b.v. literatuur (Jimerson, 2016; Pierce & Chick, 2011; Van Gasse et al., 2015)
	Gebruikersfactoren	Deels nieuw geconstrueerd, deels gebaseerd op gevalideerde schalen, vertaald naar context centrale toetsen (Bandura & Walters, 1977; Hall & Hord, 1987; Pierce & Chick, 2011; Pierce, Chick, & Gordon, 2013; Ryan & Connell, 1989; Van Gasse et al., 2015; Vanlommel et al., 2016; Verhaeghe, Vanhoof, et al., 2010b; Wayman et al., 2016)
OV4	Welke onwenselijke effecten verwachten Vlaamse schoolleiders en leraren van feedbackgebruik uit centrale toetsen?	
	Onwenselijke effecten	Nieuw geconstrueerd, o.b.v literatuur (Penninckx et al., 2017; Rogers, Barblett, & Robinson, 2018)

Een eerste versie van de vragenlijst werd in oktober 2021 voor feedback voorgelegd aan de stuurgroep en leden van het Steunpunt. Op basis hiervan werden verschillende items geherformuleerd en werden ook een aantal items toegevoegd aan de vragenlijst. Vervolgens werd de vragenlijst besproken met mensen uit het werkveld (i.c. een (gewezen) directeur secundair onderwijs en twee leraren Nederlands secundair onderwijs

(A-stroom) om na te gaan of de items eenduidig en correct interpreteerbaar waren. In december 2021 werd de vragenlijst voor dit gebruikersonderzoek herwerkt tot in haar definitieve vorm, zoals ze als online survey zou worden afgenomen. Er werd een versie ontwikkeld voor de bevraging van scholen waarvan zowel schoolleiders als leraren zouden deelnemen en een versie voor de scholen waarvan enkel de schoolleider zou deelnemen.

In een eerste gedeelte van de vragenlijst werd gevraagd naar achtergrondkenmerken. In het corpus van de vragenlijst peilden we naar zowel verwachte doelen en effecten van feedbackgebruik als naar feedbackfactoren (zoals relevantie, betekenisvolheid en gepercipieerde betrouwbaarheid en validiteit van feedback), feedbacksysteemfactoren (toegang tot feedback), schoolfactoren (zoals huidig gebruik van gestandaardiseerde toetsen, aanwezige kenmerken van beleidsvoerend vermogen en interne verwachtingen), contextfactoren (externe verwachtingen) en gebruikersfactoren (houding, motivatie, bekommernissen en persoonlijke doelmatigheid ten aanzien van feedbackgebruik) die feedbackgebruik kunnen stimuleren of hinderen. Tot slot trokken we het perspectief breder en bevroegen we een aantal specifieke actuele stellingen over het systeem van centrale toetsen.

De vragenlijst bestond voornamelijk uit gesloten vragen die verplicht te beantwoorden waren. Voor het merendeel waren dit beoordelingsvragen: stellingen die de respondent dienden te beoordelen aan de hand van een vijfpuntenschaal van (1) 'helemaal oneens' over 'eerder oneens', 'noch eens, noch oneens', 'eerder eens' tot (5) 'helemaal eens'. Bij deze items kon men ook de antwoordoptie 'niet van toepassing / ik weet het niet' aanduiden. In de inleiding bij de vragenlijst werd expliciet omschreven dat deze antwoordoptie enkel geselecteerd kon worden indien de vraag onduidelijk was of indien de respondent niet over de nodige informatie beschikte om een correcte inschatting van het item te maken. De survey bevatte ook enkele open velden om de respondent de kans te geven om aanvullende toelichtingen te geven of bijkomende opmerkingen te formuleren. Deze open velden konden wel opengelaten en overgeslagen worden.

De vragenlijst werd gedigitaliseerd en afgenomen als een online survey via Qualtrics.

4.2 STEEKPROEFTREKKING EN DATAVERZAMELING

De onderzoekspopulatie voor dit gebruikersonderzoek bestond uit alle Vlaams lagere scholen en secundaire scholen (exclusief bijzonder onderwijs). Daaruit werd een aselecte gestratificeerde representatieve

steekproef en reservesteekproef getrokken van lagere scholen en secundaire scholen, waarvan zowel de schoolleider als leraren (groep 'schoolleider en leraren') zouden deelnemen en eveneens een steekproef en reservesteekproef van lagere scholen en secundaire scholen waarvan enkel de schoolleider (groep 'schoolleiders') zou deelnemen. Voor beide groepen beoogden we een deelname van 30 lagere scholen en 30 secundaire scholen. In het lager onderwijs beoogden we de deelname van minimaal 5 teamleden per school, in het secundair onderwijs beoogden we de deelname van minimaal 12 teamleden per school. De stratificatie gebeurde naar onderwijsnet en schoolgrootte.

In november 2021 werd zowel de steekproef en reservesteekproef uit het lager onderwijs als de steekproef en reservesteekproef uit het secundair onderwijs digitaal aangeschreven en uitgenodigd tot deelname aan het surveyonderzoek. Voor het lager onderwijs schreven we 58 scholen aan om zowel de schoolleider als leraren uit te nodigen tot deelname aan het onderzoek en 56 scholen aan om enkel de schoolleider uit te nodigen. Voor het secundair onderwijs schreven we 48 scholen aan om zowel de schoolleider als leraren uit te nodigen tot deelname en 53 scholen aan om enkel de schoolleider uit te nodigen. Spoedig werd vervolgens gestart met het telefonisch contacteren van scholen van wie we tot dan toe nog geen reactie op de uitnodiging tot deelname ontvangen hadden. Net iets minder dan de helft van het aantal uitgenodigde scholen uit het lager onderwijs en net iets meer dan de helft van het aantal uitgenodigde scholen uit het secundair onderwijs beantwoordden de vraag tot deelname positief.

Begin december 2021 verstuurden we de deelnamelink naar de lagere en secundaire scholen die toegezegd hadden voor de deelname. De respons van de deelnemende scholen werd wekelijks opgevolgd. Scholen met een matige responsgraad werden herhaaldelijk via mail gevraagd de deelname tot de bevraging in de school te promoten. Scholen die de deelnamelink nog niet verspreid hadden, contacteerden we tevens telefonisch on na te gaan waarom de bevraging nog niet van start gegaan was. Voornamelijk de hectiek van de coronaperikelen bleek voor deze scholen hiertoe de reden. Als deadline voor het voltooien van de survey werd initieel 24 december 2021 vooropgesteld.

Omdat de responsgraad in het lager onderwijs eerder laag bleef, beslisten we in januari 2022 een bijkomende steekproef van 20 vervangingsscholen uit het lager onderwijs te trekken voor de bevraging van de schoolleider en leraren. Voor het aanschrijven en opvolgen van deze scholen hanteerden we een zelfde werkwijze als hierboven beschreven. Tevens contacteerden we opnieuw alle scholen uit de eerste steekproef

met een matige responsgraad om de deelnamelink nogmaals in de school te verspreiden en deelname aan de bevraging nogmaals te promoten. De bevraging werd op 20 februari 2022 definitief afgesloten.

Uiteindelijk vonden we 32 lagere scholen en 30 secundaire scholen bereid de schoolleider en leraren te laten deelnamen aan het onderzoek. Het invullen van de vragenlijst verliep echter – omwille van coronaperikelen – in alle scholen niet even vlot. Hoewel verschillende schoolleiders via telefoon de deelname van de school bevestigden, zijn in totaal 8 lagere scholen en 2 secundaire scholen niet met de vragenlijst gestart. Dit maakt dat uiteindelijk respondenten uit 24 lagere en uit 28 secundaire scholen voor de groep 'schoolleider en leraren' aan het gebruikersonderzoek hebben deelgenomen (zie Tabel 6).

In twee lagere scholen namen 2 respondenten deel aan de bevraging, in drie lagere scholen namen 3 respondenten deel aan de bevraging, in één lagere school namen 4 respondenten deel aan de bevraging, in twee lagere scholen namen 5 respondenten deel aan de bevraging en in 16 lagere scholen namen meer dan vijf respondenten deel aan de bevraging. In totaal namen in het lager onderwijs 201 respondenten uit de groep 'schoolleider en leraren' deel aan de bevraging.

Uit het secundair onderwijs namen 28 scholen deel aan de bevraging van schoolleiders en leraren. Eén school telt 4 respondenten, één school 5 respondenten en 26 scholen meer dan vijf respondenten. In totaal vulden voor het secundair onderwijs 529 respondenten uit de groep 'schoolleider en leraren' de vragenlijst in (zie Tabel 7).

De gemiddelde responsgraad voor de groep 'schoolleider en leraren' bedraagt voor het lager onderwijs 8 respondenten (met een minimum aantal respondenten per school van 2 en een maximum aantal respondenten per school van 24) en voor het secundair onderwijs 19 respondenten (met een minimum aantal respondenten per school van 4 en een maximum aantal respondenten per school van 53).

Daarnaast beantwoordden uiteindelijk 30 schoolleiders (groep 'schoolleiders') uit het lager onderwijs en 36 schoolleiders uit het secundair onderwijs de vragenlijst voor de groep 'schoolleider' (zie Tabel 6 en Tabel 7).

In totaal namen 796 respondenten verspreid over 118 scholen deel aan de online bevraging (zie Tabel 6 en Tabel 7).

Tabel 6
Respons op schoolniveau, voor zowel het lager als het secundair onderwijs

	Lager onderwijs		Secundair onderwijs		
	Schoolleider en leraren	Enkel schoolleider	Schoolleider en leraren	Enkel schoolleider	Totaal aantal betrokken scholen
	n	n	n	n	
Respons op schoolniveau					
scholen	24	30	28	36	118

Tabel 7
Respons op respondentenniveau en de mate waarin de vragenlijst al dan niet volledig werd ingevuld, voor zowel het lager als het secundair onderwijs

	Lager onderwijs				Secundair onderwijs				
	Schoolleider en leraren		Schoolleider		Schoolleider en leraren		Schoolleider		
	n	%	n	%	n	%	n	%	n
Respons op respondentenniveau									
< 50%	40	19.9	5	16.7	96	18.1	1	2.8	142
50% <100%	9	4.5	1	3.3	42	7.9	1	2.8	53
100%	152	75.6	24	80	391	73.9	34	94.4	601
Totaal	201	100%	30	100%	529	100%	36	100%	796

4.3 DATA-ANALYSE

Met het oog op het beantwoorden van de vooropgestelde onderzoeksvragen worden in dit gebruikersonderzoek vooral beschrijvende statistieken gehanteerd. Daar waar aan inferentiële statistiek gedaan wordt, hanteren we een 95%-betrouwbaarheidsinterval om statistische (on)zekerheid uit te drukken. Wanneer we aangeven dat (bijvoorbeeld) een verschil in onze observaties statistisch significant is, is er sprake van een zeer kleine kans (≤5%) dat gelijkaardige verschillen kunnen vastgesteld worden indien er geen verschil is in de gehele populatie. Dit betekent met ander woorden dat we in dergelijke gevallen op basis van berekeningen voor 95 procent zeker zijn dat dat verschil zich ook effectief voordoet in de

populatie. Waar we asterisken gebruiken om significantie aan de duiden, worden de volgende gangbare grenswaarden gebruikt: * $p < .05$, ** $p < .01$, *** $p < .001$.

4.3.1 Beschrijvende statistiek

4.3.1.1 Analyse van kenmerken van respondenten en hun scholen

Meerkeuze-items over de achtergrondkenmerken van de respondenten en de scholen waarover zij de survey invulden, worden waar mogelijk geanalyseerd aan de hand van frequentietabellen. Responsen op de open vraag met betrekking tot het aantal jaren ervaring in de functie waaraan de respondenten binnen de school de meeste tijd besteden worden mee betrokken in de beschrijvende analyses.

4.3.1.2 Analyses op itemniveau

De items uit de vragenlijst worden beschrijvend geanalyseerd door kengetallen van ligging en spreiding. Per item beschrijven we:

- *n* - het aantal respondenten dat een punt op de vijfpunts-antwoordschaal aanduidt, met andere woorden: het aantal respondenten dat een geldig antwoord geeft op het item in kwestie.
- Percentage **WN/NVT** – het aandeel respondenten dat aangeeft dat het item onduidelijk is of niet over de nodige info beschikt om een correcte inschatting van het item te maken. Dit aandeel respondenten wordt niet opgenomen in verdere berekeningen.
- **GEM** - de gemiddelde score over alle respondenten heen, berekend op basis van de gehanteerde vijfpuntenschaal.
- *SD* - de standaarddeviatie of standaardafwijking, een maat voor de spreiding ten opzichte van het gemiddelde. Ze geeft weer hoe sterk de scores van individuele respondenten van elkaar verschillen. Een kleine standaardafwijking duidt op eensgezindheid, een grotere standaardafwijking wijst erop dat respondenten eerder verdeeld zijn in hun oordeel.

4.3.1.3 Analyses op schaalniveau

Sommige theoretische concepten uit de conceptuele uitgangspunten werden bij het ontwikkelen van de vragenlijst geoperationaliseerd in verschillende items die samen een schaal beogen te vormen. Er wordt dus verondersteld dat deze items verschillende aspecten van hetzelfde achterliggende theoretisch concept meten. Het gebruik van schaalscores heeft als voordeel dat de betrouwbaarheid van het meetinstrument verhoogd wordt. Een schaal wordt verondersteld intern consistent te zijn. Om de interne consistentie van de schalen te berekenen, maken we gebruik van de Cronbachs alfa (α), een veelgebruikte maat voor interne

consistentie. In de literatuur wordt veelal vooropgesteld dat α minstens 0.60 dient te bedragen om de schaal als voldoende samenhangend te beschouwen; hoe dichter α bij waarde 1.00 ligt, hoe beter de interne consistentie van de schaal.

Voor elke respondent die minstens 75 procent van de items van de betreffende schaal heeft beantwoord (*M*, berekenen we een schaalscore.

Tabel 8 bevat het aantal items en de Cronbach's alpha-waarden van alle geconstrueerde schalen. Zoals we zien in de tabel zijn de Cronbach's alpha-waarden steeds groter dan 0.60. Zowel de Cronbach's alpha-waarden voor de schalen bij verwachte doelen, feedbackfactoren, schoolfactoren en contextfactoren bedragen telkens meer dan 0.80. We kunnen dan ook stellen dat deze schalen een goede interne consistentie vertonen. Ook uit de Cronbach's alpha-waarden voor de gehanteerde schalen bij gebruikersfactoren en gewenste effecten die meestal groter dan 0.80 en in vier gevallen groter dan 0.69 zijn, stellen we over het algemeen een goede interne consistentie vast.

Tabel 8
Schalen, het aantal items en de Cronbach's alpha-waarden

Schaal	Aantal items	α
Verwachte doelen		
- Verwachte doelen op leerlingniveau	10	0.92
- Verwachte doelen op leraarniveau	5	0.89
- Verwachte doelen op schoolniveau	10	0.92
Feedbackfactoren		
- Feedback op leerlingniveau	7	0.88
- Feedback op klasniveau	7	0.91
- Feedback op schoolniveau	8	0.93
Schoolfactoren		
- Ondersteunende relaties en samenwerking	3	0.80
- Gezamenlijke doelgerichtheid	4	0.93
- Transformationeel leiderschap	5	0.91
Contextfactoren		
- Externe verwachtingen	5	0.98

	Gebruikersfactoren		
-	Attitude: Cognitieve attitude t.a.v. feedbackgebruik	12	0.80
-	Attitude: Affectieve attitude t.a.v. feedbackgebruik	5	0.85
-	Motivatie: Autonome motivatie t.a.v. feedbackgebruik	6	0.85
-	Motivatie: Gecontroleerde motivatie t.a.v. feedbackgebruik	6	0.88
-	Motivatie: Amotivatie t.a.v. feedbackgebruik	3	0.69
-	Persoonlijke doelmatigheid: verkennen en overleggen t.a.v. feedbackgebruik	4	0.80
-	Persoonlijke doelmatigheid: interpreteren t.a.v. feedbackgebruik	4	0.84
-	Persoonlijke doelmatigheid: diagnose t.a.v. feedbackgebruik	2	0.76
-	Bekommernissen: Zelf-bekommernissen t.a.v. feedbackgebruik	5	0.81
-	Bekommernissen: Andere-bekommernissen t.a.v. feedbackgebruik	5	0.76
	Verwachte onwenselijke effecten		
-	Gewenste (onwenselijke) effecten op leerlingniveau	2	0.77
-	Gewenste (onwenselijke) effecten op leraarniveau	8	0.90
-	Gewenste (onwenselijke) effecten op schoolniveau	6	0.84
-	Gewenste (onwenselijke) effecten op onderwijssysteemniveau	6	0.88

4.3.2 Verschillen naar functie, onderwijsniveau en onderwijsnet

Aan de hand van ongepaarde t-testen en anova's (inclusief post-hoc-testen) analyseren we zowel op itemniveau als op schaalniveau de verschillen tussen groepen van respondenten. We kijken daarbij naar de functie waaraan respondenten de meeste tijd binnen de school besteden – schoolleiders (SL) en leraren (L), het onderwijsniveau – respondenten uit het lager onderwijs (LO) of het secundair onderwijs (SO) en het onderwijsnet – respondenten uit het vrij gesubsidieerd onderwijs (VG), het gemeenschapsonderwijs (G) of het officieel gesubsidieerd onderwijs (OG).

Op basis van de functie waaraan respondenten de meeste tijd besteden in de huidige school worden ze gegroepeerd in twee categorieën. Schoolleiders zijn de directeurs, alsook de beleidsondersteuners en zorgcoördinatoren. Tot de groep leraren behoren de leraren, de GOK- of SES-leraren en de zorgleraren.

We geven telkens het gemiddelde (*M*) en we rapporteren enkel over statistisch significante verschillen tussen de groepen met minimum een klein effect (d > 0.20 ; eta^2 > 0.01). We duiden hierbij de richting van het

verschil aan met symbolen. Daarbij staat ► voor 'statistisch significant groter dan' en ◄ voor 'statistisch significant kleiner dan'.

4.4 KENMERKEN VAN RESPONDENTEN EN HUN SCHOLEN

In wat volgt, beschrijven we kenmerken van respondenten en hun scholen voor zowel het lager als voor het secundair onderwijs..

4.4.1 Steekproef lager onderwijs

Tabel 9 geeft de verdeling van het aantal deelnemende lagere scholen uit de steekproef (groep 'schoolleider en leraren') over de verschillende netten weer ten opzichte van de verdeling binnen de populatie. We stellen vast dat 18 scholen uit onze steekproef van het lager onderwijs (of 75% van de scholen in de steekproef) aangesloten zijn bij de koepel van het vrij gesubsidieerd onderwijs. In de populatie binnen het lager onderwijs is het aandeel van het vrij gesubsidieerd onderwijs kleiner (59.7%). Bijgevolg zijn het gemeenschapsonderwijs en het officieel gesubsidieerd onderwijs ondervertegenwoordigd binnen de steekproef. Het aandeel van het gemeenschapsonderwijs is 8.3% binnen de steekproef ten opzichte van 18.2% binnen de populatie en het aandeel van het officieel gesubsidieerd onderwijs is 16.7% binnen de steekproef ten opzichte van 22.2% binnen de populatie.

Tabel 9
Verdeling scholen lager onderwijs naar onderwijsnet

	Steekproef		Populatie	
	n	%	N	%
Vrij gesubsidieerd onderwijs (VGO)	18	75.0	1401	59.7
Gemeenschapsonderwijs (GO)	2	8.3	426	18.2
Officieel gesubsidieerd onderwijs (OGO)	4	16.7	520	22.2
Totaal	24	100%	2347	100%

Om de verdeling volgens schoolgrootte in het lager onderwijs te rapporteren, werden op basis van het aantal leerlingen drie groepen scholen gemaakt, namelijk klein, middelgroot en groot. Hoewel de cesuren tussen deze groepen arbitrair is, baseerden we ons op de regelgeving die stelt dat een directeur van een

basisschool met minimum 180 ingeschreven leerlingen vrijgesteld is van elke onderwijsopdracht (Denis et al., 2018). Een kleine school telt bijgevolg minder dan 180 leerlingen. Een middelgrote school heeft tussen de 180 en 400 leerlingen en een grote school heeft meer dan 400 leerlingen. Tabel 10 geeft de verdeling van de scholen uit de steekproef (groep van 'schoolleider en leraren') op basis van schoolgrootte weer ten opzichte van deze verdeling binnen de populatie. We stellen vast dat 8 scholen uit onze steekproef van het lager onderwijs (of 33.3% van de scholen in de steekproef) grote scholen zijn. In de populatie binnen het lager onderwijs is het aandeel grote scholen kleiner (18.5%). Bijgevolg zijn middelgrote en kleine scholen ondervertegenwoordigd binnen de steekproef. Het aandeel middelgrote scholen is 58.3% binnen de steekproef ten opzichte van 64.0% binnen de populatie en het aandeel kleine scholen is 8.3% binnen de steekproef t.o.v. 17.6% binnen de populatie.

Tabel 10
Verdeling scholen lager onderwijs naar schoolgrootte

	Steekproef		Populatie	
	n	%	N	%
Grote school (≥400 leerlingen)	8	33.3	447	18.5
Middelgrote school (≥180 <400 leerlingen)	14	58.3	1549	64.0
Kleine school (<180 leerlingen)	2	8.3	425	17.6
Totaal	24	100%	2421	100%

Wanneer we kijken naar het geslacht van de respondenten (groep 'schoolleider en leraren' en groep 'schoolleiders') stellen we vast dat 88.3% van de respondenten uit het lager onderwijs vrouw is ten opzichte van 11.7% mannen (zie Tabel 11).

Tabel 11
Verdeling respondenten lager onderwijs naar geslacht

	n	%
Man	27	11.7
Vrouw	204	88.3
Totaal	231	100%

Wat betreft het opleidingsniveau van de respondenten (groep 'schoolleider en leraren' en groep 'schoolleiders') stellen we vast dat de meerderheid van de respondenten beschikt over een bachelorsdiploma (88.7%) (zie Tabel 12). 18% van de respondenten uit het lager onderwijs behaalde een masterdiploma en 3.5% een graduaatsdiploma.

Tabel 12
Verdeling respondenten lager onderwijs naar opleidingsniveau

	n	%
Secundair onderwijs	0	0
Bachelor	204	88.7
Graduaat	8	3.5
(Educatieve) master	18	7.8
Doctoraat	0	0
Totaal	230	100%

Daarnaast werden respondenten uit het lager onderwijs (groep 'schoolleider en leraren' en groep 'schoolleiders') gevraagd naar de functie waaraan zij binnen de school de meeste tijd besteden. Op basis van Tabel 13 kunnen we vaststellen dat de meerderheid van de respondenten leraar is (65.4%). Daarnaast is 19.9% van de respondenten directeur en 0.9% beleidsondersteuner. 13% van de respondenten neemt binnen de school een zorgfunctie op (zorg- of SES-leerkracht of –coördinator).

Op basis van de functie waaraan respondenten (groep 'schoolleider en leraren' en groep 'schoolleiders') de meeste tijd besteden in de huidige school groeperen we hen in twee categorieën (zie 4.4.1.4). Deze tweedeling naargelang 'functie' wordt eveneens weergegeven in Tabel 13. Tot de groep 'schoolleiders' behoren de schoolleiders, alsook de beleidsondersteuners en zorgcoördinatoren. Tot de groep 'leraren' behoren de leraren, de GOK- of SES-leraren en de zorgleraren. We zien in deze tabel dat 29.3% van de respondenten 'schoolleiders' zijn en dat 70.8% van de respondenten tot de groep van 'leraren' behoort.

Tabel 13
Verdeling respondenten lager onderwijs naar functie

	n	%
Groep 'Leraren'	162	70.8
Leraar kleuterklas	10	4.3
Leraar 1ste leerjaar	18	7.8
Leraar 2de leerjaar	20	8.6
Leraar 3de leerjaar	22	9.5
Leraar 4de leerjaar	26	11.3
Leraar 5de leerjaar	28	12.1
Leraar 6de leerjaar	27	11.7
GOK- of SES-leraar, zorgleraar	11	4.8
Groep 'Schoolleiders'	67	29.3
Directeur	46	19.9
Beleidsondersteuner	2	0.9
Zorgcoördinator	19	8.2
Andere	2	0.9
Totaal	231	100%

Tot slot peilden we naar de ervaring van de respondenten (groep 'schoolleider en leraren' en groep 'schoolleiders') uit het lager onderwijs in hun functie (zie Tabel 14). We stellen vast dat ongeveer een vijfde van de respondenten (20.4%) vijf jaar of minder lang tewerk gesteld is en een ander vijfde (20.8%) zes tot 10 jaar. Een derde van het aantal respondenten (33.3%) oefent de functie van waaruit ze de vragenlijst beantwoorden elf tot twintig jaar uit. Ongeveer een vierde van het aantal respondenten (25.6%) oefent de huidige functie reeds meer dan twintig jaar uit.

Tabel 14
Verdeling respondenten lager onderwijs naar ervaring

	n	%
0 – 1 jaar	21	9.1
2 – 5 jaar	26	11.3

6 – 10 jaar	48	20.8
11 – 20 jaar	77	33.3
21 – 30 jaar	39	16.9
> 30 jaar	20	8.7
Totaal	231	100%

4.4.2 Steekproef secundair onderwijs

Tabel 15 geeft de verdeling van secundaire scholen uit de steekproef (groep 'schoolleider en leraren') over de verschillende netten weer. In onze steekproef van het secundair onderwijs zijn 21 scholen (of 75% van de scholen in de steekproef) aangesloten bij de koepel van het vrij gesubsidieerd onderwijs. Wanneer we dit vergelijken met de populatie stellen we vast dat het aandeel secundaire scholen dat aangesloten is bij de koepel van het vrij gesubsidieerd onderwijs in werkelijkheid 71.1% is. Het aandeel scholen van het vrij gesubsidieerd onderwijs is in onze steekproef van het secundair onderwijs dus iets groter. Wat betreft het gemeenschapsonderwijs stellen we vast dat 4 scholen (of 14.3% van de secundaire scholen) uit onze steekproef aangesloten zijn bij deze koepel. Deze koepel is in het secundair onderwijs in werkelijkheid groter. 22.2% van de secundaire scholen uit de populatie is immers aangesloten bij deze koepel. Tet slot zijn 3 secundaire scholen uit onze steekproef aangesloten bij de koepel van het officieel gesubsidieerd onderwijs, wat neerkomt op 10.7%. Deze koepel is in de steekproef groter vermits in werkelijkheid 6.7% van de secundaire scholen aangesloten is bij de koepel van het officieel gesubsidieerd onderwijs.

Tabel 15
Verdeling scholen secundair onderwijs naar onderwijsnet

	Steekproef		Populatie	
	n	%	N	%
Vrij gesubsidieerd onderwijs (VGO)	21	75.0	684	71.1
Gemeenschapsonderwijs (GO)	4	14.3	214	22.2
Officieel gesubsidieerd onderwijs (OGO)	3	10.7	64	6.7
Totaal	28	100%	962	100%

Om de verdeling volgens schoolgrootte in het secundair onderwijs (groep 'schoolleider en leraren') te rapporteren baseren we ons op dezelfde groepen dan die voor het lager onderwijs. Hoewel de cesuren tussen deze onderwijstypes eigenlijk zouden moeten variëren, hanteren we gemakshalve toch een uniforme groepering. Een kleine school telt minder dan 180 leerlingen, een middelgrote school heeft tussen de 180 en 400 leerlingen en een grote school heeft meer dan 400 leerlingen. Zo blijkt uit Tabel 16 dat het aantal grote scholen (64.3% van de secundaire scholen) groter is dan in werkelijkheid en dat het aantal middelgrote scholen (25%) en het aantal kleine scholen (10.7%) kleiner is dan in de populatie.

Tabel 16
Verdeling scholen secundair onderwijs naar schoolgrootte

	Steekproef		Populatie	
	N	%	N	%
Grote school (≥400 leerlingen)	18	64.3	495	52.1
Middelgrote school (≥180 <400 leerlingen)	7	25	319	33.6
Kleine school (<180 leerlingen)	3	10.7	136	14.3
Totaal	28	100%	950	100%

Net als in de steekproef van het lager onderwijs zijn in de steekproef van het secundair onderwijs (groep 'schoolleider en leraren' en groep 'schoolleiders') vrouwen het sterkst vertegenwoordigd (zie Tabel 17). Het verschil tussen beide categorieën is echter minder groot dan in de steekproef van het lager onderwijs. Het aandeel mannelijke respondenten in deze steekproef bedraagt 32.2%, het aandeel vrouwelijke respondenten 67.4%.

Tabel 17
Verdeling respondenten secundair onderwijs naar geslacht

	n	%
Man	182	32.2
Vrouw	381	67.4
Andere	2	0.4
Totaal	565	100%

Tabel 18 geeft het opleidingsniveau van de respondenten uit het secundair onderwijs (groep 'schoolleider en leraren' en groep 'schoolleiders') weer. De grootste groep respondenten (49.7%) heeft een bachelordiploma. Daarnaast heeft 41.6% van de respondenten uit het secundair onderwijs een masterdiploma.

Tabel 18
Verdeling respondenten secundair onderwijs naar opleidingsniveau

	n	%
Secundair onderwijs	10	1.8
Bachelor	281	49.7
Graduaat	22	3.9
(Educatieve) master	235	41.6
Doctoraat	17	3
Totaal	565	100%

Respondenten uit het secundair onderwijs (groep 'schoolleider en leraren' en groep 'schoolleiders') werden ook gevraagd naar de functie waaraan zij momenteel binnen de school de meeste tijd besteden (zie Tabel 19). De grootste groep van de respondenten uit het secundair onderwijs die de bevraging invulden is leraar (81.5%). Daarnaast voert 17.5% van de respondenten een beleidsfunctie uit, zoals directeur, beleidsondersteuner of GOK-coördinator. Verder vinden we in de steekproef 1 of 0.2% GOK-, SES- of zorgleraar terug. Tot slot geven 5 respondenten aan een 'andere' functie op te nemen, wat neerkomt op een aandeel van 0.9%. Op basis van de functie waaraan respondenten (groep 'schoolleider en leraren' en groep 'schoolleiders') de meeste tijd besteden in de huidige school groeperen we hen in twee categorieën. Deze tweedeling naargelang 'hoofdopdracht' wordt eveneens weergegeven in Tabel 19. Schoolleiders zijn de directeurs, alsook de beleidsondersteuners en zorgcoördinatoren en vertegenwoordigen 17.7% van het aantal respondenten. Tot de groep 'Leraren' behoren de leraren, de GOK- of SES-leraren en de zorgleraren. Zij vertegenwoordigen 82.3% van het aantal respondenten uit het secundair onderwijs.

Tabel 19
Verdeling respondenten secundair onderwijs naar functie

	n	%
Groep 'Leraren'	461	82.3
Leraar 1ste graad	214	37.9
Leraar 2de graad	110	19.5
Leraar 3de graad	136	24.1
GOK- of SES-leraar, zorgleraar	1	0.2
Groep 'Schoolleiders'	99	17.7
Directeur	75	13.3
Beleidsondersteuner	13	2.3
Zorgcoördinator	11	1.9
Andere	5	0.9
Totaal	565	100%

Leraren werden gevraagd te specificeren welke vakken het grootste deel van hun lesopdracht op school uitmaakt (zie Tabel 20). Het grootste aantal respondenten (29.3%) geeft voornamelijk een ander algemeen vak dan Nederlands of wiskunde, 25.9% van de respondenten geeft voornamelijk Nederlands en 24.8% van de respondenten geeft voornamelijk wiskunde. 6.8% van de respondenten geeft voornamelijk een technisch of praktisch vak, 2.2% voornamelijk een kunstvak en 11.1% voornamelijk een ander – niet nader gespecifieerd – vak.

Tabel 20
Verdeling respondenten secundair onderwijs naar vakken die het grootste deel van de lesopdracht uitmaken

	n	%
AV Nederlands	119	25.9
AV Wiskunde	114	24.8
Ander AV dan Nederlands/wiskunde	135	29.3
Praktisch vak (PV)	16	3.5
Technisch vak (TV)	15	3.3
Kunstvak (KV)	10	2.2

Andere	51	11.1
Totaal	460	100%

We vroegen leraren uit het secundair onderwijs ook in welke onderwijsvorm zij binnen de school het meest actief zijn. Deze resultaten worden weergegeven in Tabel 21. We stellen vast dat het grootste aantal respondenten lesgeeft in de A-stroom van de eerste graad (41.3%) en in het ASO (36.3%). 13.9% van de respondenten geeft les in het TSO en 4.8% in het BSO.

Tabel 21
Verdeling leraren uit het secundair onderwijs naar onderwijsvorm

	n	%
1ste graad A-stroom	190	41.3
1ste graad B-stroom	17	3.7
2de/3de graad ASO/finaliteit doorstroom	167	36.3
2de/3de graad BSO/finaliteit arbeidsmarkt	22	4.8
2de/3de graad KSO/finaliteit doorstroom	0	0
2de/3de graad KSO/dubbele finaliteit	0	0
2de/3de graad TSO/finaliteit doorstroom	29	6.3
2de/3de graad TSO/dubbele finaliteit	35	7.6
Totaal	460	100%

Tot slot peilden we ook in de steekproef van het secundair onderwijs (groep 'schoolleider en leraren' en groep 'schoolleiders') naar de ervaring van de respondenten in de functie waaraan ze de meeste tijd besteden (zie Tabel 22). We kunnen vaststellen dat het grootste aantal respondenten elf tot twintig jaar ervaring heeft (33.3%). Ongeveer een vijfde van het aantal respondenten heeft twee tot vijf jaar ervaring (20.9%) of zes tot tien jaar ervaring (18.9%). Een kwart van het aantal respondenten (25.3%) heeft meer dan twintig jaar ervaring in de functie van waaruit ze de vragenlijst invulden.

Tabel 22
Verdeling respondenten uit het secundair onderwijs naar ervaring

	n	%
0 – 1 jaar	9	1.6
2 – 5 jaar	118	20.9
6 – 10 jaar	107	18.9
11 – 20 jaar	188	33.3
21 – 30 jaar	94	16.6
> 30 jaar	49	8.7
Totaal	565	100%

5 RESULTATEN

Om een Vlaanderen-breed inzicht te verwerven in verwachtingen bij en randvoorwaarden voor feedbackgebruik uit centrale toetsen voerden we een web-based surveyonderzoek uit bij aselecte gestratificeerde representatieve steekproef. In dit luik beschrijven we hiervan de resultaten.

5.1 PERCEPTIE VAN RELEVANTIE EN BETEKENISVOLHEID VAN FEEDBACK UIT CENTRALE TOETSEN EN VAN FAIRE EN VALIDE VERGELIJKINGEN

Om meer inzicht te verwerven in het potentieel van feedbackgebruik uit centrale toetsen is het belangrijk na te gaan wat feedback uit centrale toetsen voor schoolleiders en leraren relevant, betekenisvol en vanuit hun perspectief betrouwbaar en valide maakt.

We beschrijven in dit luik daarom achtereenvolgens welke informatie in de feedback uit centrale toetsen tegemoet kan komen aan de noden van schoolleiders en leraren en welke informatie die handvaten biedt voor betekenisgeving ontsloten kan worden (5.1.1) en met welke leerling- en schoolkenmerken rekening dient gehouden te worden in vergelijkingen van de toetsresultaten van de school zodat deze vergelijkingen als fair – en dus als valide en betrouwbaar – gepercipieerd worden (5.1.2).

5.1.1 Perceptie van relevantie en betekenisvolheid van feedback uit centrale toetsen

In de vragenlijst peilden we bij schoolleiders en leraren naar de (soorten) informatie waarover de feedback uit centrale toetsen dient te rapporteren. De aangeboden stellingen peilden zowel naar (soorten) informatie die de feedback relevant maakt, als naar (soorten) informatie die de feedback betekenisvol maakt.

Schoolleiders en leraren geven in het algemeen aan dat de feedback uit centrale toetsen zowel informatie op leerlingniveau (gem = 4.0, SD = 0.7) als – weze het in iets mindere mate – op schoolniveau (gem = 3.9, SD = 0.8) en klasniveau (gem = 3.8, SD = 0.8) dient te bevatten (zie Tabel 23).

Op leerlingniveau geven schoolleiders en leraren aan het eerder eens te zijn met de stelling dat feedback over individuele toetsresultaten (op naam) (gem = 4.0, SD = 1.0) en over het soort toetsvragen dat door elke leerling (op naam) (gem = 4.0, SD = 0.9) al dan niet correct beantwoord wordt, bijdraagt aan de relevantie van de feedback. Daarnaast zijn schoolleiders en leraren het helemaal eens dat rapporteren over de mate waarin individuele leerlingen (op naam) de getoetste eindtermen bereiken (gem = 4.3, SD = 0.8) en over de leerwinst die individuele leerlingen (op naam) boeken (gem = 4.1, SD = 0.9) de feedback op leerlingniveau betekenisvol maakt. Tevens zijn ze matig positief over het rapporteren van de mate waarin elke individuele leerling (op naam) 'sterk-gemiddeld-zwak' presteert (gem = 4.0, SD = 1.0), van de individuele toetsresultaten (op naam) van leerlingen in vergelijking met de toetsresultaten van hun medeleerlingen in dezelfde klas(groep) (gem = 3.6, SD = 1.2) en van de individuele toetsresultaten (op naam) van leerlingen in vergelijking met toetsresultaten van andere Vlaamse leerlingen met vergelijkbare achtergrondkenmerken (gem = 3.8, SD = 1.1).

Verder beoordelen schoolleiders en leraren het belang van de feedback om te rapporteren over informatie op klasniveau als gematigd positief. Toch is – in vergelijking met het belang van informatie op leerlingniveau – de inschatting van schoolleiders en leraren hier iets minder positief. Zo beoordelen schoolleiders en leraren de stellingen dat het rapporteren van gemiddelde toetsresultaten voor de verschillende klassen in de school (gem = 3.7, SD = 1.1) en van het soort toetsvragen dat door de leerlingen op klasniveau al dan niet correct werd beantwoord (gem = 4.0, SD = 0.9) als eerder wel relevant. Daarnaast zijn ze het eveneens eerder eens dat informatie over de verdeling 'sterk-gemiddeld-zwak' presteren van de leerlingen in elke klas (gem = 3.7, SD = 1.0), over de mate waarin klassen van de school de getoetste eindtermen bereiken (gem = 4.0, SD = 0.9), over de toetsresultaten van elke klas in vergelijking met toetsresultaten van andere vergelijkbare klassen in het Vlaamse onderwijs (gem = 3.7, SD = 1.1) en over de leerwinst op klasniveau (gem = 3.9, SD =

1.0) bijdraagt aan de betekenisvolheid van de feedback. Over het rapporteren over de toetsresultaten van elke klas in vergelijking met de toetsresultaten van de andere klassen op de school (gem = 3.4, SD = 1.2) maken schoolleiders en leraren een eerder neutrale inschatting.

De items die de relevantie en betekenisvolheid van de feedback op schoolniveau in kaart brengen, worden gematigd positief tot positief beoordeeld. Zo zijn schoolleiders en leraren het er eerder mee eens dat het rapporteren over de gemiddelde toetsresultaten op schoolniveau van de leerlingen in de school (gem = 3.9, SD = 1.0) en over het soort toetsvragen dat door de leerlingen op schoolniveau al dan niet correct beantwoord werd (gem = 3.9, SD = 0.9) relevant is. Men is het ook 'eerder eens' met de stellingen dat het informeren over de verdeling 'sterk-gemiddeld-zwak'-presteren van al de leerlingen van de school (gem = 3.8, SD = 1.0), over de toetsresultaten van al de leerlingen in de school in vergelijking met de toetsresultaten in alle andere Vlaamse scholen (gem = 3.7, SD = 1.1), over de toetsresultaten van de school in vergelijking met de toetsresultaten van andere vergelijkbare Vlaamse scholen (gem = 3.9, SD = 1.0) en over de leerwinst van de leerlingen op schoolniveau (gem = 4.0, SD = 0.9) de feedback betekenisvol maakt.

Wanneer we de verschillen naar functie onderzoeken, stellen we talrijke statistisch significante verschillen tussen schoolleiders en leraren vast. Het aantal verschillen neemt toe van leerling-, over klas-, naar schoolniveau. Daarenboven beoordelen schoolleiders de betreffende stellingen steeds positiever dan leraren. De verschillen zijn in omvang echter beperkt.

Als we de schaalscoresals referentie gebruiken, zien we dat schoolleiders meer informatiebehoeften dan leraren op klasniveau [gem SL = 4.0, gem L = 3.7 ; $t(717.0) = 3.5$, $p < 0.01$, $d = 0.33$] en op schoolniveau [gem SL = 4.1, gem L = 3.8 ; $t(719.0) = 4.6$, $p < 0.01$, $d = 0.44$] rapporteren.

Op leerlingniveau zijn schoolleiders het 'helemaal eens' dat het rapporteren over de individuele toetsresultaten (op naam) van de eigen leerlingen in vergelijking met toetsresultaten van andere Vlaamse leerlingen met vergelijkbare achtergrondkenmerken de feedback betekenisvol maakt. Leraren zijn het slechts 'eerder eens' met deze stelling [gem SL = 4.1, gem L = 3.8 ; $t(724.0) = 2.7$, $p < 0.01$, $d = 0.25$].

Op klasniveau zijn schoolleiders het er helemaal mee eens dat informatie over het soort toetsvragen dat door de leerlingen op klasniveau al dan niet correct beantwoord werd, de feedback relevant maakt. Leraren

zijn het hier in iets mindere mate mee eens [gem SL = 4.1, gem L = 3.9 ; $t(720.0) = 2.4$, $p < 0.01$, $d = 0.22$]. Tevens maken schoolleiders een meer positieve inschatting dan leraren van hun behoefte aan informatie over de mate waarin klassen van de school de getoetste eindtermen bereiken [gem SL = 4.3, gem L = 4.0 ; $t(723.0) = 3.9$, $p < 0.01$, $d = 0.38$]. Verder zijn schoolleiders het er in meerdere mate dan leraren mee eens dat de feedback moet informeren over de toetsresultaten van elke klas in vergelijking met de toetsresultaten van de andere klassen op de school [gem SL = 3.7, gem L = 3.3 ; $t(240.0) = 3.2$, $p < 0.01$, $d = 0.29$], over de toetsresultaten van elke klas in vergelijking met toetsresultaten van andere vergelijkbare klassen in het Vlaamse onderwijs [gem SL = 4.0, gem L = 3.6 ; $t(723.0) = 3.2$, $p < 0.01$, $d = 0.31$] en over de leerwinst van de leerlingen op klasniveau [gem SL = 4.1, gem L = 3.8 ; $t(715.0) = 2.7$, $p < 0.01$, $d = 0.25$].

Ook wat de stellingen over feedback op schoolniveau betreft beoordelen schoolleiders elke stelling positiever dan leraren. Schoolleiders ervaren statistisch significant meer informatiebehoeften dan leraren met betrekking tot de gemiddelde toetsresultaten van al de leerlingen in de school [gem SL = 4.1, gem L = 3.9 ; $t(720.0) = 2.8$, $p < 0.01$, $d = 0.26$], het soort toetsvragen dat door de leerlingen van de school al dan niet correct wordt beantwoord [gem SL = 4.1, gem L = 3.8 ; $t(718.0) = 2.7$, $p < 0.01$, $d = 0.25$], de verdeling 'sterk-gemiddeld-zwak' presteren van al de leerlingen van de school [gem SL = 4.0, gem L = 3.7 ; $t(720.0) = 2.9$, $p < 0.01$, $d = 0.27$], het aantal leerlingen dat in de school de getoetste eindtermen bereikt [gem SL = 4.4, gem L = 4.1 ; $t(723.0) = 4.3$, $p < 0.01$, $d = 0.41$], de toetsresultaten van al de leerlingen in de school in vergelijking met de toetsresultaten in alle andere Vlaamse scholen [gem SL = 4.0, gem L = 3.6 ; $t(719.0) = 3.4$, $p < 0.01$, $d = 0.32$], de toetsresultaten van de school in vergelijking met de toetsresultaten van andere vergelijkbare Vlaamse scholen [gem SL = 4.2, gem L = 3.8 ; $t(720.0) = 4.6$, $p < 0.01$, $d = 0.44$], de leerwinst van de leerlingen van de school [gem SL = 4.2, gem L = 3.9 ; $t(714.0) = 3.8$, $p < 0.01$, $d = 0.36$] en de mate waarin de school meer of minder bijdraagt aan de leerwinst van leerlingen dan andere scholen [gem SL = 4.1, gem L = 3.7 ; $t(715.0) = 4.6$, $p < 0.01$, $d = 0.43$]. Voor schoolleiders draagt deze informatie met andere woorden in sterkere mate bij aan de relevantie en betekenisvolheid van de feedback dan voor leraren.

Met betrekking tot onderwijsniveau en onderwijsnet stellen we geen statistisch significante verschillen met ten minste een klein effect vast. Het behoren tot een bepaald onderwijsniveau of onderwijsnet oefent met andere woorden geen of slechts een verwaarloosbaar effect uit op het ervaren van informatiebehoeften die feedback relevant of betekenisvol maakt.

Tabel 23
Relevantie en betekenisvolheid van feedback uit centrale toetsen op leerling-, klas- en schoolniveau

	N WN/ NVT	Gem (SD)	T-test Gem SL	L	LO	SO	Anova Gem VG	G	OG
Feedback op leerlingniveau									
Feedback op leerlingniveau moet mij informeren over:									
de individuele toetsresultaten (op naam) van onze leerlingen op de centrale toetsen	724 1.9%	4.0 (1.0)	4.0	4.0	4.0	4.0	4.0	4.2	4.2
het soort toetsvragen dat door elke leerling (op naam) al dan niet correct beantwoord werd	719 2.6%	4.0 (0.9)	4.0	4.0	4.0	4.0	4.0	3.9	4.1
de mate waarin elke individuele leerling (op naam) 'sterk-gemiddeld-zwak' presteert	730 1.0%	4.0 (1.0)	3.9	4.0	4.0	4.0	4.0	4.0	4.1
de mate waarin individuele leerlingen (op naam) de getoetste eindtermen bereiken	728 1.4%	4.3 (0.8)	4.3	4.3	4.3	4.3	4.3	4.4	4.4
de individuele toetsresultaten (op naam) van onze leerlingen in vergelijking met de toetsresultaten van hun medeleerlingen in dezelfde klas(groep)	730 1.1%	3.6 (1.2)	3.5	3.6	3.6	3.6	3.5	3.6	3.7
de individuele toetsresultaten (op naam) van onze leerlingen in vergelijking met toetsresultaten van andere Vlaamse leerlingen met vergelijkbare achtergrondkenmerken	731 0.9%	3.8 1.1	4.1 ►	3.8	3.9	3.8	3.8	3.6	3.8
de leerwinst die individuele leerlingen (op naam) boeken	727 1.5%	4.1 (0.9)	4.2	4.1	4.0	4.1	4.1	4.1	4.3
Schaal: Feedback op leerlingniveau **(α = 0.88)**	726 0.8%	4.0 (0.7)	4.0	4.0	4.0	4.0	4.0	4.0	4.1
Feedback op klasniveau:									
Feedback op klasniveau moet mij informeren over:									
gemiddelde toetsresultaten voor de verschillende klassen in de school	726 1.6%	3.7 (1.1)	3.8	3.7	3.6	3.7	3.7	3.8	3.7
het soort toetsvragen dat door de leerlingen op klasniveau al dan niet correct beantwoord werd	727 1.5%	4.0 (0.9)	4.1 ►	3.9	4.0	3.9	4.0	3.8	3.9
de verdeling 'sterk-gemiddeld-zwak'-presteren van de leerlingen in elke klas	728 1.4%	3.7 (1.0)	3.8	3.7	3.7	3.7	3.8	3.7	3.7
de mate waarin klassen van de school de getoetste eindtermen bereiken	730 1.1%	4.0 (0.9)	4.3 ►	4.0	4.0	4.1	4.1	4.0	3.9
de toetsresultaten van elke klas in vergelijking met de toetsresultaten van de andere klassen op de school	724 1.9%	3.4 (1.2)	3.7 ►	3.3	3.3	3.4	3.4	3.4	3.4

de toetsresultaten van elke klas in vergelijking met toetsresultaten van andere vergelijkbare klassen in het Vlaamse onderwijs	730 1.1%	3.7 (1.1)	4.0	►	3.6	3.8		3.7	3.7	3.5	3.6	
de leerwinst van de leerlingen op klasniveau	722 2.2%	3.9 (1.0)	4.1	►	3.8	3.9		3.9	3.9	3.8	4.0	
Schaal: Feedback op klasniveau **(α = 0.91)**	724 0.8%	3.8 (0.8)	4.0	►	3.7	3.8		3.8	3.8	3.7	3.7	
Feedback op schoolniveau												
Feedback op schoolniveau moet mij informeren over:												
gemiddelde toetsresultaten op schoolniveau van al de leerlingen in de school.	727 1.5%	3.9 (1.0)	4.1	►	3.9	3.9		3.9	3.9	4.0	3.9	
het soort toetsvragen dat door de leerlingen op schoolniveau al dan niet correct beantwoord werd	725 1.8%	3.9 (0.9)	4.1	►	3.8	3.9		3.9	3.9	3.7	3.9	
de verdeling 'sterk-gemiddeld-zwak'-presteren van al de leerlingen van de school	727 1.5%	3.8 (1.0)	4.0	►	3.7	3.7		3.8	3.8	3.3	3.6	
het aantal leerlingen dat in de school de getoetste eindtermen bereikt	730 1.1%	4.2 (0.8)	4.4	►	4.1	4.2		4.2	4.2	4.1	4.2	
de toetsresultaten van al de leerlingen in de school in vergelijking met de toetsresultaten in alle andere Vlaamse scholen	726 1.6%	3.7 (1.1)	4.0	►	3.6	3.7		3.7	3.7	3.5	3.6	
de toetsresultaten van de school in vergelijking met de toetsresultaten van andere vergelijkbare Vlaamse scholen	727 1.5%	3.9 (1.0)	4.2	►	3.8	3.9		3.8	3.9	3.6	3.9	
de leerwinst van de leerlingen op schoolniveau.	721 2.3%	4.0 (0.9)	4.2	►	3.9	4.0		4.0	4.0	3.9	4.1	
de mate waarin deze school meer of minder bijdraagt aan de leerwinst van leerlingen dan andere scholen	722 2.2%	3.8 (1.0)	4.1	►	3.7	3.8		3.8	3.8	3.7	3.7	
Schaal: Feedback op schoolniveau **(α = 0.93)**	726 1.0%	3.9 (0.8)	4.1	►	3.8	3.9		3.9	3.9	3.8	3.9	

Opmerking: Gemiddelde op basis van vijfpuntenschaal van (1) 'helemaal oneens' over 'eerder oneens', 'noch eens, noch oneens', 'eerder eens' tot (5) 'helemaal eens'.

5.1.2 Perceptie van faire en valide vergelijkingen

Scholen verschillen wat hun leerlingenpubliek betreft. De toetsresultaten van scholen op centrale toetsen zonder meer vergelijken met het Vlaamse gemiddelde kan in die zin 'onfair' lijken. We bevroegen daarom of schoolleiders en leraren het wenselijk vinden om te weten hoe toetsresultaten van de school zich verhouden tot toetsresultaten van vergelijkbare scholen (zie Tabel 24). Gemiddeld genomen geven

schoolleiders en leraren aan het eerder wel wenselijk te vinden om te weten hoe toetsresultaten van de school zich verhouden tot toetsresultaten van vergelijkbare scholen (gem = 3.7, SD = 1.0). Schoolleiders en leraren verschillen hier echter statistisch significant van mening. Schoolleiders zijn positief ten aanzien van dergelijke vergelijking, terwijl leraren hierover slechts gematigd positief zijn [gem SL = 4.0, gem L = 3.7 ; $t(588.0) = 3.4$, $p < 0.01$, $d = 0.34$].

Tabel 24
Wenselijkheid van het vergelijken van de toetsresultaten van de school met de toetsresultaten van vergelijkbare scholen

	N WN/ NVT	Gem (SD)	T-test Gem				Anova Gem		
			SL	L	LO	SO	VG	G	OG
Ik vind het wenselijk om te weten hoe de toetsresultaten van mijn school zich verhouden tot de toetsresultaten van vergelijkbare Vlaamse scholen.	594 12%	3.7 (1.0)	4.0 ►	3.7	3.8	3.7	3.7	3.5	3.7

Opmerking: Gemiddelde op basis van vijfpuntenschaal van (1) 'helemaal oneens' over 'eerder oneens', 'noch eens, noch oneens', 'eerder eens' tot (5) 'helemaal eens'.

Om gepercipieerd faire en valide schoolvergelijkingen tot stand te brengen, bevroegen we schoolleiders eveneens over de mate waarin er volgens hen rekening dient gehouden te worden met leerling- en schoolkenmerken wanneer de toetsresultaten van de school op centrale toetsen vergeleken worden met die van andere scholen (zie Tabel 25).

Wat betreft leerlingenkenmerken geven schoolleiders aan het helemaal eens te zijn rekening te houden met de thuistaal van de leerlingen (gem = 4.4, gem = 0.7) en met de aanwezigheid van leerstoornissen bij de leerlingen (gem = 4.4, gem = 0.8). Schoolleiders beoordelen het opnemen van het opleidingsniveau van de ouder(s) van de leerling (gem = 4.0, SD = 0.9), de aanwezigheid van ouderlijke steun voor de leerlingen (gem = 4.0, SD = 0.9) en de mate waarin de leerling leergemotiveerd is (gem = 4.0, SD = 0.9) in het eerlijk vergelijken van schoolresultaten als matig positief. Ook met het rekening houden met eerdere toetsresultaten (gem = 3.9, SD = 1.0) en in mindere mate met het intelligentieniveau van de leerlingen (gem = 3.6, SD = 1.0) zijn schoolleiders het eerder eens. Op de stelling om in eerlijke vergelijkingen van de toetsresultaten van de school rekening te houden met het gezinsinkomen van de ouder(s) van de leerlingen (gem = 3.4) antwoorden schoolleiders neutraal.

Daarnaast geven schoolleiders – weliswaar in iets mindere mate – eveneens aan het globaal eerder eens te zijn om ook bij de vergelijkingen van toetsresultaten op centrale toetsen tussen scholen rekening te houden met schoolkenmerken. Schoolleiders beoordelen het rekening houden met kansarme/kansrijke buurt van de school (gem = 4.1, SD = 0.9) in eerlijke schoolvergelijkingen als positief. Over het rekening houden met de stedelijke/landelijke context van de school (gem = 4.0, SD = 0.9), met de mate waarin de school zich professionaliseert (gem = 3.9, SD = 0.9), de mate waarin het schoolteam samenwerkt (gem = 3.7, SD = 1.0), de ervaring van het schoolteam (gem = 3.7, SD = 0.9) en het personeelsverloop binnen het schoolteam (gem = 3.5, SD = 1.0) zijn schoolleiders gematigd positief. Ook op de vraag aan schoolleiders uit het secundair onderwijs of zij wensen dat er in eerlijke schoolvergelijkingen rekening gehouden wordt met het onderwijsaanbod op de school geven ze aan het eerder eens te zijn (gem = 3.8, SD = 1.1). De stellingen die peilen naar het rekening houden met het totale aantal leerlingen in de school (gem = 3.2, SD = 1.2), het opleidingsniveau van het lerarenteam van de school (gem = 3.3, SD = 1.1) en het resultaat van de laatste doorlichting (gem = 3.2, SD = 1.2) in faire schoolvergelijkingen beoordelen schoolleiders als algemeen neutraal.

Als we de verschillen naar onderwijsniveau bekijken, stellen we enkele verschillen vast tussen schoolleiders uit het lager onderwijs en schoolleiders uit het secundair onderwijs. Schoolleiders uit het lager onderwijs geven – vanuit een positieve inschatting – in meerdere mate dan respondenten uit het secundair onderwijs aan het belangrijk te vinden in eerlijke schoolvergelijkingen rekening te houden met de aanwezigheid van leerstoornissen [gem LO = 4.6, gem SO = 4.3 ; $t(96.8) = 1.9$, $p = 0.05$, $d = 0.39$]. Dit verschil betreft slechts een klein verschil. Het verschil tussen schoolleiders uit het lager onderwijs en schoolleiders uit het secundair onderwijs met betrekking tot de beoordeling om in dergelijke vergelijkingen ook rekening te houden met het resultaat van de laatste doorlichting is een verschil met een medium effect. Schoolleiders uit het lager onderwijs zijn het in dit geval – vanuit een neutrale inschatting – meer eens dan schoolleiders uit het secundair onderwijs om ook met dit schoolkenmerk rekening te houden [gem LO = 3.4, gem SO = 2.9 ; $t(99.0) = 2.4$, $p = 0.02$, $d = 0.50$].

Als we verschillen naar onderwijsnet beschouwen, stellen we vast dat schoolleiders uit het vrij gesubsidieerd onderwijs in meerdere mate aangeven het eens te zijn dan schoolleiders uit het officieel gesubsidieerd onderwijs om rekening te houden met de thuistaal van de leerlingen [gem VG = 4.4, gem OG = 3.5 ; $F(3.0,18.4) = 3.3$, $p = 0.05$, $eta^2 = 0.14$].

Tabel 25
Elementen om in rekening te brengen om tot faire vergelijking van toetsresultaten te komen

	N WN/N VT	Gem (SD)	T-test Gem LO		SO	Anova Gem VG	G	OG	
In faire vergelijkingen van de toetsresultaten van de school, dient rekening gehouden te worden met:									
de thuistaal van de leerlingen	101 0.0%	4.4 (0.7)	4.4		4.4	4.4	4.3	3.5	VG►OG
de aanwezigheid van leerstoornissen bij de leerlingen	101 0.0%	4.4 (0.8)	4.6	►	4.3	4.4	4.0	4.5	
de kansarme/kansrijke buurt van de school	101 0.0%	4.1 (0.9)	4.1		4.1	4.1	3.7	4.3	
de leermotivatie van de leerlingen	101 0.0%	4.0 (0.9)	4.1		3.9	3.9	3.3	3.3	
de aanwezigheid van ouderlijke steun voor de leerlingen	101 0.0%	4.0 (0.9)	4.1		4.0	4.0	3.0	3.3	
het opleidingsniveau van de ouder(s) van de leerlingen	101 0.0%	4.0 (0.9)	3.9		4.1	3.9	3.7	3.5	
de stedelijke/landelijke context van de school	101 0.0%	4.0 (0.9)	3.9		4.0	3.9	3.7	3.5	
de eerdere toetsresultaten van de leerlingen	100 1.0%	3.9 (1.0)	3.9		3.8	3.9	4.0	3.0	
de mate waarin het schoolteam zich professionaliseert	101 0.0%	3.9 (1.0)	4.0		3.8	3.9	3.3	4.0	
het onderwijsaanbod van de school (*)	65 0.0%	3.8 (1.1)				3.8	3.7	2.5	
de mate waarin het schoolteam samenwerkt	101 0.0%	3.7 (1.0)	3.9		3.6	3.7	3.3	4.0	
de ervaring van het schoolteam	101 0.0%	3.7 (0.9)	3.9		3.6	3.8	3.0	3.0	
het intelligentieniveau van de leerlingen	101 0.0%	3.6 (1.0)	3.7		3.6	3.4	4.3	3.3	
het personeelsverloop binnen het schoolteam	101 0.0%	3.5 (1.0)	3.7		3.4	3.6	3.3	3.0	
het gezinsinkomen van de ouder(s) van de leerlingen	101 0.0%	3.4 (1.1)	3.4		3.5	3.1	3.7	3.5	
het opleidingsniveau van het lerarenteam van de school	101 0.0%	3.3 (1.1)	3.0		3.1	3.4	3.3	2.8	
het totale aantal leerlingen in de school	101	3.1	3.0		3.1	3.0	3.0	3.0	

	0.0%	(1.2)					
het resultaat van de laatste doorlichting van de school	101 0.0%	3.1 (1.2)	3.4	►	2.9	3.1	3.0	2.8

Opmerking: Gemiddelde op basis van vijfpuntenschaal van (1) 'helemaal oneens' over 'eerder oneens', 'noch eens, noch oneens', 'eerder eens' tot (5) 'helemaal eens'.
Opmerking (*): Enkel bevraagd bij schoolleiders S.O.

5.2 DOELEN VOOR FEEDBACKGEBRUIK UIT CENTRALE TOETSEN

Een belangrijke vraag om inzicht te verwerven in feedbackgebruik uit centrale toetsen binnen scholen is het 'waarom' van feedbackgebruik. We onderscheiden hierin doelen op leerling-, leraar- en schoolniveau.

5.2.1 Doelen voor feedbackgebruik op leerlingniveau

Schoolleiders en leraren blijken gemiddeld genomen een neutrale beoordeling te maken van doelen op leerlingniveau (gem = 3.4 , SD = 0.7) (zie Tabel 26). Op itemniveau geven ze aan het eerder eens te zijn de feedback te willen gaan gebruiken als een bron om een zicht te krijgen op sterktes en zwaktes van leerlingen (gem = 3.9, SD = 0.9), om remediëring voor leerlingen die getoetste eindtermen nog niet bereiken, vorm te geven (gem = 3.8, SD = 0.8), om lesonderwerpen te identificeren die extra uitleg behoeven (gem = 3.8, SD = 0.9), tijdens de klassenraad (gem = 3.6, SD = 1.0) en om het studieadvies voor leerlingen te motiveren (gem = 3.5, SD = 1.1). Daarnaast maken schoolleiders en leraren een neutrale inschatting met betrekking tot feedbackgebruik om aangepaste leerdoelen voor individuele leerlingen op te stellen (gem = 3.1, SD = 1.1), om leerlingen in differentiatiegroepen te verdelen (gem = 3.2, SD = 1.1), in oudergesprekken (gem = 3.3, SD = 1.0) en om de attestering van leerlingen te motiveren (gem = 3.3, SD = 1.1). Daarenboven geven ze aan eerder geen gebruik van de feedback te zullen maken om te beslissen of een leerling mag overgaan naar het volgend (leer)jaar (gem = 2.8, SD = 1.2).

Als we verschillen naar functie onderzoeken, stellen we statistisch significante verschillen tussen schoolleiders en leraren vast met betrekking tot doelen voor feedbackgebruik op leerlingniveau. Leraren hebben gemiddeld een neutrale houding ten aanzien van het gebruik van feedback om te beslissen of een leerling mag overgaan naar het volgende (leer)jaar, terwijl schoolleiders hier gematigd negatief tegenover staan [gem SL = 2.4, gem L = 3.0 ; $t(590.0) = -4.95$, $p < 0.01$, $d = -0.49$]. Daarnaast hebben leraren een gematigd positieve houding ten aanzien van feedbackgebruik om advies voor leerlingen te motiveren, terwijl schoolleiders hier eerder neutraal tegenover staan [gem SL = 3.2, gem L = 3.5 ; $t(194.27) = -2.8$, $p < 0.01$, $d =$

0.28]. Verder zijn leraren ook iets positiever dan hun schoolleiders over het gebruik van feedback in oudergesprekken [gem SL = 3.1 gem L = 3.4 ; $t(592.0) = -2.1$, $p = 0.04$, $d = 0.20$] en om de attestering van leerlingen te motiveren [gem SL = 3.0, gem L = 3.4 ; $t(590.0) = -3.7$, $p < 0.01$, $d = 0.36$].

Ook wat betreft verschillen naar onderwijsniveau stellen we kleine statistisch significante verschillen vast. Zo blijken respondenten uit het lager onderwijs gemiddeld gematigd positief om met de feedback aan de slag te gaan om leerlingen in differentiatiegroepen verdelen, terwijl respondenten uit het secundair onderwijs dit item gemiddeld als neutraal beoordelen [gem LO = 3.5, gem SO = 3.1 ; $t(602.0) = 4.3$, $p < 0.01$, $d = 0.39$] en zijn ze gemiddeld iets positiever dan respondenten uit het secundair onderwijs om met de feedback aan de slag te gaan om aangepaste leerdoelen voor individuele leerlingen op te stellen [gem LO = 3.3, gem SO = 3.0 ; $t(596.0) = 3.9$, $p < 0.01$, $d = 0.36$]. Respondenten uit het secundair onderwijs zijn op hun beurt gemiddeld gematigd positief om de feedback te gebruik om het studieadvies voor leerlingen te motiveren, terwijl respondenten uit het lager onderwijs dit doel gemiddeld net wat minder positief beoordelen [gem LO = 3.3, gem SO = 3.5 ; $t(602.0) = -2.6$, $p = 0.01$, $d = 0.24$]. Daarnaast zijn respondenten uit het secundair onderwijs gemiddeld iets minder negatief dan respondenten uit het lager onderwijs ten aanzien van het gebruik van de feedback om te beslissen of een leerling mag overgaan naar het volgende (leer)jaar [gem LO = 2.6, gem SO = 2.9 ; $t(959.0) = -3.3$, $p < 0.01$, $d = 0.30$].

Tot slot blijkt er op itemniveau ook een klein, maar statistisch significante verschil tussen de verschillende onderwijsnetten. Zo zijn de respondenten uit het gemeenschapsonderwijs op leerlingniveau – vanuit een neutrale inschatting – gemiddeld iets positiever dan respondenten uit het vrij gesubsidieerd onderwijs ten aanzien van het gebruik van feedback om te beslissen of een leerling mag overgaan naar het volgende (leer)jaar [gem VG = 2.8, gem G = 3.3 ; $F(12.7, 721.0) = 4.7$, $p < 0.01$, $eta^2 = 0.02$].

Tabel 26
Doelen voor feedbackgebruik op leerlingniveau

	N WN/ NVT	Gem (SD)	T-test Gem SL	L	LO	SO	Anova Gem VG	G	OG
Ik ga de feedback uit centrale toetsen gebruiken:									
als een bron om een zicht te krijgen op sterktes en zwaktes van leerlingen.	630 5.3%	3.9 (0.9)	3.9	3.9	3.8	3.9	3.9	4.0	3.9
om aangepaste leerdoelen voor individuele leerlingen op te stellen.	598 10.1%	3.1 (1.1)	3.2	3.0	3.3 ▸	3.0	3.0	3.3	2.9

om remediëring voor leerlingen die getoetste eindtermen nog niet bereiken, vorm te geven.	622 6.5%	3.8 (0.8)	3.8		3.8	3.8		3.8		3.8	3.9	3.9		
om leerlingen in differentiatiegroepen te verdelen.	604 9.2%	3.2 (1.1)	3.2		3.2	3.5	►	3.1		3.2	3.3	3.3		
om lesonderwerpen te identificeren die extra uitleg behoeven.	619 7.0%	3.8 (0.9)	3.8		3.8	3.8		3.8		3.8	3.9	3.9		
om te beslissen of een leerling mag overgaan naar het volgende (leer)jaar.	597 10.2%	2.8 (1.2)	2.4	◄	3.0	2.6	◄	2.9		2.8	3.3	3.0	VG◄G	
tijdens de klassenraad.	604 9.2%	3.6 (1.0)	3.5		3.6	3.4		3.6		3.5	3.7	3.8		
in oudergesprekken.	599 9.9%	3.3 (1.0)	3.1	◄	3.4	3.2		3.4		3.3	3.5	3.4		
om de attestering van leerlingen te motiveren.	597 10.2%	3.3 (1.1)	3.0	◄	3.4	3.2		3.3		3.2	3.6	3.4		
om het studieadvies voor leerlingen te motiveren.	604 9.2%	3.5 (1.1)	3.2	◄	3.5	3.3	◄	3.5		3.4	3.7	3.5		
Schaal: Doelen op leerlingniveau (α = 0.92)	588 3.8%	3.4 (0.7)	3.3		3.4	3.4		3.4		3.4	3.6	3.5		

5.2.2 Doelen voor feedbackgebruik op leraarniveau

Ook op leraarniveau (zie Tabel 27) blijken schoolleiders en leraren gemiddeld genomen een neutrale beoordeling te maken van mogelijke doelen van feedbackgebruik (gem = 3.4, SD = 0.8). Hierbij maken schoolleiders en leraren een gematigd positieve inschatting van feedbackgebruik om instructie-aanpak te evalueren (gem = 3.5, SD = 0.9), om te reflecteren over het eigen functioneren (gem = 3.6, SD = 1.0) en om zichzelf te professionaliseren (gem = 3.5, SD = 1.0). Verder hebben ze een neutrale inschatting van feedbackgebruik om een eigen visie te ontwikkelen over wat goede klaspraktijk is (gem = 3.2, SD = 1.0) en geven ze aan eerder niet met feedback aan de slag te gaan om keuzes met betrekking tot de eigen werkzaamheden op school te verantwoorden (gem = 2.9, SD = 1.1).

Wat betreft verschillen tussen de verschillende groepen, stellen we enkel kleine statistisch significante verschillen naar functie vast. Zo zijn schoolleiders het er – gemiddeld genomen – in meerdere mate dan leraren mee eens de feedback te gebruiken om instructie-aanpak te evalueren [gem SL = 3.7, gem L = 3.5 ; $t(227.47) = 2.2$, $p = 0.03$, $d = 0.21$] en om een eigen visie te ontwikkelen over wat een goede klaspraktijk is [gem SL = 3.4, gem L = 3.2 ; $t(608.0) = 2.3$, $p = 0.02$, $d = 0.22$]. Leraren schatten gemiddeld met feedback aan

de slag gaan om te reflecteren over het eigen functioneren dan weer net iets positiever in dan schoolleiders [gem SL = 3.3, gem L = 3.7 ; $t(187.8) = -2.8$, $p < 0.01$, $d = 0.29$].

Tabel 27
Doelen voor feedbackgebruik op leraarniveau

	N WN/ NVT	Gem (SD)	T-test Gem SL		L	LO	SO	Anova Gem VG	G	OG
Ik ga de feedback uit centrale toetsen gebruiken:										
om instructie-aanpak te evalueren.	615 7.5%	3.5 (0.9)	3.7	►	3.5	3.6	3.5	3.5	3.6	3.6
om te reflecteren over mijn eigen functioneren.	623 6.3%	3.6 (1.0)	3.3	◄	3.7	3.5	3.6	3.6	3.6	3.7
om een eigen visie te ontwikkelen over wat een goede klaspraktijk is.	615 7.5%	3.2 (1.0)	3.4	►	3.2	3.3	3.2	3.2	3.2	3.3
om mezelf te professionaliseren.	615 7.5%	3.5 (1.0)	3.5		3.5	3.5	3.6	3.5	3.6	3.7
om keuzes met betrekking tot mijn werkzaamheden op school te verantwoorden.	605 9.0%	2.9 (1.1)	3.0		2.9	3.0	2.9	2.9	3.1	3.1
Schaal: Doelen op leraarniveau (α = 0.89)	604 4.7%	3.4 (0.8)	3.4		3.4	3.4	3.4	3.3	3.4	3.5

5.2.3 Doelen voor feedbackgebruik op schoolniveau

Schoolleiders en leraren beoordelen ook de mogelijkheden tot feedbackgebruik op schoolniveau (gem = 3.1, SD = 0.8) eerder neutraal (zie Tabel 28). Zij verwachten eerder wel met de feedback als een bijkomende bron van kwaliteitszorg (gem = 3.7, SD = 0.9) en als een bijkomende manier om aan de kwaliteitsverwachtingen van het Referentiekader onderwijskwaliteit (OK-kader) te werken (gem = 3.5, SD = 1.0) aan de slag te gaan. Ze maken een neutrale beoordeling van de stellingen 'ik ga de feedback gebruiken om concrete actiepunten te formuleren voor het pedagogisch beleid (gem = 3.4, SD = 1.1), om aan visie-ontwikkeling te doen binnen de school (gem = 3.3, SD = 1.0), om te bespreken wat wij als schoolteam verstaan onder goed onderwijs (gem = 3.4, SD = 1.0) en om onze leer- en onderwijsprestaties te verantwoorden aan de onderwijsinspectie (gem = 3.2, SD = 1.1). Tegelijkertijd geven ze aan eerder niet met de feedback aan de slag te gaan om instructiekwaliteit van het lerarenteam in kaart te brengen (gem = 2.8, SD = 1.1), om nieuwe leerlingen aan te trekken (gem = 2.2, SD = 1.1), om naar de buitenwereld toe een positief imago van onze school te creëren (gem = 2.3, SD = 1.1) en om onze leer- en onderwijsprestaties te verantwoorden aan het schoolbestuur (gem = 2.7, SD = 1.1).

Wat betreft doelen voor feedbackgebruik op schoolniveau, stellen we statistisch significante verschillen vast tussen schoolleiders en leraren. Zo blijkt op schaalniveau dat schoolleiders gemiddeld neutraal en leraren gemiddeld gematigd negatief zijn ten aanzien van doelen op schoolniveau [gem SL = 3.4, gem L = 2.9 ; $t(527.0) = 6.9$, $p < 0.01$, $d = 0.70$]. De functie heeft in dit geval een medium effect op de inschatting van de doelen voor feedbackgebruik. Op itemniveau stellen we verder zowel grote, medium als kleine verschillen tussen schoolleiders en leraren vast. Zo zijn schoolleiders gemiddeld gematigd positief om feedback te gebruiken om concrete actiepunten te formuleren voor het pedagogisch beleid, terwijl leraren hierover gemiddeld neutraal zijn [gem SL = 4.0, gem L = 3.1 ; $t(271.25) = 9.0$, $p < 0.01$, $d = 0.84$]. Bovendien zijn schoolleiders gemiddeld meer positief om feedback te gebruiken als een bijkomende bron van kwaliteitszorg [gem SL = 4.3, gem L = 3.5 ; $t(294.17) = 9.7$, $p < 0.01$, $d = 0.89$] en als een bijkomende manier om aan de kwaliteitsverwachtingen van het Referentiekader onderwijskwaliteit te werken [gem SL = 4.2, gem L = 3.3 ; $t(299.34) = 10.7$, $p < 0.01$, $d = 0.99$]. Daarnaast verwachten schoolleiders in meerdere mate dan leraren met de feedback aan de slag te gaan om aan visie-ontwikkeling te doen [gem SL = 3.8, gem L = 3.1 ; $t(560.0) = 7.5$, $p < 0.01$, $d = 0.73$], om te bespreken wat als schoolteam verstaan wordt [gem SL = 3.9, gem L = 3.2 ; $t(263.18) = 7.7$, $p < 0.01$, $d = 0.73$] en om de leer- en onderwijsprestaties te verantwoorden aan de onderwijsinspectie [gem SL = 3.5, gem L = 3.1 ; $t(583.0) = -3.8$, $p < 0.01$, $d = 0.36$]. Verder maken schoolleiders gemiddeld een neutrale en leraren een gemiddeld gematigd negatieve inschatting om met feedback aan de slag te gaan om instructiekwaliteit van het lerarenteam in kaart te brengen [gem SL = 3.3, gem L = 2.6 ; $t(529.0) = 6.6$, $p < 0.01$, $d = 0.65$]. Beide groepen zijn het eerder niet eens om de feedback te gebruiken om de leer- en onderwijsprestaties te verantwoorden aan het schoolbestuur. Hierover zijn schoolleiders het echter in mindere mate oneens dan leraren [gem SL = 2.9, gem L = 2.6 ; $t(537.0) = 3.2$, $p < 0.01$, $d = 0.31$]. Eveneens zijn schoolleiders en leraren gemiddeld gematigd negatief om feedback te gebruiken om nieuwe leerlingen aan te trekken [gem SL = 2.0, gem L = 2.3 ; $t(571.0) = -2.28$, $p = 0.02$, $d = -0.22$] en om naar de buitenwereld toe een positief imago van de school te creëren [gem SL = 2.1, gem L = 2.4 ; $t(572.0) = -2.6$, $p < 0.01$, $d = 0.26$]. Opmerkelijk hierbij is dat de bevraagde leraren deze items minder negatief inschatten.

Ook wat betreft verschillen naar onderwijsniveau stellen we kleine statistisch significante verschillen vast. Respondenten uit het lager onderwijs en respondenten uit het secundair onderwijs zijn gemiddeld neutraal in hun inschatting om met feedback op schoolniveau aan de slag te gaan. Binnen deze neutrale inschatting blijken respondenten uit het lager onderwijs echter gemiddeld iets positiever dan respondenten uit het secundair onderwijs [gem LO = 3.2, gem SO = 3.0 ; $t(388.48) = 3.1$, $p < 0.01$, $d = 0.28$]. Op itemniveau geven zowel respondenten uit het lager onderwijs als in mindere mate respondenten uit het secundair onderwijs

aan eerder wel de feedback te zullen gebruiken als een bijkomende bron van kwaliteitszorg [gem LO = 3.9, gem SO = 3.6 ; t(398.18) = 2.8, p <0.01 , d = 0.24]. Daarnaast zijn respondenten uit het lager onderwijs gemiddeld positiever dan respondenten uit het secundair onderwijs ten aanzien van het gebruik van feedback om als een bijkomende manier aan de kwaliteitsverwachtingen van het Referentiekader onderwijskwaliteit te werken [gem LO = 3.7, gem SO = 3.4 ; t(379.81) = 3.8, p <0.01 , d = 0.34], om aan visie-ontwikkeling te doen binnen de school [gem LO = 3.5, gem SO = 3.2. ; t(564.0) = 2.8, p <0.01 , d = 0.26] of om te bespreken wat het schoolteam verstaat onder goed onderwijs [gem LO = 3.6, gem SO = 3.3 ; t(382.58) = 2.7, p <0.01 , d = 0.24]. Respondenten uit het secundair onderwijs zijn daarentegen minder negatief dan respondenten uit het lager onderwijs ten aanzien van het gebruik van de feedback om naar de buitenwereld toe een positief imago van de school te creëren [gem LO = 2.2, gem SO = 2.4 ; t(576.0) = -2.6, p = 0.01 , d = 0.24].

Tot slot blijken er voor doelen voor feedbackgebruik op schoolniveau ook kleine, maar statistisch significante verschillen tussen de verschillende onderwijsnetten. Zo schatten de bevraagde respondenten uit het gemeenschapsonderwijs feedbackgebruik om nieuwe leerlingen aan te trekken [gem VG = 2.2, gem G = 2.7, gem OG = 2.1 ; F(19.2, 560.5) = 8.8, p <0.01, eta^2 = 0.03] en om naar de buitenwereld toe een positief imago van de school te creëren [gem VG = 2.3, gem G = 2.9, gem OG = 2.2 ; F(20.1, 607.8) = 8.6, p <0.01, eta^2 = 0.03] gemiddeld iets minder negatief in dan respondenten uit het vrij gesubsidieerd en officieel gesubsidieerd onderwijs.

Tabel 28
Doelen voor feedbackgebruik op schoolniveau

	N WN/ NVT	Gem (SD)	T-test Gem SL		L	LO		SO	Anova Gem VG	G	OG
Ik ga de feedback uit centrale toetsen gebruiken:											
om concrete actiepunten te formuleren voor het pedagogisch beleid.	575 13.5%	3.4 (1.1)	4.0	►	3.1	3.5		3.3	3.3	3.4	3.4
als een bijkomende bron van kwaliteitszorg.	596 10.4%	3.7 (0.9)	4.3	►	3.5	3.9	►	3.6	3.6	3.6	3.6
als een bijkomende manier om aan de kwaliteitsverwachtingen van het Referentiekader onderwijskwaliteit (OK-kader) te werken.	559 15.9%	3.5 (1.0)	4.2	►	3.3	3.7	►	3.4	3.4	3.4	3.4
om aan visie-ontwikkeling te doen binnen de school.	566 14.9%	3.3 (1.0)	3.8	►	3.1	3.5	►	3.2	3.2	3.2	3.3

om te bespreken wat wij als schoolteam verstaan onder goed onderwijs.	583 12.3%	3.4 (1.0)	3.9	►	3.2	3.6	►	3.3	3.3	3.3	3.4		
om instructiekwaliteit van het lerarenteam in kaart te brengen.	535 19.5%	2.8 (1.1)	3.3	►	2.6	2.9		2.7	2.7	2.8	2.7		
om nieuwe leerlingen aan te trekken.	578 13.1%	2.2 (1.1)	2.0	◄	2.3	2.1		2.2	2.2	2.7	2.1	VG◄G G►OG	
om naar de buitenwereld toe een positief imago van onze school te creëren.	578 13.1%	2.3 (1.1)	2.1	◄	2.4	2.2	◄	2.4	2.3	2.9	2.2	VG◄G G►OG	
om onze leer- en onderwijsprestaties te verantwoorden aan de onderwijsinspectie.	589 11.4%	3.2 (1.1)	3.5	►	3.1	3.4		3.2	3.2	3.3	3.3		
om onze leer- en onderwijsprestaties te verantwoorden aan het schoolbestuur.	543 18.3%	2.7 (1.1)	2.9	►	2.6	2.8	►	2.6	2.7	2.7	2.7		
Schaal: Doelen op schoolniveau (α = 0.92)	533 7.6%	3.1 (0.8)	3.4	►	2.9	3.2	►	3.0	3.0	3.1	3.0		

5.3 RANDVOORWAARDEN VOOR FEEDBACKGEBRUIK UIT CENTRALE TOETSEN

Om inzicht te verwerven in randvoorwaarden voor feedbackgebruik uit centrale toetsen, brengen we in kaart in welke mate specifieke beïnvloedende feedbacksysteemfactoren (5.3.1), schoolfactoren (5.3.2), contextfactoren (5.3.3) en gebruikersfactoren (5.3.4) voor feedbackgebruik binnen Vlaamse scholen gerealiseerd of wenselijk zijn.

5.3.1 Toegang tot de feedback

Zowel mogelijke onwenselijke effecten van als doelen voor feedbackgebruik uit centrale toetsen, kunnen het gevolg zijn van of bepaald worden door wie toegang heeft tot de feedback. Daarenboven is toegang tot de feedback een noodzakelijke basisvoorwaarde om met de feedback aan de slag te kunnen gaan. We bevroegen daarom schoolleiders en leraren over wie er – volgens hun inschatting – toegang mag hebben tot verschillende vormen van de feedback uit centrale toetsen. We beschrijven achtereenvolgens hun inschatting hieromtrent ten aanzien van niet-geanonimiseerde toetsresultaten van een individuele leerling (5.3.1.1), van toetsresultaten van de deelnemende klassen (5.3.1.2) en van toetsresultaten van de school (5.3.1.2).

5.3.1.1 Toegang tot de niet-geanonimiseerde toetsresultaten van een individuele leerling

Schoolleiders en leraren zijn het gemiddeld genomen in grote mate eens dat de betrokken vakleraar (gem = 4.6, SD = 0.6), de zorgcoördinator/de leerlingbegeleider (gem = 4.4, SD = 0.9) en de schoolleiding (gem = 4.3, SD = 1.0) toegang zouden krijgen tot de niet-geanonimiseerde toetsresultaten van een individuele leerling (zie Tabel 29). Ook over de toegang voor de leerling zelf (gem = 4.1, SD = 1.2), de ouders van de leerling (gem = 4.1, SD = 1.2) en al de leraren die aan de leerling lesgeven (gem = 4.1, SD = 1.1) zijn schoolleiders en leraren het gemiddeld helemaal eens. Daarnaast geven schoolleiders en leraren aan het gemiddeld eerder eens te zijn met de toegang tot deze vorm van feedback voor het CLB (gem = 3.6, SD = 1.3) en blijken ze gemiddeld neutraal in hun inschatting over de toegang voor alle leraren van de betrokken vakgroep (gem = 3.4, SD = 1.3). Verder maken schoolleiders en leraren gemiddeld een gematigd negatieve inschatting over de toegang tot deze feedback voor al de leraren van het lerarenteam (gem = 2.9, SD = 1.4), de onderwijsinspectie (gem = 2.7, SD = 1.4), de pedagogische begeleidingsdienst (gem 2.6, SD = 1.4), het schoolbestuur (gem = 2.5, SD = 1.3) en de schoolgemeenschap (gem = 2.2, SD = 1.2).

Als we verschillen naar functie bekijken, stellen we ten aanzien van de gemiddelde inschatting van schoolleiders en leraren statistisch significante verschillen vast. Leraren maken een positieve inschatting van de toegang tot de individuele toetsresultaten voor de leerling zelf [gem SL = 3.8, gem L = 4.2 ; $t(186193.98) = -3.2$, $p < 0.01$, $d = -0.34$] en voor de ouders van de leerling [gem SL = 3.9, gem L = 4.2 ; $t(533.0) = -2.6$, $p < 0.01$, $d = 0.27$], terwijl schoolleiders hiervan een gematigd positieve inschatting maken. Verder geven schoolleiders in kleine mate aan het meer eens te zijn dan leraren met toegang tot de feedback voor de zorgcoördinator/leerlingbegeleider [gem SL = 4.6, gem L = 4.3 ; $t(279.21) = 2.8$, $p < 0.01$, $d = 0.26$]. Daarnaast is er een effect van functie op de inschatting ten aanzien van de toegang van de schoolleider tot de feedback. Schoolleiders beoordelen het betreffende item beduidend positiever dan leraren [gem SL = 4.7, gem L = 4.2 ; $t(436.62) = 7.1$, $p < 0.01$, $d = 0.60$]. Tot slot blijken schoolleiders negatiever in hun beoordeling dan leraren van de toegang tot deze feedback voor de schoolgemeenschap [gem SL = 1.9, gem L = 2.2 ; $t(257.5) = -3.5$, $p < 0.01$, $d = -0.33$] en de pedagogische begeleidingsdienst [gem SL = 2.1, gem L = 2.8 ; $t(249.18) = -5.0$, $p < 0.01$, $d = -0.48$].

Wat betreft verschil naar onderwijsniveau, stellen we grotere statistisch significante verschillen vast. Zo heeft het onderwijsniveau een groot effect op de beoordeling van de toegang tot deze informatie van de leerling zelf. Respondenten uit het lager onderwijs zijn het gemiddeld eerder eens met de toegang voor de leerling en respondenten uit het secundair onderwijs zijn het hier gemiddeld helemaal mee eens [gem LO = 3.4, gem SO = 4.4 ; $t(243.39) = -9.4$, $p < 0.01$, $d = 0.93$]. Daarnaast zijn respondenten uit het secundair onderwijs het gemiddeld eveneens in meerdere mate dan respondenten uit het lager onderwijs eerder eens met de toegang tot deze informatie voor de ouders van de leerling [gem LO = 3.9, gem SO = 4.2 ; $t(257.14) = -7.7$, p

<0.01 , *d* = 0.74]. Respondenten uit het lager onderwijs maken op hun beurt gemiddeld een meer positieve inschatting van de toegang tot deze informatie voor al de leraren die aan deze leerling lesgeven [gem LO = 4.5, gem SO = 4.0 ; *t*(501.79) = 6.4, *p* <0.01 , *d* = 0.53], voor al de leraren van het lerarenteam [gem LO = 3.3, gem SO = 2.7 ; *t*(600.0) = 5.7, *p* <0.01 , *d* = 0.52] en voor de zorgcoördinator/leerlingbegeleider [gem LO = 4.7, gem SO = 4.3 ; *t*(578.24) = 6.7, *p* <0.01 , *d* = 0.55]. Daarnaast geven respondenten uit het lager onderwijs ook te kennen het gemiddeld in kleine mate meer eens te zijn dan respondenten uit het secundair onderwijs met toegang tot deze informatie voor de schoolleiding [gem LO = 4.5, gem SO = 4.2 ; *t*(434.62) = 3.7, *p* <0.01 , *d* = 0.31] en het CLB [gem LO = 3.8, gem SO = 3.5 ; *t*(385.54) = 2.8, *p* <0.01 , *d* = 0.24].

Naar onderwijsnet stellen we geen statistisch significante verschillen vast.

Tabel 29
Toegang tot de niet-geanonimiseerde toetsresultaten van een individuele leerling

	N WN/ NVT	Gem (SD)	T-test Gem SL		L	LO		SO	Anova Gem VG	G	OG
Wat mij betreft mag er toegang zijn tot de niet-geanonimiseerde toetsresultaten van een individuele leerling voor:											
de leerling zelf	601 1.0%	4.1 (1.2)	3.8	◄	4.2	3.4	◄	4.4	4.2	4.2	4.1
de ouders van deze leerling	603 0.7%	4.1 (1.2)	3.9	◄	4.2	3.5	◄	4.4	4.1	4.4	4.1
de betrokken vakleraar	601 1.0%	4.6 (0.6)	4.6		4.6	4.6		4.6	4.6	4.6	4.7
al de leraren die aan deze leerling lesgeven	604 0.5%	4.1 (1.1)	4.2		4.1	4.5	►	4.0	4.1	4.0	4.2
alle leraren van de betrokken vakgroep	427 0.9%	3.4 (1.3)	3.2		3.5				3.4	3.6	3.4
al de leraren van het lerarenteam	602 0.8%	2.9 (1.4)	2.9		2.8	3.3	►	2.7	2.8	3.0	3.0
de zorgcoördinator/leerlingbegeleider	605 0.3%	4.4 (0.9)	4.6	►	4.3	4.7	►	4.3	4.4	4.3	4.4
de schoolleiding	605 0.3%	4.3 (1.0)	4.7	►	4.2	4.5	►	4.2	4.2	4.2	4.3
het schoolbestuur	596 1.8%	2.5 (1.3)	2.4		2.6	2.5		2.5	2.5	2.6	2.5

de schoolgemeenschap	595 2.0%	2.2 (1.2)	1.9	◄	2.2	2.3		2.1	2.2	2.3	2.3
de pedagogische begeleidingsdienst	596 1.8%	2.6 (1.4)	2.1	◄	2.8	2.7		2.6	2.6	2.8	3.0
het CLB	601 1.0%	3.6 (1.3)	3.4		3.6	3.8	►	3.5	3.5	3.8	3.7
de onderwijsinspectie	594 2.1%	2.7 (1.4)	2.5		2.7	2.7		2.6	2.7	2.8	2.9

5.3.1.2 Toegang tot de toetsresultaten van de deelnemende klassen

In het algemeen zijn schoolleiders en leraren het gemiddeld helemaal eens dat de betrokken vakleraar (gem = 4.5, SD = 0.7), de zorgcoördinator/leerlingbegeleider (gem = 4.3, SD = 1.0), de schoolleiding (gem = 4.3, SD = 1.0) en al de leraren die aan deze klas lesgeven (gem = 4.2, SD = 1.0) toegang behoren te hebben tot de toetsresultaten van de deelnemende klassen (zie Tabel 30). Daarnaast zijn de bevraagde schoolleiders en leraren uit het secundair onderwijs het er gemiddeld tevens eerder mee eens dat alle leraren van de betrokken vakgroep toegang tot deze feedback hebben (gem = 3.8, SD = 1.2). Schoolleiders en leraren maken gemiddeld een neutrale inschatting van de toegang tot de klasresultaten voor het CLB (gem = 3.4, SD = 1.4), voor al de leraren van het lerarenteam (gem = 3.2, SD = 1.4) en voor de betrokken leerlingen (gem = 3.0, SD = 1.5). Daarnaast gaan ze gemiddeld eerder niet akkoord met de toegang voor de onderwijsinspectie (gem = 2.9, SD = 1.4), de pedagogische begeleidingsdienst (gem = 2.8, SD = 1.4), de ouders van de betrokken leerlingen (gem = 2.8, SD = 1.5), het schoolbestuur (gem = 2.8, SD = 1.3), de overheid (gem = 2.6, SD = 1.4) en de schoolgemeenschap (gem = 2.4, SD = 1.3) tot deze feedback. Daarenboven zijn schoolleiders en leraren het gemiddeld helemaal oneens met de toegang van het grote publiek tot deze feedback (gem = 1.6, SD = 1.0).

Als we verschillen naar functie bekijken, stellen we ten aanzien van de inschatting van schoolleiders en leraren statistisch significante verschillen vast. Zo blijken schoolleiders gemiddeld nog positiever dan leraren ten aanzien van toegang voor al de leraren die aan de betrokken klas lesgeven [gem SL = 4.4, gem L = 4.1 ; $t(277.3) = 3.1$, $p < 0.01$, $d = 0.29$], de zorgcoördinator/leerlingbegeleider [gem SL = 4.5, gem L = 4.2 ; $t(289.25) = 4.5$, $p < 0.01$, $d = 0.41$] en de schoolleiding [gem SL = 4.7, gem L = 4.2 ; $t(390.9) = 7.3$, $p < 0.01$, $d = 0.63$] tot deze feedback. Schoolleiders uit het secundair onderwijs zijn het daarenboven gemiddeld in meerdere mate dan leraren uit het secundair onderwijs eens dat ook alle leraren van de betrokken vakgroep toegang krijgen tot deze feedback [gem SL SO = 4.1, gem L SO = 3.7 ; $t(418.0) = 2.2$, $p = 0.03$, $d = 0.29$]. Daarnaast zijn

leraren het gemiddeld in meerdere mate dan schoolleiders eens met de toegang van de betrokken leerlingen tot de feedback [gem SL = 2.7, gem L = 3.0 ; $t(587.0) = -2.2$, $p = 0.03$, $d = 0.21$] en zijn leraren gemiddeld iets minder negatief in hun beoordeling van toegang voor de schoolgemeenschap [gem SL = 2.2, gem L = 2.5 ; $t(584.0) = -2.2$, $p = 0.03$, $d = -0.22$] en de pedagogische begeleidingsdienst [gem SL = 2.7, gem L = 3.0 ; $t(586.0) = -2.5$, $p = 0.01$, $d = 0.25$] tot de feedback met klasresultaten. Daarenboven zijn schoolleiders gemiddeld nog negatiever in hun inschatting dan leraren ten aanzien van toegang voor het grote publiek tot deze feedback [gem SL = 1.5, gem L = 1.7 ; $t(285.39) = -2.8$, $p < 0.01$, $d = -0.26$].

Ook wat onderwijsniveau betreft, stellen we statistisch significante verschillen tussen respondenten uit het lager en het secundair onderwijs vast. Respondenten uit het secundair onderwijs zijn – vanuit een neutrale beoordeling – gemiddeld positiever dan respondenten uit het lager onderwijs ten aanzien van toegang voor de betrokken leerlingen [gem LO = 2.5, gem SO = 3.2 ; $t(374.8) = -5.7$, $p < 0.01$, $d = -0.50$] en de ouders van de betrokken leerlingen [gem LO = 2.4, gem SO = 3.0 ; $t(381.7) = -5.1$, $p < 0.01$, $d = 0.45$] tot deze feedback. Respondenten uit het lager onderwijs zijn het op hun beurt – vanuit een gematigd positieve en positieve inschatting – gemiddeld meer eens dan respondenten uit het secundair onderwijs met de toegang voor al de leraren die aan de betrokken klassen lesgeven [gem LO = 4.4, gem SO = 4.1 ; $t(423.8) = 3.6$, $p < 0.01$, $d = 0.31$], al de leraren van het lerarenteam [gem LO = 3.6, gem SO = 3.0 ; $t(381.4) = 5.5$, $p < 0.01$, $d = 0.48$], de schoolleiding [gem LO = 4.4, gem SO = 4.2 ; $t(425.7) = 3.2$, $p < 0.01$, $d = 0.27$] en het CLB [gem LO = 3.6, gem SO = 3.3 ; $t(377.1) = 2.6$, $p = 0.01$, $d = 0.22$].

Naar onderwijsnet stellen we geen statistisch significante verschillen vast.

Tabel 30
Toegang tot de toetsresultaten van de deelnemende klassen

	N WN/ NVT	Gem (SD)	T-test Gem SL		L	LO		SO	Anova Gem VG	G	OG
Wat mij betreft mag er toegang zijn tot de toetsresultaten van de deelnemende klassen voor:											
de betrokken leerlingen	591 2.6%	3.0 (1.5)	2.7	◄	3.0	2.5	◄	3.2	3.0	3.2	2.6
de ouders van de betrokken leerlingen	592 2.5%	2.8 (1.5)	2.6		2.9	2.4	◄	3.0	2.9	3.1	2.3
de betrokken vakleraar	594 2.1%	4.5 (0.7)	4.5		4.6	4.5		4.6	4.6	4.5	4.5

al de leraren die aan deze klas lesgeven	597 4.6%	4.2 (1.0)	4.4	►	4.1	4.4	►	4.1	4.2	4.0	4.2
alle leraren van de betrokken vakgroep	422 2.1%	3.8 (1.2)	4.1	►	3.7				3.8	3.7	3.6
al de leraren van het lerarenteam	594 2.1%	3.2 (1.4)	3.3		3.1	3.6	►	3.0	3.1	3.2	3.2
de zorgcoördinator/leerlingbegeleider	597 1.6%	4.3 (1.0)	4.5	►	4.2	4.6	►	4.1	4.3	4.2	4.1
de schoolleiding	598 1.5%	4.3 (1.0)	4.7	►	4.2	4.4	►	4.2	4.2	4.1	4.2
het schoolbestuur	589 3.0%	2.8 (1.3)	2.9		2.7	2.8		2.8	2.8	2.7	2.6
de schoolgemeenschap	590 2.8%	2.4 (1.3)	2.2	◄	2.5	2.5		2.4	2.4	2.6	2.4
de pedagogische begeleidingsdienst	592 2.5%	2.9 (1.4)	2.7	◄	3.0	3.0		2.9	3.0	2.9	3.1
het CLB	598 1.5%	3.4 (1.4)	3.3		3.5	3.6	►	3.3	3.4	3.6	3.4
de onderwijsinspectie	591 2.6%	2.9 (1.4)	2.9		3.0	3.0		2.9	2.9	3.1	3.1
de overheid	588 3.1%	2.6 (1.4)	2.5		2.6	2.6		2.6	2.6	2.4	2.5
het grote publiek	588 3.1%	1.6 (1.0)	1.5	◄	1.7	1.6		1.7	1.7	1.7	1.5

Op de vraag of leraren rechtstreeks toegang wensen tot de resultaten van hun klas op de centrale toetsen zonder dat hun schoolleiding daarbij tussenpersoon is beoordelen respondenten in het algemeen gematigd positief (gem = 3.7, SD = 1.2) (zie Tabel 31). Schoolleiders zijn het hier echter gemiddeld in mindere mee eens dan leraren [gem SL = 3.1, gem L = 3.8 ; $t(200.2) = -5.8$, $p < 0.01$, $d = -0.59$].

Tabel 31
Wenselijkheid naar rechtstreekse toegang van leraren tot de resultaten van hun klas op centrale toetsen

	N WN/ NVT	Gem (SD)	T-test Gem SL		L	LO	SO	Anova Gem VG	G	OG
Leraren moeten rechtstreeks toegang hebben tot de resultaten van hun klas op centrale toetsen zonder dat hun schoolleiding daarbij tussenpersoon is.	579 3.7%	3.7 (1.2)	3.1	◄	3.8	3.7	3.7	3.8	3.6	3.9

5.3.1.3 Toegang tot de toetsresultaten van de school

Schoolleiders en leraren beoordelen de toegang tot de toetsresultaten van de school voor de betrokken vakleraar (gem = 4.4, SD = 0.8), de schoolleiding (gem = 4.4, SD = 0.8), de zorgcoördinator/leerlingbegeleider (gem : 4.3, SD = 1.0) en al de leraren van de betrokken vakgroep (gem = 4.2, SD = 1.0) in het algemeen als positief en de toegang voor al de leraren van het lerarenteam (gem = 4.0, SD = 1.2), het schoolbestuur (gem = 3.6, SD = 1.3) en het CLB (gem = 3.6, SD = 1.3) als gematigd positief (zie Tabel 32). Schoolleiders en leraren geven aan gemiddeld neutraal te zijn ten aanzien van de toegang tot de schoolresultaten voor de onderwijsinspectie (gem = 3.4, SD = 1.4), de pedagogische begeleidingsdienst (gem = 3.4, SD = 1.4) de schoolgemeenschap (gem = 3.2, SD = 1.4), de betrokken leerlingen (gem = 3.1, SD = 1.4), de ouders van de betrokken leerlingen (gem = 3.1, SD = 1.5) en de overheid (gem = 3.0, SD = 1.4). Schoolleiders en leraren zijn gemiddeld gematigd negatief in hun beoordeling van de toegang tot deze feedback voor al de leerlingen van de school (gem = 2.7, SD = 1.5), de ouders van al de leerlingen van de school (gem = 2.6, SD = 1.4) en het grote publiek (gem = 2.1, SD = 1.3).

Als we verschillen naar functie bekijken, stellen we kleine statistisch significante verschillen vast tussen schoolleiders en leraren, waarbij de schoolleiders telkens gemiddeld een positievere inschatting maken. Zo zijn schoolleiders het gemiddeld in beperkte mate meer dan leraren eens met toegang tot de toetsresultaten van de school voor al de leraren van de betrokken vakgroep [gem SL = 4.4, gem L = 4.2 ; $t(152.3) = 2.6$, $p = 0.01$, $d = 0.29$], al de leraren van het lerarenteam [gem SL = 4.2, gem L = 3.9 ; $t(592.0) = 2.6$, $p < 0.01$, $d = 0.27$], de zorgcoördinator/leerlingbegeleider [gem SL = 4.5, gem L = 4.3 ; $t(287.5) = 3.3$, $p < 0.01$, $d = 0.30$] en de schoolleiding [gem SL = 4.6, gem L = 4.3 ; $t(318.0) = 4.6$, $p < 0.01$, $d = 0.41$]. Daarnaast beoordelen schoolleiders de toegang voor het schoolbestuur gemiddeld als gematigd positief, terwijl leraren hier gemiddeld een neutrale beoordeling aan geven [gem SL = 3.8, gem L = 3.5 ; $t(247.5) = 2.5$, $p = 0.01$, $d = 0.24$].

Ook naar onderwijsniveau stellen we statistisch significante verschillen vast. Zo zijn respondenten uit het secundair onderwijs gemiddeld minder negatief dan respondenten uit het lager onderwijs ten aanzien van toegang tot schoolresultaten voor de betrokken leerlingen [gem LO = 2.6, gem SO = 3.3 ; $t(588.0) = -5.5$, $p < 0.01$, $d = -0.51$], al de leerlingen van de school [gem LO = 2.2, gem SO = 2.9 ; $t(361.1) = -5.1$, $p < 0.01$, $d = -0.45$], de ouders van de betrokken leerlingen [gem LO = 2.7, gem SO = 3.3 ; $t(590.0) = -4.5$, $p < 0.01$, $d = -0.41$] en de ouders van al de leerlingen van de school [gem LO = 2.3, gem SO = 2.8 ; $t(350.8) = -3.9$, $p < 0.01$, $d = -0.34$]. Daarnaast zijn respondenten uit het lager onderwijs het in meerdere mate dan respondenten uit het

secundair onderwijs eens met de toegang voor al de leraren van het lerarenteam [gem LO = 4.3, gem SO = 3.9 ; $t(403.0) = 3.6$, $p < 0.01$, $d = 0.31$] en de zorgcoördinator/leerlingbegeleider [gem LO = 4.6, gem SO = 4.2 ; $t(462.1) = 6.1$, $p < 0.01$, $d = 0.50$] tot deze feedback.

Naar onderwijsnet stellen we geen statistisch significante verschillen vast.

Tabel 32
Toegang tot de toetsresultaten van de school

	N WN/ NVT	Gem (SD)	T-test Gem SL		L	LO		SO	Anova Gem VG	G	OG
Wat mij betreft mag er toegang zijn tot de toetsresultaten van de school voor:											
de betrokken leerlingen	590 2.8%	3.1 (1.4)	3.0		3.1	2.6	◄	3.3	3.2	3.1	2.8
al de leerlingen van de school	592 2.5%	2.7 (1.5)	2.6		2.7	2.2	◄	2.9	2.8	2.6	2.3
de ouders van de betrokken leerlingen	592 2.5%	3.1 (1.5)	3.0		3.1	2.7	◄	3.3	3.1	3.1	2.9
de ouders van al de leerlingen van de school	593 2.3%	2.6 (1.4)	2.7		2.6	2.3	◄	2.8	2.7	2.5	2.4
de betrokken vakleraar	593 2.3%	4.4 (0.8)	4.5		4.4	4.5		4.4	4.5	4.5	4.4
al de leraren van de betrokken vakgroep	423 1.9%	4.2 (1.0)	4.4	►	4.2				4.3	4.0	4.2
al de leraren van het lerarenteam	598 1.5%	4.0 (1.2)	4.2	►	3.9	4.3	►	3.9	4.0	3.8	4.1
de zorgcoördinator/leerlingbegeleider	597 1.6%	4.3 (1.0)	4.5	►	4.3	4.6	►	4.2	4.3	4.2	4.3
de schoolleiding	598 1.5%	4.4 (0.8)	4.6	►	4.3	4.5		4.4	4.4	4.2	4.4
het schoolbestuur	592 2.5%	3.6 (1.3)	3.8	►	3.5	3.4		3.7	3.6	3.4	3.3
de schoolgemeenschap	590 2.8%	3.1 (1.4)	3.0		3.1	3.0		3.1	3.1	3.1	2.9
de pedagogische begeleidingsdienst	591 2.6%	3.4 (1.4)	3.4		3.4	3.5		3.4	3.4	3.4	3.3
het CLB	596	3.6	3.5		3.6	3.8	►	3.5	3.6	3.7	3.6

	1.8%	(1.3)									
de onderwijsinspectie	592 2.5%	3.4 (1.4)	3.5	3.4	3.4	3.4	3.3	3.4	3.6		
de overheid	589 3.0%	3.0 (1.4)	3.0	3.0	3.0	3.0	3.0	2.9	3.0		
het grote publiek	588 3.1%	2.1 (1.3)	1.9	2.1	2.0	2.1	2.1	2.0	1.9		

5.3.2 Schoolfactoren met betrekking tot feedbackgebruik

Met betrekking tot schoolfactoren die feedbackgebruik uit centrale toetsen stimuleren of hinderen, bevroegen we – op basis van de conceptuele uitgangspunten – percepties over het actueel gebruik van gestandaardiseerde toetsen (5.3.2.1), over het ervaren van schoolinterne verwachtingen met betrekking tot feedbackgebruik uit centrale toetsen (5.3.2.2) en over aspecten van het beleidsvoerend vermogen van scholen en transformationeel leiderschap met betrekking tot feedbackgebruik uit centrale toetsen (5.3.2.3).

5.3.2.1 Actueel gebruik van gestandaardiseerde toetsen

Met betrekking tot het actueel gebruik van gestandaardiseerde toetsen, bevroegen we specifieke respondenten zowel over het huidige gebruik van gestandaardiseerde toetsen op school als over de mate waarin op de school momenteel aan de slag wordt gegaan met de feedback uit die gestandaardiseerde toetsen (zie Tabel 33).

Gemiddeld genomen geven schoolleiders en leraren aan eerder geen gebruik te maken van gestandaardiseerde diagnose-instrumenten (gem = 2.9, SD = 1.3) en schatten ze het gebruik van toetsen in een leerlingvolgsysteem (gem = 3.5, SD = 1.5) en van paralleltoetsen (gem = 3.5, SD = 1.3) gemiddeld eerder positief in. Schoolleiders en leraren uit het lager onderwijs geven daarenboven gemiddeld aan het er helemaal mee eens te zijn dat op de school gebruik gemaakt wordt van IDP- of OVSG-toetsen (gem = 4.5, SD = 1.0). Hierbij aansluitend geven schoolleiders en leraren gemiddeld genomen aan dat op de school eerder niet intensief met de feedback/resultaten van gestandaardiseerde diagnose-instrumenten (gem = 2.8, SD = 1.2) en paralleltoetsen (gem = 2.9, SD = 1.2) aan de slag wordt gegaan. Tevens beoordelen ze het intensief aan de slag gaan met de feedback/resultaten van toetsen in een leerlingvolgsysteem (gem = 3.2, SD = 1.2) gemiddeld als neutraal. Schoolleiders en leraren uit het lager onderwijs beantwoorden de stelling 'op onze school gaan we intensief aan de slag met de feedback/resultaten van IDP- of OVSG-toetsen' eveneens gemiddeld als neutraal (gem = 3.4, SD = 1.1).

Als we verschillen naar functie bekijken, stellen we enkele kleine verschillen tussen schoolleiders en leraren vast. Opvallend is dat schoolleiders gemiddeld aangeven zowel in meerdere mate gebruik te maken van toetsen in een leerlingvolgsysteem dan leraren [gem SL = 3.7, gem L = 3.4 ; $t(533.0) = 2.2$, $p = 0.03$, $d = 0.22$] als in meerdere mate intensief aan de slag te gaan met de feedback/resultaten van toetsen in een leerlingvolgsysteem dan leraren [gem SL = 3.5, gem L = 3.1 ; $t(220.1) = 3.7$, $p < 0.01$, $d = 0.37$]. Leraren daarentegen geven dan weer gemiddeld aan in meerdere mate gebruik te maken van parelleltoetsen dan schoolleiders [gem SL = 3.1, gem L = 3.6 ; $t(184.9) = -4.1$, $p < 0.01$, $d = -0.43$].

Daarnaast stellen we ook verschillen naar onderwijsniveau vast. Respondenten uit het lager onderwijs geven gemiddeld aan in meerdere mate gebruik te maken van de toetsen in een leerlingvolgsysteem dan respondenten uit het secundair onderwijs [gem LO = 4.6, gem SO = 2.9 ; $t(513.3) = 19.9$, $p < 0.01$, $d = 1.73$]. Tegelijkertijd geven respondenten uit het lager onderwijs eveneens aan in meerdere mate gebruik te maken van gestandaardiseerde diagnose-instrumenten [gem LO = 3.6, gem SO = 2.7 ; $t(286.0) = 8.2$, $p < 0.01$, $d = 0.79$] en in meerdere mate intensief aan de slag te gaan met de feedback/resultaten van gestandaardiseerde diagnose-instrumenten [gem LO = 3.4, gem SO = 2.5 ; $t(488.0) = 7.4$, $p < 0.01$, $d = 0.74$] en met de feedback/resultaten van een leerlingvolgsysteem [gem LO = 3.5, gem SO = 3.1 ; $t(394.6) = 4.8$, $p < 0.01$, $d = 0.42$] dan respondenten uit het secundair onderwijs. Daarnaast is er ook een verschil tussen beide groepen wat betreft het gebruik van paralleltoetsen. Respondenten uit het secundair onderwijs beoordelen het gebruik van paralleltoetsen op de school positiever dan respondenten uit het lager onderwijs [gem LO = 3.1, gem SO = 3.6 ; $t(188.6) = -3.6$, $p < 0.01$, $d = -0.38$].

Ook naar onderwijsnet stellen we statistisch significante verschillen vast. Zo geven respondenten uit het officieel gesubsidieerd lager onderwijs gemiddeld aan in meerdere mate intensief aan de slag te gaan met de feedback/resultaten van (IDP- of) OVSG-toetsen dan respondenten uit het vrij gesubsidieerd lager onderwijs en het lager onderwijs van het gemeenschapsonderwijs [gem VG = 3.3, gem G = 2.9, gem OG = 4.1 ; $F(2.0, 30.7) = 10.0$, $p < 0.01$, $eta^2 = 0.07$].

Tabel 33
Actueel gebruik van gestandaardiseerde toetsen

	N WN/N VT	Gem (SD)	T-test Gem SL		L	LO		SO	Anova Gem VG	G	OG	
Op onze school maken we gebruik van:												
de toetsen in een leerlingvolgsysteem.	540 14.8%	3.5 (1.5)	3.7	►	3.4	4.6	►	2.9	3.4	3.4	3.5	
IDP- of OVSG-toetsen.	176 5.9%	4.5 (1.0)	4.5		4.5				4.4	4.4	4.6	
paralleltoetsen.	534 15.6%	3.5 (1.3)	3.1	◄	3.6	3.1	◄	3.6	3.6	3.4	3.7	
gestandaardiseerde diagnose-instrumenten.	510 19.6%	2.9 (1.3)	3.1		2.9	3.6	►	2.7	2.8	3.0	3.3	
Op onze school gaan we intensief aan de slag met de feedback/resultaten van:												
toetsen in een leerlingvolgsysteem.	567 10.6%	3.2 (1.2)	3.5	►	3.1	3.5	►	3.0	3.0	3.3	3.3	
IDP- of OVSG-toetsen.	163 12.4%	3.4 (1.1)	3.6		3.4				3.3	2.9	4.1	VG◄OG G◄OG
paralleltoetsen.	510 19.4%	2.9 (1.2)	2.7		2.9	2.9		2.9	2.9	2.9	3.1	
gestandaardiseerde diagnose-instrumenten.	490 22.7%	2.8 (1.2)	2.9		2.7	3.4	►	2.5	2.7	2.8	3.1	

5.3.2.2 Schoolinterne verwachtingen met betrekking tot feedbackgebruik

Indien relevant werd aan specifieke doelgroepen het ervaren van schoolinterne verwachtingen met betrekking tot feedbackgebruik bevraagd (ze Tabel 34).

Over het algemeen schatten schoolleiders en leraren de duidelijkheid aan schoolinterne verwachtingen gematigd negatief in. Zo is het voor leraren uit het lager en secundair onderwijs gemiddeld eerder niet duidelijk wat de schoolleiding (gem = 2.2, SD = 1.2), collega's (gem = 2.2, SD = 1.1), collega's in de vakgroep (gem = 2.2, SD = 1.2) en leerlingen (gem = 2.0, SD = 1.0) van hen verwachten met betrekking tot feedbackgebruik. Ook schoolleiders geven aan dat het voor hen eerder niet duidelijk is wat collega-directeuren (gem = 2.2, SD = 1.2) en leraren (gem = 2.4, SD = 1.1) van hen verwachten bij feedbackgebruik.

Ook de verwachtingen de ouders van de leerlingen (gem = 2.0, SD = 1.0) zijn voor schoolleiders en leraren gemiddeld eerder niet duidelijk.

Hoewel respondenten uit het lager onderwijs en respondenten uit het secundair onderwijs de schoolinterne verwachtingen bij feedbackgebruik gemiddeld eerder negatief inschatten, stellen we tussen beide groepen wel statistisch significante verschillen vast. Leraren uit het lager onderwijs zijn positiever dan leraren uit het secundair onderwijs over de duidelijkheid van verwachtingen van de schoolleiding [gem LO = 2.8, gem SO = 2.0 ; $t(461.0) = 7.2$, $p < 0.01$, $d = 0.73$], van collega's [gem LO = 2.8, gem SO = 2.0 ; $t(116.0) = 6.1$, $p < 0.01$, $d = 0.70$] en van leerlingen [gem LO = 2.3, gem SO = 1.9 ; $t(139.0) = 2.5$, $p < 0.01$, $d = 0.43$] bij feedbackgebruik. Daarnaast zijn ook schoolleiders uit het lager onderwijs positiever dan schoolleiders uit het secundair onderwijs over de stelling die peilt naar de duidelijkheid van verwachtingen van leraren voor schoolleiders bij feedbackgebruik [gem LO = 2.8, gem SO = 2.1 ; $t(138.0) = 3.8$, $p < 0.01$, $d = 0.65$]. Bovendien schatten schoolleiders en leraren uit het lager onderwijs de duidelijkheid van verwachtingen van ouders [gem LO = 2.3, gem SO = 1.9 ; $t(460.0) = 4.5$, $p < 0.01$, $d = 0.48$] bij feedbackgebruik positiever in dan schoolleiders en leraren uit het secundair onderwijs.

Naar onderwijsnet stellen we geen statistisch significante verschillen tussen de groepen vast.

Tabel 34
Schoolinterne verwachtingen bij feedbackgebruik

	N WN/NVT	Gem (SD)	T-test Gem SL	L	LO	SO	Anova Gem VG	G	OG
Omtrent het gebruik van de feedback uit centrale toetsen, is het voor mij (i.c. een leraar uit het L.O. of het S.O.) duidelijk:									
wat mijn schoolleiding van mij verwacht.	463 9.9%	2.2 (1.2)			2.8 ▸	2.0	2.2	2.4	2.5
wat mijn collega's van mij verwachten.	414 10.4%	2.2 (1.1)			2.8 ▸	2.0	2.1	2.4	2.4
Omtrent het gebruik van de feedback uit centrale toetsen, is het voor mij (i.c. een schoolleider uit het S.O.) duidelijk:									
wat mijn collega-directeuren van mij verwachten.	61 6.1%	2.2 (1.2)					2.3	3.0	1.5
Omtrent het gebruik van de feedback uit centrale toetsen, is het voor mij (i.c. een schoolleider uit het L.O of het S.O.) duidelijk:									
wat de leraren van mij verwachten.	140 8.5%	2.4 (1.1)			2.8 ▸	2.1	2.5	2.8	2.4

wat de leerlingen van mij verwachten.	141 7.8%	2.0 (1.0)			2.3 ► 1.9	2.1	2.3	2.0	
Omtrent het gebruik van de feedback uit centrale toetsen, is het voor mij (i.c. een leraar uit het S.O.) duidelijk:									
wat mijn collega's in de vakgroep van mij verwachten.	309 12.0%	2.2 (1.2)				2.1	2.5	2.2	
Omtrent het gebruik van de feedback uit centrale toetsen, is het voor mij (i.c. een leraar of schoolleider uit het L.O of het S.O.) duidelijk:									
wat de ouders van mij verwachten.	462 11.8%	2.0 (1.0)	2.0	2.0	2.3 ► 1.9	2.0	2.0	2.1	

5.3.2.3 Beleidsvoerend vermogen met betrekking tot feedbackgebruik

Uit de theoretische uitgangspunten blijkt dat zowel ondersteunende relaties en samenwerking en geïnformeerde besluitvoering tot wezenlijke verschillen kan leiden inzake geïnformeerde besluitvoering van scholen. Daarom werd ook in de survey het beleidsvoerend vermogen van scholen met betrekking tot feedbackgebruik geoperationaliseerd in 'ondersteunende relaties en samenwerking' en 'gezamenlijke doelgerichtheid'.

Wat betreft de ondersteunende relaties en samenwerking (zie Tabel 35) stellen we een eerder gematigd positief schaalgemiddelde vast (gem = 3.6, SD = 1.1). Daarbij geven respondenten aan het gemiddeld eerder eens te zijn met de stellingen dat er op de school iemand aanwezig is bij wie men terecht kan voor vragen over het interpreteren van de feedback (gem = 3.5, SD = 1.3) of voor het helpen van lessen te trekken uit de feedback (gem = 3.5, SD = 1.2). Daarnaast geven respondenten ook aan zich gemiddeld eerder wel ondersteund te voelen door collega's bij het interpreteren en analyseren van de feedback (gem = 3.7, SD = 1.1).

Wat betreft verschillen naar onderwijsniveau voor 'ondersteunende relaties en samenwerking' stellen we zowel voor respondenten uit het lager als uit het secundair een eerder positieve inschatting in. Hoewel het verschil tussen beide groepen klein is, blijken respondenten uit het lager onderwijs gemiddeld toch iets positiever dan respondenten uit het secundair onderwijs [gem LO = 3.8 , gem SO = 3.5 ; $t(396.7) = 3.3$, $p < 0.01$, $d = 0.31$]. Op itemniveau blijken zowel respondenten uit lager onderwijs als – in mindere mate – respondenten uit het secundair onderwijs de stelling zich ondersteund te voelen door collega's bij het interpreteren en analyseren van de feedback gemiddeld als het meest positief te beoordelen [gem LO = 3.9 , gem SO = 3.7 ; $t(385.8) = 2.4$, $p = 0.02$, $d = 0.22$]. Daarnaast valt op dat respondenten uit het lager onderwijs

gemiddeld eerder positief zijn over de stellingen dat op de school iemand aanwezig is bij wie men met vragen over het interpreteren van de feedback terecht kan [gem LO = 3.8 , gem SO = 3.4 ; $t(37.2.5) = 3.6$, $p < 0.01$, $d = 0.34$] en bij wie men terecht kan om geholpen te worden bij trekken van lessen uit de feedback [gem LO = 3.8 , gem SO = 3.4 ; $t(373.8) = 3.3$, $p < 0.01$, $d = 0.31$], terwijl respondenten uit het secundair onderwijs deze stellingen gemiddeld als neutraal beoordelen. Naar onderwijsnet merken we geen statistisch significante verschillen op met betrekking tot ondersteunende relaties en samenwerking ten aanzien van feedbackgebruik.

Tabel 35
Ondersteunende relaties en samenwerking ten aanzien van feedbackgebruik

	N WN/ NVT	Gem (SD)	T-test Gem SL	L	LO		SO	Anova Gem VG	G	OG
Op onze school is er iemand aanwezig bij wie ik met vragen over het interpreteren van de feedback uit centrale toetsen terecht kan.	486 22.2 %	3.5 (1.3)	3.6	3.5	3.8	►	3.4	3.6	3.4	3.4
Op onze school is er iemand aanwezig die mij kan helpen lessen te trekken uit de feedback van centrale toetsen.	487 22.1%	3.5 (1.2)	3.7	3.5	3.8	►	3.4	3.6	3.4	3.3
Ik voel me door mijn collega's ondersteund bij het interpreteren en analyseren van de feedback uit centrale toetsen.	524 16.2%	3.7 (1.1)	3.7	3.8	3.9	►	3.7	3.8	3.5	3.8
Schaal: Ondersteunende relaties en samenwerking m.b.t. feedbackgebruik (α = 0.80)	469 15.8%	3.6 (1.1)	3.7	3.6	3.8	►	3.5	3.7	3.4	3.5

Met betrekking tot de gezamenlijke doelgerichtheid van de school ten aanzien van feedbackgebruik (zie Tabel 36) beoordelen schoolleiders en leraren over het algemeen gematigd negatief (gem = 2.7, SD = 0.9). Schoolleiders en leraren zijn het gemiddeld eerder oneens met de stellingen dat het gebruik van feedback op de school een prioriteit is (gem = 2.7, SD = 1.1), dat er op de school een duidelijke visie is over hoe de feedback kan gebruikt worden (gem = 2.6, SD = 1.1) en dat men het op de school onderling eens is over de redenen om (gem = 2.8, SD = 1.1) en de wijze waarop (gem = 2.7, SD = 1.0) met de feedback aan de slag te gaan.

Ook voor 'gezamenlijke doelgerichtheid' stellen we statistisch significante verschillen naargelang de functie die respondenten uitoefenen vast op zowel schaal- als itemniveau. Schoolleiders beoordelen de gezamenlijke doelgerichtheid gemiddeld iets minder negatief dan leraren [gem SL = 2.9, gem L = 2.7 ; $t(404.0) = 2.0$, $p =$

0.04 , *d* = 0.23], al is het verschil klein en het gemiddelde niet van die aard dat er uit de antwoorden van schoolleiders een gezamenlijke doelgerichtheid met betrekking tot feedbackgebruik blijkt. Hoewel ook op itemniveau een eerder afwezigheid van gezamenlijke doelgerichtheid blijkt, zijn schoolleiders gemiddeld iets minder negatief dan leraren wat betreft de stellingen dat men het op de school onderling eens is over de redenen waarom [gem SL = 3.0 , gem L = 2.7 ; *t*(412.0) = 2.5, *p* = 0.01 , *d* = 0.27] en de wijze waarop [gem SL = 2.9, gem L = 2.7 ; *t*(407.0) = 2.7, *p* <0.01 , *d* = 0.29] met de feedback aan de slag kan gegaan worden. Hoewel we met betrekking tot 'gezamenlijke doelgerichtheid' voor zowel respondenten uit het lager onderwijs als voor respondenten uit het secundair onderwijs een gematigd negatief schaalgemiddelde vaststellen, beoordelen respondenten uit het lager onderwijs de gezamenlijke doelgerichtheid met betrekking tot feedback gemiddeld iets positiever dan respondenten uit het secundair onderwijs [gem LO = 2.9 , gem SO = 2.6 ; *t*(406.0) = 3.7, *p* <0.01 , *d* = 0.38]. Op itemniveau stellen we vast dat respondenten uit het lager onderwijs gemiddeld neutraal zijn over de stellingen dat op de school het gebruik van de feedback een prioriteit is [gem LO = 3.0 , gem SO = 2.6 ; *t*(441.0) = 3.8, *p* <0.01 , *d* = 0.38] en dat men het op de school onderling eens is over de redenen waarom met de feedback aan de slag te gaan [gem LO = 3.0 , gem SO = 2.6 ; *t*(415.0) = 3.6, *p* <0.01 , *d* = 0.37] en gemiddeld eerder oneens zijn met de stellingen dat er op de school een visie is over hoe de feedback gebruikt kan worden [gem LO = 2.8 , gem SO = 2.5 ; *t*(430.0) = 2.9, *p* <0.01 , *d* = 0.30] en dat men her op de school het onderling eens is over de wijzen waarop met de feedback aan de slag kan gegaan worden [gem LO = 2.9 , gem SO = 2.6 ; *t*(410.0) = 3.1, *p* <0.01 , *d* = 0.32]. Respondenten uit het secundair onderwijs schatten deze vier stellingen telkens gemiddeld als gematigd negatief in. Naar onderwijsnet merken we geen statistisch significante verschillen op met betrekking tot gezamenlijke doelgerichtheid ten aanzien van feedbackgebruik.

Tabel 36
Gezamenlijke doelgerichtheid ten aanzien van feedbackgebruik

	N WN/ NVT	Gem (SD)	T-test Gem SL		L	LO		SO	Anova Gem VG	G	OG
Op onze school is gebruik van de feedback uit centrale toetsen een prioriteit.	443 29.1%	2.7 (1.1)	2.9		2.7	3.0	►	2.6	2.7	2.7	3.0
Op onze school is er een duidelijke visie over hoe de feedback uit centrale toetsen gebruikt kan worden.	432 30.9%	2.6 (1.1)	2.6		2.6	2.8	►	2.5	2.6	2.6	2.6
Op onze school zijn we het onderling eens over de redenen waarom we met de feedback uit centrale toetsen aan de slag gaan.	417 33.3%	2.8 (1.1)	3.0	►	2.7	3.0	►	2.6	2.7	2.8	3.0
Op onze school zijn we het onderling eens over de wijzen waarop we met de feedback uit centrale toetsen aan de slag kunnen gaan.	412 34.1%	2.7 (1.0)	2.9	►	2.7	2.9	►	2.6	2.7	2.7	2.9

Schaal: Gezamenlijke doelgerichtheid m.b.t. feedbackgebruik (α = 0.93)	408 31.7%	2.7 (0.9)	2.9	►	2.7	2.9	►	2.6	2.7	2.7	2.9

5.3.2.4 Transformationeel leiderschap met betrekking tot feedbackgebruik

De vragenlijst peilde eveneens naar aspecten van het transformationeel leiderschap met betrekking tot feedbackgebruik (zie Tabel 37).

Wat betreft transformationeel leiderschap stellen we een neutraal schaalgemiddelde vast (gem = 3.1, SD = 0.9). Schoolleiders en leraren beoordelen de stelling dat de schoolleider de leraren op de school coacht bij het gebruik van feedback gemiddeld als gematigd negatief (gem = 2.9, SD = 1.1). Schoolleiders en leraren staan gemiddeld neutraal ten opzichte van de stellingen dat de schoolleiding op de school het gebruik van feedback aanmoedigt (gem = 3.2, SD = 1.1), dat de schoolleiding op de school de feedback met de teamleden bespreekt (gem = 3.1, SD = 1.2), dat de schoolleiding op de school ervoor zorgt dat leraren voldoende vaardigheden hebben om met de feedback aan de slag te kunnen (gem = 3.1, SD = 1.1) en dat de schoolleiding op de school oog heeft voor de informatiebehoeften van leren (gem = 3.4, SD = 1.1).

Schoolleiders beoordelen transformationeel leiderschap gemiddeld als eerder positief, terwijl het schaalgemiddelde van leraren neutraal is [gem SL = 3.9 , gem L = 3.0 ; $t(264.8) = 5.5$, $p < 0.01$, $d = 0.59$].

Op itemniveau komt dit transformationeel leiderschap vooral tot uiting in de stellingen dat op de school de schoolleiding leraren coacht bij het gebruik van de feedback [gem SL = 3.3 , gem L = 2.7 ; $t(428.0) = 5.4$, $p < 0.01$, $d = 0.61$] en dat op de school de schoolleiding oogt heeft voor de informatiebehoeften van leraren [gem SL = 3.9 , gem L = 3.3 ; $t(293.9) = 6.9$ $p < 0.01$, $d = 0.65$]. Verder beoordelen schoolleiders de stellingen dat de schoolleiding op de school het gebruik van de feedback aan moedigt [gem SL = 3.4 , gem L = 3.1 ; $t(411.0) = 2.6$, $p = 0.01$, $d = 0.28$], dat de schoolleiding op de school de feedback bespreekt met de teamleden [gem SL = 3.5 , gem L = 3.0 ; $t(413.0) = 4.3$, $p < 0.01$, $d = 0.47$] en dat de schoolleiding op de school ervoor zorgt dat leraren voldoende vaardigheden hebben om met de feedback uit de centrale toetsen aan de slag te kunnen gaan [gem SL = 3.3 , gem L = 2.9 ; $t(450.0) = 3.5$, $p < 0.01$, $d = 0.39$] positiever dan leraren, maar wel steeds neutraal.

Tevens stellen we naar onderwijsniveau statistisch significante verschillen tussen respondenten op. Hoewel de gemiddelde schaalscore voor transformationeel leiderschap voor beide groepen neutraal blijkt, schatten

schoolleiders en leraren uit het lager onderwijs het transformationeel leiderschap binnen de school met betrekking tot feedbackgebruik gemiddeld net iets positiever in dan schoolleiders en leraren uit het secundair onderwijs [gem LO = 3.3 , gem SO = 3.0 ; $t(361.2.0) = 3.4$, $p < 0.01$, $d = 0.34$]. Ook op itemniveau stellen we enkele kleine verschillen vast. Zo valt het op dat respondenten uit het lager onderwijs gemiddeld aangeven gematigd positief te zijn over de stellingen dat de schoolleiding op de school het gebruik van de feedback aanmoedigt [gem LO = 3.5 , gem SO = 3.0 ; $t(380.6) = 4.2$, $p < 0.01$, $d = 0.42$] en de feedback bespreekt met de teamleden [gem LO = 3.5 , gem SO = 3.0 ; $t(416.0) = 4.3$, $p < 0.01$, $d = 0.44$], terwijl respondenten uit het secundair onderwijs beide stelling gemiddeld als gematigd negatief beoordeelden. Respondenten uit zowel het lager onderwijs als – in mindere mate – uit het secundair onderwijs zijn gemiddeld neutraal over de stelling dat de schoolleiding op de school ervoor zorgt dat leraren voldoende vaardigheden hebben om met de feedback aan de slag te kunnen gaan [gem LO = 3.2 , gem SO = 3.0 ; $t(381.3) = 2.1$, $p = 0.04$, $d = 0.24$]. Daarenboven zijn schoolleiders en leraren uit het lager onderwijs het gemiddeld eerder eens dat de schoolleiding op de school leraren coacht bij het gebruik van feedback, terwijl schoolleiders en leraren uit het secundair onderwijs hierover gemiddeld eerder neutraal staan [gem LO = 3.0 , gem SO = 2.8 ; $t(432.0) = 2.4$, $p = 0.02$, $d = 0.24$].

Tot slot constateren we op itemniveau een statistisch significant verschil naar onderwijsnet. Zo stellen we een klein verschil vast tussen respondenten uit het vrij gesubsidieerd onderwijs en het officieel gesubsidieerd onderwijs. Respondenten uit het officieel gesubsidieerd onderwijs zijn immers – hoewel nog steeds een neutrale inschatting – in meerdere mate dan respondenten uit het vrij gesubsidieerd onderwijs positief over de stelling dat de schoolleiding op de school het gebruik van de feedback aanmoedigt [gem VG = 3.1 , gem OG = 3.6 ; $F(3.6, 1.2) = 3.1$, $p = 0.05$, $eta^2 = 0.02$].

Tabel 37
Transformationeel leiderschap met betrekking tot feedbackgebruik

	N WN/ NVT	Gem (SD)	T-test Gem SL		L	LO		SO	Anova Gem VG	G	OG	
De schoolleiding op onze school moedigt het gebruik van de feedback uit centrale toetsen aan.	416 33.4%	3.2 (1.1)	3.4	►	3.1	3.5	►	3.0	3.1	3.3	3.6	VG◄OG
De schoolleiding op deze school bespreekt de feedback uit centrale toetsen met de teamleden.	418 33.1%	3.1 (1.2)	3.5	►	3.0	3.5	►	3.0	3.1	3.0	3.3	
De schoolleiding op deze school zorgt ervoor dat leraren voldoende vaardigheden hebben om met de feedback uit centrale toetsen aan de slag te kunnen gaan.	456 27.0%	3.1 (1.1)	3.3	►	2.9	3.2	►	3.0	3.0	3.0	3.0	

De schoolleiding op deze school coacht leraren bij het gebruik van de feedback uit centrale toetsen.	434 30.1%	2.9 (1.1)	3.3	►	2.7	3.0	►	2.8	2.8	2.7	2.9
De schoolleiding op deze school heeft oog voor de informatiebehoeften van leraren.	523 16.3%	3.4 (1.1)	3.9	►	3.3	3.4		3.5	3.5	3.1	3.3
Schaal: Transformationeel leiderschap (α = 0.91)	401 26.4%	3.1 (0.9)	3.5	►	3.0	3.3	►	3.0	3.1	3.0	3.2

5.3.3 Schoolexterne verwachtingen met betrekking tot feedbackgebruik

In contextfactoren die feedback kunnen versterken of belemmeren richten we ons op het ervaren van schoolexterne verwachtingen ten aanzien van feedbackgebruik uit centrale toetsen. In de vragenlijst werden een aantal items opgenomen om dit concept te meten (zie Tabel 38). Respondenten geven aan dat externe verwachtingen bij feedbackgebruik op schaalniveau (gem = 1.9, SD = 0.9) voor hen helemaal niet duidelijk zijn. Ook op itemniveau schatten respondenten het gemiddeld negatief in dat verwachtingen van de pedagogische begeleidingsdienst (gem = 1.9, SD = 1.0) en van de scholengemeenschap (gem = 1.9, SD = 1.0) bij feedbackgebruik duidelijk zijn. Respondenten zijn het bovendien gemiddeld eerder oneens met de stellingen dat verwachtingen van de overheid (gem = 2.0, SD = 1.0), van de onderwijsinspectie (gem = 2.0, SD = 1.0) en van het schoolbestuur (gem = 2.0, SD = 1.0) bij feedbackgebruik duidelijk zijn.

Als we de verschillen naar functie bekijken, stellen we vast op itemniveau enkele statistisch significante verschillen tussen schoolleiders en leraren vast. Zo beoordelen schoolleiders de duidelijkheid aan verwachtingen bij feedbackgebruik van de overheid [gem SL = 2.2, gem L = 1.9 ; $t(555) = 2.4$, $p = 0.01$, $d = 0.24$], van de onderwijsinspectie [gem SL = 2.3, gem L = 1.9 ; $t(193.6) = 3.1$, $p < 0.01$, $d = 0.33$] en van het schoolbestuur [gem SL = 2.1, gem L = 1.9 ; $t(551.0) = 2.1$, $p = 0.04$, $d = 0.20$] gemiddeld minder negatief dan leraren. Ook wat betreft de verschillen naar onderwijsniveau stellen we statistisch significante verschillen vast. Op schaalniveau geven respondenten uit het lager onderwijs (gem = 2.3) aan dat externe verwachtingen ten aanzien van feedbackgebruiker voor hen gemiddeld eerder onduidelijk zijn, terwijl dergelijke verwachtingen voor respondenten uit het secundair onderwijs gemiddeld helemaal onduidelijk zijn (gem = 1.8) [$t(263.4) = 5.7$, $p < 0.01$, $d = 0.55$]. Ook op itemniveau stellen we tussen beide onderwijsniveaus statistisch significante verschillen vast. Zo zijn respondenten uit het lager onderwijs (gem = 2.4) het gemiddeld meer eens dan respondenten uit het secundair onderwijs (gem = 1.8) met de stelling dat de verwachtingen van de onderwijsinspectie met betrekking tot feedbackgebruik duidelijk is [$t(255.4) = 6.2$, $p < 0.01$, $d = 0.60$] en beoordelen respondenten uit het lager onderwijs (gem = 2.2) de duidelijkheid van verwachtingen met betrekking tot feedbackgebruik van de scholengemeenschap gemiddeld positiever dan respondenten uit het secundair onderwijs (gem = 1.8) [$t(262.6) = 5.2$, $p < 0.01$, $d = 0.50$]. Daarenboven geven

respondenten uit het lager onderwijs gemiddeld een kleine positievere inschatting van de duidelijkheid in de verwachtingen van het schoolbestuur [gem SL = 2.3, gem L = 1.9 ; $t(554) = 4.4$, $p < 0.01$, $d = 0.40$], van de overheid [gem SL = 2.3, gem L = 1.9 ; $t(271.5) = 5.1$, $p < 0.01$, $d = 0.49$] en van de pedagogische begeleidingsdienst [gem SL = 2.3, gem L = 1.8 ; $t(555.0) = 5.4$, $p < 0.01$, $d = 0.49$] bij feedbackgebruik dan respondenten uit het secundair onderwijs. Met betrekking tot het ervaren van duidelijke externe verwachtingen bij feedbackgebruik stellen we geen statistisch significante verschillen tussen de netten vast.

Tabel 38
Schoolexterne verwachtingen bij feedbackgebruik

	N WN/ NVT	Gem (SD)	T-test Gem SL		L	LO		SO	Anova Gem VG	G	OG
Omtrent het gebruik van de feedback uit centrale toetsen, is het voor mij duidelijk:											
wat de overheid van mij verwacht.	560 8.9%	2.0 (1.0)	2.2	►	1.9	2.3	►	1.9	1.9	2.2	2.2
wat de onderwijsinspectie van mij verwacht.	557 9.4%	2.0 (1.0)	2.3	►	1.9	2.4	►	1.8	1.9	2.2	2.0
wat de pedagogische begeleidingsdienst van mij verwacht.	557 9.4%	1.9 (1.0)	2.1		1.9	2.3	►	1.8	1.9	2.1	2.0
wat de scholengemeenschap van mij verwacht.	556 9.6%	1.9 (1.0)	2.0		1.9	2.2	►	1.8	1.8	2.0	2.1
wat het schoolbestuur van mij verwacht.	556 9.6%	2.0 (1.0)	2.1	►	1.9	2.3	►	1.9	1.9	2.1	2.1
Schaal: Schoolexterne verwachtingen met betrekking tot feedbackgebruik (α = 0.98)	552 8.6%	1.9 (0.9)	2.1		1.9	2.3	►	1.8	1.9	2.1	2.1

5.3.4 Gebruikersfactoren met betrekking tot feedbackgebruik

Bij de conceptuele uitgangspunten onderscheiden we voor gebruikersfactoren die feedbackgebruik kunnen faciliteren en/of belemmeren de attitude ten aanzien van feedbackgebruik, de motivatie ten aanzien van feedbackgebruik, de persoonlijke doelmatigheidsbeleving ten aanzien van feedbackgebruik en bekommernissen ten aanzien van feedbackgebruik. We beschrijven hieronder de inschatting van de aanwezigheid ervan door schoolleiders en leraren.

5.3.4.1 Attitude ten aanzien van feedbackgebruik

Met betrekking tot de attitude ten aanzien van feedbackgebruik onderscheiden we een affectieve en een cognitieve component (zie Tabel 39). We stellen vast dat de affectieve attitude ten aanzien van feedbackgebruik uit centrale toetsen van schoolleiders en leraren over het algemeen neutraal is (gem = 3.0, SD = 0.6). Hierbij staan respondenten eveneens gemiddeld neutraal tegenover de stellingen dat ze de komst van centrale toetsen toejuichen (gem = 3.2, SD = 1.1) en de komst van centrale toetsen hen niet afschrikt (gem = 3.4, SD = 1.1). De stellingen dat de komst van centrale toetsen respondenten enthousiast maakt (gem = 2.8, SD = 1.0) en dat ze de komst van centrale toetsen *niet* betreuren (gem = 2.8, SD = 1.1) beoordelen schoolleiders en leraren als eerder oneens. De cognitieve attitude ten aanzien van feedbackgebruik uit centrale toetsen blijkt op basis van de schaalscore gematigd negatief (gem = 2.9, SD = 0.7). Zo stellen we op itemniveau vast dat respondenten aangeven gemiddeld eerder oneens te zijn met de stellingen dat de feedback op klasniveau waardevolle informatie is om als input te gebruiken voor functionerings- en evaluatiegesprekken van leraren (gem = 2.3, SD = 1.1) en waardevolle informatie is om als input te gebruiken voor functionerings- en evaluatiegesprekken van schoolleiders (gem = 2.4, SD = 1.1). Daarnaast geven respondenten gemiddeld aan eerder liever niet te vertrouwen op de feedback dan op de toetsen die momenteel reeds gebruik worden om leerlingen te evalueren (gem = 2.6, SD = 1.0) en vinden schoolleiders en leraren gemiddeld dat de feedback eerder niet veel toevoegt aan wat ze intuïtief aanvoelen met betrekking tot de leerprestaties van leerlingen (gem = 2.7, SD = 1.0). Verder zijn respondenten het er gemiddeld eveneens eerder mee oneens dat centrale toetsen waardevol zullen zijn om tot beter lesgeven te komen (gem = 2.9, SD = 1.1), dat centrale toeten veel zullen leren over de onderwijsaanpak van leraren in de school (gem = 2.9, SD = 1.1) en dat centrale toetsen *niet* aanvoelen als een vorm van controle (gem = 2.8, SD = 1.2) en geven ze aan eerder geen voorstander te zijn van het invoeren van het systeem van centrale toetsen (gem = 2.9, SD = 1.2).

Daarnaast stellen we vast dat schoolleiders en leraren de stellingen dat centrale toetsen de autonomie *niet* bedreigen (gem = 3.1, SD = 1.2), dat centrale toetsen een meerwaarde zijn voor de goede werking van de school (gem = 3.1, SD = 1.1) en dat centrale toetsen veel zullen leren over de leerprestaties van leerlingen (gem = 3.3, SD = 1.0) als neutraal beoordelen. Ondanks de gematigd negatieve tot neutrale beoordeling van deze stellingen geven schoolleiders en leraren tegelijkertijd aan de feedback uit centrale eerder wel te gaan gebruiken (gem = 3.7, SD = 0.8)..

Met betrekking tot de affectieve attitude ten aanzien van feedbackgebruik stellen we geen statistisch significante verschillen vast tussen schoolleiders en leraren. Met betrekking tot de cognitieve attitude ten aanzien van feedbackgebruik stellen we wel statistisch significante verschillen tussen schoolleiders en

leraren vast. De gemiddelden van schoolleiders liggen daarbij telkens hoger dan deze van leraren. Op basis van het schaalgemiddelde voor de cognitieve houding ten aanzien van feedbackgebruik stellen we een neutrale houding vast bij schoolleiders en een gematigd negatieve houding bij leraren [gem SL = 3.0, gem L = 2.9 ; $t(571.0) = 2.7$, $p < 0.01$, $d = 0.26$]. Op itemniveau geven schoolleiders gemiddeld aan een neutrale houding te hebben met betrekking tot de stelling 'de feedback uit centrale toetsen voegt veel toe aan wat ik intuïtief aanvoel met betrekking tot de leerprestaties van leerlingen', terwijl leraren deze stelling gemiddeld eerder als oneens beantwoordden [gem SL = 3.0, gem L = 2.6 ; $t(521.0) = 3.5$, $p < 0.01$, $d = 0.37$]. Bovendien blijkt dat leraren gemiddeld in meerdere mate een negatieve inschatting dan schoolleiders maken voor de stellingen dat de feedback op klasniveau waardevolle informatie is om als input te gebruiken voor functionerings- en evaluatiegesprekken van leraren [gem SL = 2.6, gem L = 2.3 ; $t(560.0) = 3.0$, $p < 0.01$, $d = 0.30$], dat de feedback uit centrale toetsen op schoolniveau waardevolle informatie is om als input te gebruiken voor functionerings- en evaluatiegesprekken van schoolleiders [gem SL = 2.7, gem L = 2.3 ; $t(557.0) = 3.8$, $p < 0.01$, $d = 0.38$] en dat men liever vertrouwt op de feedback uit centrale toetsen dan op de toetsen die momenteel al gebruikt worden om leerlingen te evalueren [gem SL = 2.8, gem L = 2.5 ; $t(557.0) = 2.8$, $p < 0.01$, $d = 0.28$]. Ook de stelling 'ik ga de feedback uit centrale toetsen gebruiken' beoordelen schoolleiders gemiddeld positiever dan leraren [gem SL = 3.9, gem L = 3.7 ; $t(536.0) = 2.4$, $p = 0.02$, $d = 0.24$].

Wanneer we de verschillen naar onderwijsniveau bekijken stellen we op schaalniveau vast dat respondenten uit het lager onderwijs gemiddeld een neutrale, maar wel een iets positievere affectieve attitude [gem LO = 3.1, gem SO = 3.0 ; $t(506.3) = 3.2$, $p < 0.01$, $d = 0.24$] en cognitieve attitude [gem LO = 3.3, gem SO = 3.0 ; $t(517.3) = 4.5$, $p < 0.01$, $d = 0.35$] ten aanzien van feedbackgebruik rapporteren dan respondenten uit het secundair onderwijs. Dit verschil weerspiegelt zich eveneens in de gemiddelde scores op itemniveau. Met betrekking tot de affectieve attitude valt op dat respondenten uit het lager onderwijs de komst van centrale toetsen meer toejuichen dan respondenten uit het secundair onderwijs [gem LO = 3.4, gem SO = 3.2 ; $t(502.8) = 2.7$, $p < 0.01$, $d = 0.20$], dat de komst van centrale toetsen hen minder sceptisch maakt dan respondenten uit het secundair onderwijs [gem LO = 3.2, gem SO = 2.8 ; $t(492.0) = 3.9$, $p < 0.01$, $d = 0.30$] en dat ze het er in mindere mate dan respondenten uit het secundair onderwijs mee eens zijn de komst van centrale toetsen niet te betreuren [gem LO = 2.6, gem SO = 2.8 ; $t(496.28) = -2.7$, $p < 0.01$, $d = -0.21$]. Met betrekking tot de cognitieve attitude stellen we vast dat respondenten uit het lager onderwijs aangeven meer overtuigd te zijn dan respondenten uit het secundair onderwijs dat centrale toetsen een meerwaarde betekenen voor de goede werking van de school [gem LO = 3.3, gem SO = 3.0 ; $t(500.2) = 3.7$, $p < 0.01$, $d = 0.28$] en dat ze positiever zijn dan respondenten uit het secundair onderwijs met betrekking tot de waarde van centrale toetsen om tot beter lesgeven te komen [gem LO = 3.1, gem SO = 2.8 ; $t(485.5) = 3.5$, $p < 0.01$, $d = 0.27$]. Verder

geven respondenten uit het lager onderwijs aan dat ze centrale toetsen minder als autonomiebedreigend ervaren dan respondenten uit het secundair onderwijs [gem LO = 3.6, gem SO = 3.0 ; $t(499.4) = 7.0$, $p < 0.01$, $d = 0.53$]. Dit verschil tussen beide groepen betreft een medium effect. Tot slot stellen we ook vast dat er geen (wezenlijke) verschillen met betrekking tot de attituden ten aanzien van feedbackgebruik tussen de onderwijsnetten zijn. Zowel de affectieve attitude als de cognitieve attitude ten aanzien van feedbackgebruik van respondenten uit het vrij gesubsidieerd onderwijs, het gemeenschapsonderwijs of het officieel gesubsidieerd onderwijs is met andere woorden hetzij neutraal, hetzij gematigd negatief.

Tabel 39
Attitude ten aanzien van feedbackgebruik

	N WN/ NVT	Gem (SD)	T-test Gem SL	L	LO		SO	Anova Gem VG	G	OG
Ik juich de komst van centrale toetsen toe.	784 1.5%	3.2 (1.1)	3.1	3.3	3.4	►	3.2	3.2	3.3	3.5
De komst van centrale toetsen maakt me enthousiast.	789 0.9%	2.8 (1.0)	2.7	2.9	3.0		2.8	2.8	2.9	3.0
Ik betreur de komst van centrale toetsen *niet**.	779 2.1%	2.8 (1.1)	2.9	2.7	2.6	◄	2.8	2.8	2.7	2.5
De komst van centrale toetsen maakt me *niet** sceptisch.	778 2.3%	2.9 (1.1)	2.8	3.0	3.2	►	2.8	2.9	2.8	3.2
De komst van centrale toetsen schrikt me *niet** af.	787 1.1%	3.4 (1.1)	3.4	3.3	3.5		3.3	3.3	3.4	3.5
Schaal: Affectieve attitude (α = 0.85)	781 0.8%	3.0 (0.6)	3.3	3.0	3.1	►	3.0	3.0	3.0	3.1
Centrale toetsen zullen een meerwaarde zijn voor de goede werking van de school.	767 3.6%	3.1 (1.1)	3.1	3.1	3.3	►	3.0	3.1	3.1	3.4
Centrale toetsen zullen waardevol zijn om tot beter lesgeven te komen.	770 3.3%	2.9 (1.1)	3..0	2.9	3.1	►	2.8	2.8	2.8	3.3
Centrale toetsen zullen mij veel leren over de leerprestaties van de leerlingen.	767 2.5%	3.3 (1.0)	3.4	3.3	3.4		3.3	3.3	3.3	3.5
Centrale toetsen zullen mij veel leren over de onderwijsaanpak van leraren in de school.	769 3.4%	2.9 (1.1)	3.0	2.9	2.9		2.9	2.9	3.0	3.2
Centrale toetsen zullen mijn autonomie *niet** bedreigen.	769 3.4%	3.1 (1.2)	3.2	3.1	3.6	►	3.0	3.1	3.2	3.4
Centrale toetsen voelen voor mij *niet** aan als een vorm van controle.	782 1.8%	2.8 (1.2)	2.8	2.8	3.0		2.7	2.8	2.8	3.1

De feedback uit centrale toetsen op klasniveau is waardevolle informatie om als input te gebruiken voor functionerings- en evaluatiegesprekken van leraren.	566 5.8%	2.3 (1.1)	2.6	►	2.3	2.3		2.4	2.3	2.4	2.2
De feedback uit centrale toetsen op schoolniveau is waardevolle informatie om als input te gebruiken voor functionerings- en evaluatiegesprekken van schoolleiders.	653 6.3%	2.4 (1.1)	2.7	►	2.3	2.3		2.4	2.3	2.4	2.2
De feedback uit centrale toetsen voegt *veel* * toe aan wat ik intuïtief aanvoel met betrekking tot de leerprestaties van leerlingen.	527 12.3%	2.7 (1.0)	3.0	►	2.6	2.8		2.6	2.6	2.6	2.8
Ik vertrouw liever op *de feedback uit centrale toetsen dan op de toetsen die wij momenteel al gebruiken om leerlingen te evalueren** .	563 6.5%	2.6 (1.0)	2.8	►	2.5	2.7		2.6	2.5	2.6	2.7
Ik ga de feedback uit centrale toetsen gebruiken.	540 10.1%	3.7 (0.8)	3.9	►	3.7	3.8		3.7	3.7	3.8	3.9
Ik ben een voorstander van het invoeren van het systeem van centrale toetsen.	565 5.8%	2.9 (1.2)	2.9		2.9	3.1		2.9	2.9	3.0	3.1
Schaal: Cognitieve attitude (α = 0.74)	576 0.3%	2.9 (0.7)	3.0	►	2.9	3.0	►	2.8	2.9	2.9	3.0

5.3.4.2 Motivatie ten aanzien van feedbackgebruik

Op basis van de conceptuele uitgangspunten worden vijf types van motivatie onderscheiden, waarbij de beschrijving van deze types vertrekt vanuit het onderscheid tussen autonome en gecontroleerde motivatie enerzijds en amotivatie anderzijds. Autonoom gemotiveerde feedbackgebruikers 'willen' met feedback aan de slag gaan en gecontroleerd gemotiveerde feedbackgebruikers 'moeten' met feedback aan de slag gaan. In beide gevallen is er sprake van motivatie. Geamotiveerde feedbackgebruikers kenmerken zich in een afwezigheid van motivatie voor feedbackgebruik.

Op schaalniveau blijkt dat schoolleiders en leraren in het algemeen neutraal zijn in hun inschatting in de mate waarin ze zich autonoom gemotiveerd voelen (gem = 3.4, SD = 0.7) voor feedbackgebruik (zie Tabel 40). Hun inschatting van de mate waarin ze gecontroleerd gemotiveerd (gem = 2.5, SD = 0.8) en geamotiveerd (gem= 2.5, SD = 0.9) zijn voor feedbackgebruik is dan weer gematigd negatief. Wat betreft de autonome motivatie voor feedbackgebruik zijn respondenten het gemiddeld eerder eens met de stellingen met feedback aan de slag te gaan om te willen bijleren (gem = 3.7, SD = 0.9), met feedback aan de slag te gaan als een weloverwogen keuze (gem = 3.7, SD = 0.9) en met feedback aan de slag te gaan om beter zicht te krijgen op de leerprestaties van leerlingen (gem = 3.8, SD = 0.9). Tevens stellen we vast dat schoolleiders

en leraren aangeven gemiddeld eerder wel met de feedback aan de slag te gaan omdat ze dat een interessante activiteit vinden (gem = 3.6, SD = 0.9), terwijl ze tegelijkertijd aangeven gemiddeld eerder niet plannen met feedback aan de slag te gaan omdat ze dat een aangename bezigheid vinden (gem = 2.3, SD = 1.0). Op de items die gecontroleerde motivatie meten, antwoorden respondenten overwegend gemiddeld gematigd negatief. Zo zijn respondenten het eerder oneens met feedback aan de slag te gaan omdat ze willen dat anderen hen een waardevol lid van het team vinden (gem = 2.8, SD = 1.1) of omdat ze zich schuldig zouden voelen als ze niet met de feedback aan de slag zouden gaan (gem = 2.8, SD = 1.2). Daarenboven zijn respondenten het in nog meerdere mate oneens met feedback aan de slag te gaan omdat ze zich zouden schamen als ze het niet zouden doen (gem = 2.2, SD = 1.1), omdat ze anderen de indruk willen geven een goede collega/schoolleider te zijn (gem = 2.2, SD = 1.0) of omdat anderen anders teleurgesteld in hen zullen zijn (gem = 2.2, SD = 1.0). Ten aanzien van de inschatting om met feedback aan de slag te gaan omdat dit verondersteld wordt te doen zijn respondenten neutraal (gem = 3.0, SD = 1.1). Ook op de items bij amotivatie antwoorden respondenten in het algemeen gematigd negatief. Respondenten zijn het eerder oneens met de stelling 'de reden om met de feedback aan de slag te gaan, zijn voor mij niet duidelijk' (gem = 2.6, SD = 1.0), 'bij het aan de slag gaan met de feedback ga ik het gevoel hebben mijn tijd te verdoen' (gem = 2.5, SD 1.0) en 'ik zie niet in welk verschil het gebruik van feedback zou maken' (gem = 2.5, SD = 1.1).

Wanneer we verschillen naar functie bekijken, stellen we vast dat de gemiddelde beoordeling van alle gehanteerde schalen statisch significant verschillend zijn tussen schoolleiders en leraren. Zo geven schoolleiders aan eerder wel autonoom gemotiveerd te zijn voor feedbackgebruik, terwijl leraren hiervoor een neutrale inschatting maken [gem SL = 3.7, gem L = 3.3 ; $t(663.0) = 6.1$, $p < 0.01$, $d = 0.58$] en geven leraren aan in meerdere mate gecontroleerd gemotiveerd te zijn voor feedbackgebruik dan leraren [gem L = 2.6, gem SL = 2.3 ; $t(659.0) = -4.6$, $p < 0.01$, $d = -0.46$]. Daarenboven maken schoolleiders een sterkere gematigde negatieve inschatting van de mate waarin men geamotiveerd is voor feedbackgebruik dan leraren [gem SL = 2.2, gem L = 2.6 ; $t(633.0) = -4.3$, $p < 0.01$, $d = -0.41$]. Met betrekking tot intrinsieke regulatie voor autonome motivatie geven schoolleiders gemiddeld aan aan de slag gaan met feedback boeiender te vinden dan leraren [gem SL = 3.8, gem L = 3.3 ; $t(237.9) = -5.7$, $p < 0.01$, $d = -0.54$]. Daarnaast is er een klein verschil tussen schoolleiders en leraren voor de stelling 'ik vind de feedback interessant om mee aan de slag te gaan'. Schoolleiders en – in mindere mate – leraren zijn beiden gematigd positief over deze stelling [gem SL = 3.9, gem L = 3.5 ; $t(659.0) = 4.7$, $p < 0.01$, $d = 0.44$]. Met betrekking tot de items die geïdentificeerde regulatie bij autonome motivatie meten, geven zowel schoolleiders als – in mindere mate – leraren aan gemiddeld gematigd positief te zijn. Zo beoordelen schoolleiders de stellingen 'ik plan met de feedback aan de slag te gaan, want ik wil beter zicht krijgen op leerprestaties van leerlingen' [gem SL = 4.0, gem L = 3.7 ; $t(674.0) =$

4.4, p <0.01 , d = 0.42], 'ik plan met de feedback aan de slag te gaan omdat ik wil bijleren' [gem SL = 4.0, gem L = 3.7 ; t(273.1) = 4.7, p <0.01 , d = 0.42] en 'aan de slag gaan met de feedback uit centrale toetsen zal voor mij een weloverwogen keuze zijn' [gem SL = 4.0, gem L = 3.6 ; t(677.0) = -5.5, p <0.01 , d = 0.52] positiever dan leraren. Ook met betrekking tot de items die gecontroleerde motivatie meten, stellen we kleine statistisch significante verschillen tussen schoolleiders en leraren op itemniveau vast. Nu zijn het echter leraren die vanuit een gemiddeld gematigd negatieve of neutrale houding de betreffende items toch positiever percipiëren dan schoolleiders. Zo geven leraren aan gematigd negatief te zijn betreffende de stellingen 'ik plan met de feedback aan de slag te gaan omdat ik anderen de indruk wil geven een goede collega te zijn' [gem L = 2.3, gem SL = 1.9 ; t(665.0) = -3.5, p <0.01 , d = -0.33] en 'ik plan met de feedback aan de slag te gaan omdat anderen anders teleurgesteld in hen zouden zijn [gem L = 2.3, gem SL = 1.9 ; t(668.0) = -4.3, p <0.01 , d = -0.41], terwijl schoolleiders beide stellingen negatief inschatten. Daarnaast blijken leraren een neutrale houding en schoolleiders een gematigd negatieve houding te hebben over de stelling 'ik plan met de feedback aan de slag te gaan omdat ik verondersteld wordt dit te doen [gem L = 3.1, gem SL = 2.7 ; t(211.7.0) = -3.5, p <0.01 , d = -0.34]. Hoewel beide groepen het eerder eens zijn met de stellingen feedbackgebruik te plannen omdat ze zich zouden schamen als ze het niet zouden doen [gem L = 2.3, gem SL = 2.0 ; t(617.0) = -3.6, p <0.01 , d = -0.33] en omdat ze willen dat anderen hen een waardevol lid van het team vinden [gem L = 2.9, gem SL = 2.4 ; t(664.0) = -4.4, p <0.01 , d = -0.42], schatten leraren beide items toch iets positiever in dan schoolleiders. Wat betreft amotivatie ten aanzien van feedbackgebruik maken schoolleiders een gemiddeld sterkere gematigd negatieve beoordeling van de stellingen 'de redenen om met de feedback aan de slag te gaan zijn niet duidelijk voor mij' [gem SL = 2.3, gem L = 2.6 ; t(245.2) = -3.3, p <0.01 , d = -0.30] en 'ik zie niet in welk verschil het gebruik van de feedback zou maken' [gem SL = 2.2, gem L = 2.5 ; t(663.0) = -2.6, p <0.01 , d = -0.23]. Daarnaast geven schoolleiders in (nog) sterkere mate dan leraren aan het oneens te zijn dat ze bij het aan de slag gaan met feedback het gevoel gaan hebben de tijd te verdoen [gem SL = 2.1, gem L = 2.6 ; t(656.0) = -5.6, p <0.01 , d = -0.53].

Met betrekking tot verschillen naar onderwijsniveau, stellen we voornamelijk statistisch significante verschillen vast voor autonome motivatie ten aanzien van feedbackgebruik. Op schaalniveau geven respondenten uit het lager onderwijs vanuit een neutrale houding gemiddeld aan autonomer gemotiveerd te zijn dan respondenten uit het secundair onderwijs om met feedback aan de slag te gaan [gem LO = 3.5, gem SO = 3.4 ; t(668.0) = 2.3 ,p = 0.02, d = 0.20]. Op itemniveau vertaalt zich dat in een gemiddeld positievere beoordeling van respondenten uit het lager onderwijs van de stellingen feedbackgebruik uit centrale toetsen boeiender [gem LO = 3.6, gem SO = 3.3 ; t(433.0) = 3.8 ,p <0.01, d = 0.31] en interessanter [gem LO = 3.8, gem SO = 3.6 ; t(433.3) = 2.6 ,p <0.01, d = 0.21] te vinden dan respondenten uit het secundair onderwijs.

Wat betreft verschillen tussen de onderwijsnetten stellen we geen statistisch significante verschillen vast met ten minste een klein effect.

Tabel 40
Motivatie ten aanzien van feedbackgebruik

	N WN/ NVT	Gem (SD)	T-test Gem SL		L	LO		SO	Anova Gem VG	G	OG
Intrinsieke regulatie											
Ik vind de feedback uit centrale toetsen erg interessant om mee aan de slag te gaan.	666 5.7%	3.6 (0.9)	3.9	►	3.5	3.8	►	3.6	3.6	3.5	3.8
Ik plan met de feedback uit centrale toetsen aan de slag te gaan omdat ik dat een aangename bezigheid vind.	681 3.5%	2.3 (1.0)	2.5		2.2	2.4		2.2	2.2	2.4	2.5
Ik vind aan de slag gaan met de feedback uit centrale toetsen boeiend.	654 7.4%	3.4 (1.0)	3.8	►	3.3	3.6	►	3.3	3.3	3.3	3.6
Geïdentificeerde regulatie											
Ik plan met de feedback uit centrale toetsen aan de slag te gaan, want ik wil beter zicht krijgen op de leerprestaties van de leerlingen.	681 3.5%	3.8 (0.9)	4.0	►	3.7	3.8		3.7	3.7	3.7	3.9
Aan de slag gaan met de feedback uit centrale toetsen zal voor mij een weloverwogen keuze zijn.	684 3.1	3.7 (0.9)	4.0	►	3.6	3.7		3.7	3.6	3.6	3.7
Ik plan met de feedback uit centrale toetsen aan de slag te gaan omdat ik wil bijleren.	682 3.4%	3.7 (0.9)	4.0	►	3.7	3.8		3.7	3.7	3.7	4.0
Schaal: Autonome motivatie (α = 0.85)	670 1.8%	3.4 (0.7)	3.7	►	3.3	3.5	►	3.4	3.4	3.4	3.6
Geïntrojecteerde regulatie											
Ik zou me schuldig voelen als ik niet met de feedback uit centrale toetsen aan de slag zou gaan.	684 3.2%	2.8 (1.2)	2.6		2.8	2.9		2.7	2.8	2.8	2.7
Ik plan met de feedback uit centrale toetsen aan de slag te gaan omdat ik me zou schamen als ik het niet zou doen.	680 3.7%	2.2 (1.1)	2.0	◄	2.3	2.2		2.2	2.3	2.4	2.2
Ik plan met de feedback uit centrale toetsen aan de slag te gaan omdat ik anderen de indruk wil geven dat ik een goede collega/schoolleider ben.	672 4.8%	2.2 (1.0)	1.9	◄	2.3	2.2		2.2	2.2	2.3	2.3
Ik plan met de feedback uit centrale toetsen aan de slag te gaan, want ik wil dat anderen mij een waardevol lid van het team vinden.	671 5.0%	2.8 (1.1)	2.4	◄	2.9	2.7		2.8	2.8	2.9	3.0
Externe regulatie											

Ik plan met de feedback uit centrale toetsen aan de slag te gaan omdat ik verondersteld word dit te doen.	684 3.1%	3.0 (1.1)	2.7	◄	3.1	3.0	3.0	3.1	3.0	3.0
Ik plan met de feedback uit centrale toetsen aan de slag te gaan omdat anderen anders teleurgesteld zullen zijn in mij.	675 4.4%	2.2 (1.0)	1.9	◄	2.3	2.2	2.2	2.2	2.3	2.1
Schaal: Gecontroleerde motivatie (α = 0.88)	666 1.8%	2.5 (0.8)	2.3	◄	2.6	2.5	2.5	2.6	2.6	2.5
Geen regulatie										
De redenen om met de feedback uit centrale toetsen aan de slag te gaan, zijn voor mij niet duidelijk.	682 3.4%	2.6 (1.0)	2.3	◄	2.6	2.4	2.6	2.6	2.4	2.3
Bij het aan de slag gaan met de feedback uit centrale toetsen ga ik het gevoel hebben mijn tijd te verdoen.	663 6.1%	2.5 (1.0)	2.1	◄	2.6	2.4	2.5	2.6	2.3	2.3
Ik zie niet in welk verschil het gebruik van de feedback uit centrale toetsen zou maken.	670 5.1%	2.5 (1.1)	2.2	◄	2.5	2.3	2.5	2.5	2.4	2.3
Schaal: Amotivatie (α = 0.69)	640 1.4%	2.5 (0.9)	2.2	◄	2.6	2.4	2.5	2.6	2.4	2.3

5.3.4.3 Persoonlijke doelmatigheidsbeleving ten aanzien van feedbackgebruik

De persoonlijke doelmatigheid van schoolleiders en leraren inzake feedbackgebruik beschrijft de mate waarin de respondenten zich bekwaam achten om met de feedback aan de slag te gaan (zie Tabel 41). Respondenten beoordelen op schaalniveau de persoonlijke doelmatigheid met betrekking tot verkennen en overleggen als gemiddeld neutraal (gem = 3.1, SD = 0.8), met betrekking tot interpreteren als gemiddeld positief (gem = 4.0, SD = 0.7) en met betrekking tot diagnosticeren als gemiddeld gematigd positief (gem = 3.8, SD = 0.8). Ook het item dat de bekwaamheid met betrekking tot het bedenken van gepaste acties op basis van verklaringen voor verschillen in toetsresultaten beoordelen schoolleiders en leraren gemiddeld gematigd positief (gem = 3.8, SD = 0.8). Wat betreft de persoonlijke doelmatigheid met betrekking tot verkennen en overleggen geven schoolleiders en leraren op itemniveau aan dat ze zich over het algemeen gemiddeld eerder wel bekwaam voelen om met leerlingen over de feedback te communiceren (gem = 3.7, SD = 0.9) en om de feedback met collega's te bespreken (gem = 3.7, SD = 0.9). Tegelijkertijd geven respondenten gemiddeld aan eerder (nog) geen goed zicht te hebben op de bruikbaarheid van de feedback (gem = 2.5, SD = 1.1) en op wat ze uit de feedback kunnen leren (gem = 2.6, SD = 1.1). De items die persoonlijke doelmatigheid met betrekking tot analyseren meten, worden door de respondenten allemaal als gemiddeld positief beoordeeld. Respondenten vinden zichzelf (eerder) bekwaam om grafieken te begrijpen (gem = 4.2, gem = 0.7), om tabellen te begrijpen (gem = 4.2, gem = 0.7), om statistische gegevens te interpreteren (gem = 4.0, SD = 0.9), om centrale toetsresultaten van leerlingen te interpreteren (gem = 3.9, SD = 0.9) en om de

juiste vaststellingen uit de feedback te halen (gem = 3.6, SD = 0.9). Ook de stellingen over de persoonlijke doelmatigheid met betrekking tot diagnose beoordelen respondenten gemiddeld eerder positief. Zo geven ze aan zichzelf eerder bekwaam te vinden om mogelijke verklaringen voor verschillen in centrale toetsresultaten verder te onderzoeken aan de hand van bijkomende informatie (gem = 3.7, SD = 0.8) en om verschillen in centrale toetsresultaten te verklaren (gem = 3.9, SD = 0.8).

Wanneer we naar verschillen tussen schoolleiders en leraren kijken, stellen we slechts een enkel statistisch significant verschil vast. Leraren blijken zichzelf gemiddeld bekwamer dan schoolleiders in te schatten voor het interpreteren van centrale toetsresultaten van leerlingen [gem L = 3.9, gem SL = 3.7 ; $t(198.2) = -2.3$, $p = 0.02$, $d = -0.23$]. Daarnaast stellen we vast dat respondenten uit het lager onderwijs de persoonlijke doelmatigheid met betrekking tot verkennen en overleggen gemiddeld beperkt positiever inschatten dan respondenten uit het secundair onderwijs [gem LO = 3.3, gem SO = 3.1 ; $t(564.0) = 2.3$, $p = 0.02$, $d = 0.22$]. Ook op itemniveau stellen we hieromtrent kleine gemiddelde verschillen vast tussen respondenten uit het lager onderwijs en respondenten uit het secundair onderwijs. Respondenten uit het lager onderwijs geven – vanuit een gematigd negatieve inschatting – aan een beter zicht te hebben dan respondenten uit het secundair onderwijs op de bruikbaarheid van de feedback [gem LO = 2.8, gem SO = 2.5 ; $t(576.0) = 4.1$, $p < 0.01$, $d = 0.38$] en op wat de feedback kan leren [gem LO = 2.9, gem SO = 2.6 ; $t(577.0) = 3.5$, $p < 0.01$, $d = 0.32$]. Met betrekking tot verschillen naar onderwijsnet stellen we geen statistisch significante verschillen met ten minste een klein effect vast.

Tabel 41
Persoonlijke doelmatigheidsbeleving ten aanzien van feedbackgebruik

	N WN/ NVT	Gem (SD)	T-test Gem SL	L	LO		SO	Anova Gem VG	G	OG
Ik heb goed zich op:										
de bruikbaarheid van de feedback uit centrale toetsen.	578 9.7%	2.5 (1.1)	2.7	2.5	2.8	►	2.4	2.5	2.7	2.5
wat ik uit de feedback uit centrale toetsen kan leren.	579 9.5%	2.6 (1.1)	2.7	2.6	2.9	►	2.5	2.6	2.7	2.5
Ik vind mezelf bekwaam om:										
met leerlingen over de feedback uit centrale toetsen te communiceren.	590 7.8%	3.7 (0.9)	3.5	3.7	3.6		3.7	3.7	3.8	3.6
de feedback uit centrale toetsen te bespreken met collega's.	604 5.6%	3.7 (0.9)	3.8	3.7	3.8		3.7	3.8	3.8	3.6

Schaal: PDM m.b.t verkennen en overleggen (α = 0.80)	566 3.9%	3.1 (0.8)	3.2		3.1	3.3	►	3.1	3.1	3.3	3.1
de juiste vaststellingen uit de feedback uit centrale toetsen te halen.	575 10.2%	3.6 (0.9)	3.6		3.6	3.6		3.6	3.6	3.6	3.4
toetsresultaten van leerlingen op centrale toetsen te interpreteren.	599 6.4%	3.9 (0.9)	3.7	◄	3.9	4.0		3.8	3.9	3.9	3.7
statistische gegevens te interpreteren.	618 3.4%	4.0 (0.9)	3.9		4.0	4.0		4.0	4.0	4.0	3.7
tabellen te begrijpen.	624 2.5%	4.2 (0.7)	4.1		4.2	4.1		4.2	4.2	4.2	4.0
grafieken te begrijpen.	623 2.7%	4.2 (0.7)	4.1		4.2	4.1		4.2	4.2	4.2	4.0
Schaal: PDM m.b.t interpreteren (α = 0.84)	601 2.6%	4.0 (0.7)	3.9		4.0	4.0		4.0	4.0	4.0	3.8
verschillen in toetsresultaten op centrale toetsen te verklaren.	599 6.4%	3.9 (0.8)	3.8		3.9	4.0		3.8	3.9	3.9	3.6
mogelijke verklaringen voor verschillen in toetsresultaten op centrale toetsen aan de hand van bijkomende informatie verder te onderzoeken.	593 7.3%	3.7 (0.8)	3.8		3.7	3.7		3.7	3.7	3.8	3.6
Schaal: PDM m.b.t diagnose (α = 0.76)	580 4.6%	3.8 (0.8)	3.8		3.8	4.0		4.0	3.8	3.8	3.6
o.b.v. verklaringen voor verschillen in toetsresultaten op centrale toetsen gepaste acties te bedenken.	600 6.3%	3.7 (0.8)	3.8		3.7	3.7		3.7	3.7	3.7	3.8

5.3.4.4 Bekommernissen ten aanzien van feedbackgebruik

Om de bereidheid tot feedbackgebruik in kaart te brengen, bevroegen we respondenten over aanwezige bekommernissen ten aanzien van feedbackgebruik (zie Tabel 42).

Op schaalniveau geven schoolleiders en leraren aan dat zelf-bekommernissen (gem = 3.6, SD = 0.6) over het algemeen gemiddeld sterker aanwezig zijn andere-bekommernissen (gem = 3..3, SD = 0.6). Wat betreft zelf-bekommernissen op het niveau van bewustwording antwoorden schoolleiders en leraren gemiddeld eerder neutraal op de stellingen op dit moment niet te weten wat het gebruik van feedback kan inhouden (gem = 3.2, SD = 1.2) en het zo druk te hebben dat men momenteel weinig aandacht kan besteden aan de mogelijkheden van feedbackgebruik (gem = 3.4, SD = 1.1). Zelf-bekommernissen op het niveau van informatie blijken gemiddeld het vaakst gerapporteerd door schoolleiders en leraren. Schoolleiders en leraren het

eerder eens met de stellingen de mogelijkheden van feedbackgebruik verder te willen verkennen (gem = 3.8, SD = 0.8) en meer informatie te willen hebben over hoe met feedback aan de slag te kunnen (gem = 4.0, SD = 0.9). Schoolleiders en leraren geven met andere woorden blijk van interesse in bijkomende informatie over feedbackgebruik. Met betrekking tot taak-bekommernissen valt op dat schoolleiders en leraren gemiddeld eerder wel vrezen over onvoldoende tijd te beschikken om grondig met de feedback aan de slag te gaan (gem = 3.6, SD = 1.1), terwijl ze het eerder oneens zijn met de stelling 'ik vrees dat ik niet in staat ben om met de feedback aan de slag te gaan' (gem = 2.0, SD = 1.0). Wat betreft andere-bekommernissen geven schoolleiders en leraren in beperkte mate aan bekommerd te zijn over de invloed die centrale toetsen op de leerlingen (gem = 3.4, SD = 1.2) en op de samenwerking tussen de leraren op de school (gem = 3.1, SD = 1.2) zullen hebben. Hoewel schoolleiders en leraren gemiddeld neutraal antwoorden op de stelling 'ik wil in het gebruik van feedback samenwerken met collega's uit andere scholen' (gem = 3.0, SD = 1.1), geven ze tegelijkertijd aan gemiddeld eerder gematigd positief te zijn om collega's leraren te willen helpen om met de feedback aan de slag te gaan (gem = 3.5, SD = 1.0) en zelf te willen weten hoe anderen met de feedback aan de slag gaan (gem = 3.8, SD = 0.9).

Wanneer we de verschillen in bekommernissen naar functie bekijken, merken we enkele statistisch significante verschillen tussen leraren en schoolleiders op ten aanzien van zelf-bekommernissen en andere-bekommernissen. Zo stellen we kleine verschillen vast tussen schoolleiders en leraren met betrekking tot zelf-bekommernissen. Leraren geven gemiddeld in grotere mate dan schoolleiders aan zich zorgen te maken over de hoeveelheid tijd die feedbackgebruik van hen zal vergen [gem L = 3.6, gem SL = 3.3 ; $t(245.1) = -3.4$, $p < 0.01$, $d = -0.31$] en zijn schoolleiders het gemiddeld helemaal eens om de mogelijkheden van feedbackgebruik verder te willen verkennen, terwijl leraren het hiermee slechts gemiddeld eerder eens zijn [gem L = 3.7, gem SL = 4.1 ; $t(687.0) = 5.0$, $p < 0.01$, $d = 0.47$]. Met betrekking tot andere-bekommernissen geven schoolleiders op schaalniveau schoolleiders – vanuit een gemiddeld gematigd positieve inschatting – aan in sterkere mate dan leraren bekommerd te zijn om 'anderen' [gem L = 3.3, gem SL = 3.7 ; $t(269.1) = 8.5$, $p < 0.01$, $d = 0.76$]. Tevens blijkt dat schoolleiders gemiddeld een grotere bereidheid tonen met betrekking tot samenwerking voor feedbackgebruik dan leraren. Schoolleiders geven immers in meerdere mate dan leraren aan te willen weten hoe anderen met de feedback aan de slag gaan [gem L = 3.7, gem SL = 4.1 ; $t(687.0) = 5.2$, $p < 0.01$, $d = 0.51$] en te willen samenwerken met collega's uit andere scholen in het gebruik van de feedback [gem L = 2.8, gem SL = 3.6 ; $t(674.0) = 8.5$, $p < 0.01$, $d = 0.76$]. Daarnaast blijken er op itemniveau kleine statistisch significante verschillen naar onderwijsniveau met betrekking tot aanwezige bekommernissen. Respondenten uit het secundair onderwijs geven gemiddeld aan bezorgd te zijn over de invloed die centrale toetsen op de leerlingen [gem LO = 3.0, gem SO = 3.5 ; $t(384.8) = -4.9$, $p < 0.01$, $d = 0.41$]

zullen hebben, terwijl respondenten uit het lager onderwijs zich hier slechts gematigd om bekommeren. Ook wat betreft bekommernissen over de invloed die centrale toetsen op de samenwerking tussen leraren in de school zullen hebben, hebben respondenten uit het secundair onderwijs gemiddeld een neutrale houding, terwijl respondenten uit het lager onderwijs hierover minder bekommerd zijn [gem LO = 2.8, gem SO = 3.4 ; $t(419.8) = -4.7$, $p < 0.01$, $d = -0.38$]. Respondenten uit het lager onderwijs zijn het dan gemiddeld weer in meerdere mate dan respondenten uit het secundair onderwijs eens om collega-leraren te willen helpen om met feedback aan de slag te gaan [gem LO = 3.7, gem SO = 3.4 ; $t(681.0) = 2.9$, $p < 0.01$, $d = 0.24$]. Tot slot blijken er ook enkele statistisch significante verschillen tussen de verschillende onderwijsnetten. Zo stellen we vast dat zelf-bekommernissen bij respondenten uit het vrij gesubsidieerd onderwijs over het algemeen – in beperkte mate – sterker aanwezig zijn dan bij respondenten uit het officieel gesubsidieerd onderwijs [gem VG = 3.6, gem OG = 3.5 ; $F(4.2, 235.9) = 5.6$, $p < 0.01$, $eta^2 = 0.02$] en dat respondenten uit het vrij gesubsidieerd onderwijs in meerdere mate dan respondenten uit het officieel gesubsidieerd onderwijs aangeven bezorgd te zijn over de invloed van centrale toetsen op de samenwerking tussen leraren in de school dan [gem VG = 3.2, gem OG = 2.7 ; $F(2.0, 113.5) = 6.2$, $p < 0.01$, $eta^2 = 0.02$].

Tabel 42
Bekommernissen ten aanzien van feedbackgebruik

	N WN/ NVT	Gem (SD)	T-test Gem SL		L	LO		SO	Anova Gem VG	G	OG	
Samenwerking												
Ik wil weten hoe anderen met de feedback uit centrale toetsen aan de slag gaan.	694 1.7	3.8 (0.9)	4.1	►	3.7	3.8		3.8	3.7	3.8	3.7	
Ik wil in het gebruik van de feedback uit centrale toetsen samenwerken met collega's uit andere scholen.	681 3.5	3.0 (1.1)	3.6	►	2.8	3.1		3.0	2.9	2.9	2.8	
Ik wil collega-leraren helpen om met feedback uit centrale toetsen aan de slag te gaan.	683 3.3	3.5 (1.0)	4.2		3.3	3.7	►	3.4	3.4	3.4	3.4	
Ik ben bezorgd over de invloed die centrale toetsen op de samenwerking tussen leraren in onze school zullen hebben.	694 1.7	3.1 (1.2)	3.1		3.1	2.8	◄	3.2	3.2	2.9	2.7	VG►OG
Gevolgen voor de leerlingen												
Ik ben bezorgd over de invloed die centrale toetsen op onze leerlingen zullen hebben.	695 1.6	3.4 (1.2)	3.4		3.4	3.0	◄	3.5	3.4	3.4	3.1	
Schaal: Andere-bekommernissen (α = 0.76)	684 0.7	3.3 (0.6)	3.7	►	3.3	3.3		3.4	3.3	3.3	3.2	
Beheersing												

Ik vrees dat ik niet in staat ben om met de feedback uit centrale toetsen aan de slag te gaan.	610 4.7	2.0 (1.0)	1.9		2.0	2.0		2.0	2.0	2.1	2.0		
Ik vrees over onvoldoende tijd te beschikken om grondig met de feedback uit centrale toetsen aan de slag te gaan.	696 1.4	3.6 (1.1)	3.4		3.6	3.5		3.6	3.6	3.4	3.4		
Schaal: Taak-bekommernissen													
Persoonlijke betrokkenheid													
Ik maak me zorgen over hoeveel tijd feedbackgebruik uit centrale toetsen van mij zal vergen.	689 2.4	3.6 (1.1)	3.3	◄	3.6	3.6		3.6	3.6	3.3	3.6		
Informatie													
Ik zou meer informatie willen hebben over hoe ik met de feedback uit centrale toetsen aan de slag kan.	694 1.7	4.0 (0.9)	4.1		4.0	4.0		4.0	4.0	3.	4.0		
Ik zou de mogelijkheden van feedbackgebruik uit centrale toetsen verder willen verkennen.	694 1.7	3.8 (0.8)	4.1	►	3.7	3.8		3.7	3.7	3.6	3.7		
Bewustwording													
Ik weet op dit moment niet wat het gebruik van feedback uit centrale toetsen kan inhouden	689 2.4	3.2 (1.2)	3.1		3.3	3.1		3.3	3.3	3.0	3.2		
Ik heb het zo druk, dat ik momenteel weinig aandacht kan besteden aan de mogelijkheden van feedbackgebruik uit centrale toetsen.	691 2.1	3.4 (1.1)	3.3		3.4	3.4		3.4	3.5	3.1	3.2		
Schaal: Zelf-bekommernissen (α = 0.81)	693 1.00	3.6 (0.6)	3.6		3.6	3.6		3.6	3.6	3.4	3.5	VG►OG	

5.4 VERWACHTE ONWENSELIJKE EFFECTEN VAN CENTRALE TOETSEN

Verwachte onwenselijke effecten van centrale toetsen werden zowel bevraagd op leerling- en leraarniveau, als op school- en onderwijssysteemniveau.

5.4.1 Verwachte onwenselijke effecten van centrale toetsen op leerlingniveau

Op leerlingniveau (zie Tabel 43) beoordelen schoolleiders en leraren verwachte onwenselijke effecten van centrale toetsen gemiddeld als neutraal (gem = 3.2, SD = 1.0). Schoolleiders en leraren beoordelen de items 'ik verwacht dat centrale toetsen een bedreiging vormen voor de brede evaluatie van leerlingen' (gem = 3.4, SD = 1.2) en 'ik verwacht dat centrale toetsen tot competitie leiden tussen leerlingen' (gem = 3.1, SD = 1.2) eveneens als neutraal.

Als we naar verschillen naar functie kijken, merken we een statistisch significant verschil tussen schoolleiders en leraren op ten aanzien van de stelling 'als de centrale toetsen vorm krijgen zoals gepland, verwacht ik dat ze tot competitie tussen leerlingen leiden'. Leraren beoordelen deze stelling gemiddeld neutraal, terwijl schoolleiders deze stelling gemiddeld als eerder oneens beoordelen [gem SL = 2.9, gem L = 3.2 ; $t(573.0) = -2.5$, $p = 0.01$, $d = -0.25$]. Wat betreft verschillen naar onderwijsniveau ten aanzien van verwachte onwenselijke effecten van centrale toetsen op leerlingniveau zijn de verschillen tussen enerzijds respondenten uit het lager onderwijs en anderzijds respondenten uit het secundair onderwijs gemiddeld klein [gem LO = 2.9, gem SO = 3.4 ; $t(561.0) = -5.2$, $p < 0.01$, $d = -0.48$]. Op itemniveau geven leraren en schoolleiders uit het secundair onderwijs aan in meerdere mate dan leraren en schoolleiders uit het lager onderwijs aan te verwachten dat centrale toetsen een bedreiging vormen voor de brede evaluatie van leerlingen [gem LO = 3.1, gem SO = 3.5 ; $t(568.0) = -3.7$, $p < 0.01$, $d = -0.34$] en tot competitie tussen leerlingen leiden [gem LO = 2.7, gem SO = 3.2 ; $t(576.0) = -4.7$, $p < 0.01$, $d = -0.43$]. Tot slot stellen we vast dat de impact van onderwijsnet op de verwachtingen naar onwenselijke effecten van centrale toetsen op leerlingniveau onbestaande of verwaarloosbaar klein is ($eta^2 \leq 0.01$).

Tabel 43
Verwachte onwenselijke effecten van centrale toetsen op leerlingniveau

	N WN/ NVT	Gem (SD)	T-test Gem SL	L	LO	SO	Anova Gem VG	G	OG
Als de centrale toetsen vorm krijgen zoals gepland, verwacht ik dat ze op leerlingniveau:									
een bedreiging vormen voor de brede evaluatie van leerlingen.	570 5.8%	3.4 (1.2)	3.7	3.3	3.1 ◄	3.5	3.4	3.2	2.9
tot competitie leiden tussen leerlingen .	578 4.5%	3.1 (1.2)	2.9 ◄	3.2	2.7 ◄	3.2	3.1	3.4	2.9
Schaal: Verwachte onwenselijke effecten op leerlingenniveau (α = 0.77)	563 3.4%	3.2 (1.0)	3.1	3.3	2.9 ◄	3.4	3.2	3.3	2.9

5.4.2 Verwachte onwenselijke effecten van centrale toetsen op leraarniveau

Verder geven schoolleiders en leraren aan het gemiddeld eerder eens te zijn onwenselijke effecten te verwachten op leraarniveau (gem = 3.5, SD = 0.9) (zie Tabel 44). Zo zijn ze het er gemiddeld eerder mee eens dat centrale toetsen leraren aanzetten om leerlingen extra voor te bereiden op de centrale toetsen (gem = 3.8, SD = 1.1), voor extra planlast zorgen (gem = 3.8, SD = 1.1), leiden tot minder aandacht voor leerinhouden die niet getoetst worden in centrale toetsen (gem = 3.7, SD = 1.1) en voor minder ruimte zorgen voor een persoonlijke inbreng van leraren in de klaspraktijk (gem = 3.5, SD = 1.2). Op de stellingen dat centrale toetsen

demotiverend werken voor leraren (gem = 3.3, SD = 1.1), tot competitie leiden tussen leraren (gem = 3.1, SD = 1.2), een bedreiging vormen voor de pedagogische vrijheid van leraren (gem = 3.4, SD = 1.2) en zorgen voor onnodige twijfel over de evaluatiecompetenties van leraren (gem = 3.4, SD = 1.2) antwoorden respondenten dan weer gemiddeld neutraal.

Als we naar verschillen naar functie kijken, merken we enkele statistisch significante verschillen tussen schoolleiders en leraren op. Zo verwachten leraren gemiddeld in meerder mate dan schoolleiders dat centrale toetsen voor onnodige twijfel over de evaluatiecompetenties van leraren [gem SL = 3.2, gem L = 3.5 ; $t(567.0) = -2.3$, $p < 0.01$, $d = -0.23$] en voor extra planlast zorgen [gem SL = 3.6, gem L = 3.9 ; $t(582.0) = -2.4$, $p = 0.02$, $d = -0.24$]. Naar onderwijsniveau stellen we zowel voor elk item als voor de gehele schaal [gem LO = 3.1, gem SO = 3.7 ; $t(573.0) = -7.9$, $p < 0.01$, $d = -0.72$] statistisch significante verschillen vast tussen leraren en schoolleiders uit het lager onderwijs enerzijds en leraren en schoolleiders uit het secundair onderwijs anderzijds. Op itemniveau betreffen deze verschillen de verwachtingen dat centrale toetsen voor onnodige twijfel zorgen over de evaluatiecompetenties van leraren [gem LO = 3.0, gem SO = 3.6 ; $t(570.0) = -5.0$, $p < 0.01$, $d = -0.45$], voor extra planlast zorgen [gem LO = 3.5, gem SO = 4.0 ; $t(585.0) = -4.4$, $p < 0.01$, $d = -0.39$], demotiverend werken voor leraren [gem LO = 3.0, gem SO = 3.4 ; $t(338.8) = -3.2$, $p < 0.01$, $d = -0.29$] en tot competitie tussen leraren leiden [gem LO = 2.8, gem SO = 3.3 ; $t(577.0) = -4.3$, $p < 0.01$, $d = -0.39$]. Daarenboven verschillen respondenten uit het secundair onderwijs in nog meerdere dan respondenten uit het lager onderwijs met betrekking tot de verwachtingen dat centrale toetsen leiden tot minder aandacht voor leerinhouden die niet getoetst worden [gem LO = 3.3, gem SO = 3.9 ; $t(563.0) = -6.3$, $p < 0.01$, $d = -0.58$], voor minder ruimte zorgen voor een persoonlijke inbreng van leraren in de klaspraktijk [gem LO = 2.9, gem SO = 3.8 ; $t(576.0) = -8.2$, $p < 0.01$, $d = -0.74$], een bedreiging vormen voor de pedagogische vrijheid van leraren [gem LO = 2.9, gem SO = 3.5 ; $t(579.0) = -6.0$, $p < 0.01$, $d = -0.55$] en leraren aanzetten om leerlingen extra voor te bereiden op de centrale toetsen [gem LO = 3.3, gem SO = 4.1 ; $t(290.8) = -8.0$, $p < 0.01$, $d = -0.74$]. Tot slot stellen we vast dat de impact van onderwijsnet op de verwachtingen naar onwenselijke effecten van centrale toetsen ook op leraarniveau onbestaande of verwaarloosbaar klein is ($eta^2 \leq 0.01$).

Tabel 44
Verwachte onwenselijke effecten van centrale toetsen op leraarniveau

	N WN/ NVT	Gem (SD)	T-test Gem				Anova Gem		
			SL	L	LO	SO	VG	G	OG
Als de centrale toetsen vorm krijgen zoals gepland, verwacht ik dat ze op leraarniveau:									

leiden tot minder aandacht voor leerinhouden die niet getoetst worden in centrale toetsen.	565 6.6%	3.7 (1.1)	3.7		3.7	3.3	◄	3.9	3.7	3.7	3.9
voor minder ruimte zorgen voor een persoonlijke inbreng van leraren in de klaspraktijk.	578 4.5%	3.5 (1.2)	3.6		3.6	2.9	◄	3.8	3.6	3.3	3.4
een bedreiging vormen voor de pedagogische vrijheid van leraren.	581 4.0%	3.4 (1.2)	3.2		3.4	2.9	◄	3.5	3.4	3.3	3.2
zorgen voor onnodige twijfel over de evaluatiecompetenties van leraren.	572 5.5%	3.4 (1.2)	3.2	◄	3.5	3.0	◄	3.6	3.5	3.3	3.3
leraren aanzetten om leerlingen extra voor te bereiden op de centrale toetsen.	585 3.3%	3.8 (1.1)	3.0		3.6	3.3	◄	4.1	3.8	3.9	4.0
voor extra planlast zorgen.	587 3.0%	3.8 (1.1)	3.6	◄	3.9	3.5	◄	4.0	3.9	3.7	3.8
demotiverend werken voor leraren.	574 5.1%	3.3 (1.1)	3.3		3.3	3.0	◄	3.4	3.3	3.1	2.9
tot competitie leiden tussen leraren.	579 4.3%	3.1 (1.2)	3.0		3.2	2.8	◄	3.3	3.2	3.2	3.0
Schaal: Verwachte onwenselijke effecten op leraarniveau (α = 0.90)	575 1.7%	3.5 (0.9)	3.4		3.6	3.1	◄	3.7	3.6	3.4	3.4

5.4.3 Verwachte onwenselijke effecten van centrale toetsen op schoolniveau

Ook op schoolniveau (zie Tabel 45) geven schoolleiders en leraren gemiddeld aan eerder wel onwenselijke effecten te verwachten van centrale toetsen (gem = 3.5, SD = 1.2). Respondenten geven aan het eerder eens te zijn dat centrale toetsen tot openbare *rankings* van scholen leiden (gem = 3.9, SD = 1.1), tot competitie tussen scholen (gem = 3.9, SD = 1.1) leiden, voor te grote druk zorgen op het schoolteam om goede prestaties te realiseren op centrale toetsen (gem = 3.7, SD = 1.1) en een bedreiging vormen voor de autonomie van scholen (gem = 3.4, SD = 1.2). In mindere mate verwachten ze dat centrale toetsen scholen aanzetten tot het manipuleren van resultaten (gem = 3.0, SD = 1.3) en op een ondoordachte manier gebruikt worden door de schoolleiding om de onderwijsprestaties van leraren te evalueren (gem = 3.0, SD = 1.2).

Als we naar verschillen naar functie kijken, merken we op schaalniveau een statistisch significant verschil tussen schoolleiders en leraren op. Leraren verwachten gemiddeld in meerdere mate onwenselijke effecten van centrale toetsen op leraarniveau [gem SL = 3.3, gem L = 3.6 ; $t(550.0) = -2.3$, $p < 0.01$, $d = -0.25$]. Behalve dit kleine verschil verwachten leraren tevens in meerdere mate dan schoolleiders dat centrale toetsen op een ondoordachte manier gebruikt worden door de schoolleiding om de onderwijsprestaties van leraren te evalueren [gem SL = 2.5, gem L = 3.2 ; $t(553.0) = -6.0$, $p < 0.01$, $d = -0.61$]. Ook op schoolniveau [gem LO = 3.2,

gem SO = 3.7 ; $t(201.4) = -3.0$, $p < 0.01$, $d = -0.33$] stellen we statistisch significante verschillen tussen respondenten uit het lager en secundair onderwijs vast. Zo verschillen beide groepen gemiddeld in kleine mate wat de verwachtingen van centrale toetsen betreffen tot het aanzetten van scholen tot het manipuleren van resultaten [gem LO = 2.7, gem SO = 3.1 ; $t(552.0) = -2.7$, $p < 0.01$, $d = -0.25$], tot het op een ondoordachte manier gebruikmaken van centrale toetsen door de schoolleiding om de onderwijsprestaties van leraren te evalueren [gem LO = 2.8, gem SO = 3.2 ; $t(556.0) = -3.1$, $p < 0.01$, $d = -0.29$], tot het leiden tot openbare *rankings* van scholen [gem LO = 3.6, gem SO = 4.1 ; $t(243.2) = -4.7$, $p < 0.01$, $d = -0.46$] en tot het vormen van een bedreiging voor de autonomie van scholen [gem LO = 3.0, gem SO = 3.6 ; $t(569.0) = -5.2$, $p < 0.01$, $d = -0.48$]. Daarnaast geven respondenten uit het secundair onderwijs aan gemiddeld nog sterker te verschillen met respondenten uit het lager onderwijs in hun verwachtingen dat centrale toetsen voor te grote druk zorgen op het schoolteam om goede prestaties te realiseren op de centrale toetsen [gem LO = 3.2, gem SO = 3.8 ; $t(285.0) = -5.8$, $p < 0.01$, $d = -0.54$] en tot competitie leiden tussen scholen [gem LO = 3.5, gem SO = 4.1 ; $t(241.6) = -5.4$, $p < 0.01$, $d = -0.53$]. Tevens blijken respondenten naar onderwijsnet niet te verschillen in hun verwachtingen ten aanzien van onwenselijke effecten van centrale toetsen op schoolniveau.

Tabel 45
Verwachte onwenselijke effecten van centrale toetsen op schoolniveau

	N WN/ NVT	Gem (SD)	T-test Gem SL	L	LO	SO	Anova Gem VG	G	OG
Als de centrale toetsen vorm krijgen zoals gepland, verwacht ik dat ze op schoolniveau:									
scholen aanzetten tot het manipuleren van resultaten.	554 8.4%	3.0 (1.3)	2.8	3.0	2.7 ◄	3.1	3.0	3.0	2.8
voor te grote druk zorgen op het schoolteam om goede prestaties te realiseren op de centrale toetsen.	580 4.1%	3.7 (1.1)	3.5	3.7	3.2 ◄	3.8	3.7	3.6	3.6
op een ondoordachte manier gebruikt worden door de schoolleiding om de onderwijsprestaties van leraren te evalueren.	558 7.8%	3.0 (1.2)	2.5 ◄	3.2	2.8 ◄	3.2	3.1	3.2	3.1
tot openbare rankings van scholen leiden.	573 5.3%	3.9 (1.1)	3.9	3.9	3.6 ◄	4.1	4.0	3.9	3.9
tot competitie leiden tussen scholen.	576 4.8%	3.9 (1.1)	3.9	3.9	3.5 ◄	4.1	4.0	3.8	3.8
een bedreiging vormen voor de autonomie van scholen.	571 5.6%	3.4 (1.2)	3.3	3.5	3.0 ◄	3.6	3.5	3.4	3.3

Schaal: Verwachte onwenselijke effecten op schoolniveau (α = 0.84)	555 2.5%	3.5 (1.2)	3.3	◄	3.6	3.2	◄	3.7	3.6	3.5	3.4				

5.4.4 Verwachte onwenselijke effecten van centrale toetsen op onderwijssysteemniveau

In de meeste mate verwachten schoolleiders en leraren onwenselijke effecten op onderwijssysteemniveau (gem = 3.6, SD = 1.0) (zie Tabel 46). Zo geven ze aan het gemiddeld eerder eens te zijn dat centrale toetsen tot oneerlijke vergelijkingen tussen scholen leiden (gem = 3.7, SD = 1.2), tot grotere segregatie tussen scholen leiden (gem = 3.7, SD = 1.1), de focus van de onderwijsinspectie te sterk richten op output-data (gem = 3.8, SD = 1.1), de invulling van onderwijskwaliteit vernauwen tot cognitieve leerresultaten (gem = 3.5, SD = 1.2) en leerprestaties van leerlingen verengen tot de toetsresultaten van de leerlingen op centrale toetsen voor wiskunde en Nederlands (gem = 3.6, SD = 1.2). Respondenten beantwoorden de stelling dat centrale toetsen een bedreiging vormen voor de pedagogische vrijheid van scholen gemiddeld als neutraal (gem = 3.4, SD = 1.2).

Naar functie blijken de respondenten niet te verschillen in hun verwachtingen ten aanzien van onwenselijke effecten van centrale toetsen op onderwijssysteemniveau. Wel stellen we statistisch significante verschillen naar onderwijsniveau vast [gem LO = 3.2, gem SO = 3.8 ; $t(254.5) = -5.9$, $p < 0.01$, $d = -0.58$]. Zo geven respondenten uit het secundair onderwijs gemiddeld in meerdere mate dan respondenten uit het lager onderwijs aan te verwachten dat centrale toetsen de invulling van onderwijskwaliteit vernauwen tot cognitieve leerresultaten [gem LO = 3.1, gem SO = 3.7 ; $t(292.5) = -5.1$, $p < 0.01$, $d = -0.47$], leerprestaties van leerlingen verengen tot de toetsresultaten van de leerlingen op de centrale toetsen voor wiskunde en Nederlands [gem LO = 3.2, gem SO = 3.7 ; $t(289.0) = -5.2$, $p < 0.01$, $d = -0.49$], een bedreiging vormen voor de pedagogische vrijheid van scholen [gem LO = 3.0, gem SO = 3.6 ; $t(571.0) = -5.0$, $p < 0.01$, $d = -0.46$], tot grotere segregatie leiden tussen scholen [gem LO = 3.2, gem SO = 3.8 ; $t(254.7) = -5.1$, $p < 0.01$, $d = -0.49$], de focus van de onderwijsinspectie te sterk richten op output-data [gem LO = 3.5, gem SO = 4.0 ; $t(263.2) = -4.6$, $p < 0.01$, $d = -0.44$] en leiden tot oneerlijke vergelijking [gem LO = 3.3, gem SO = 3.9 ; $t(269.7) = -5.6$, $p < 0.01$, $d = -0.53$]. Tot slot stellen we vast dat de impact van onderwijsnet op de verwachtingen naar onwenselijke effecten van centrale toetsen ook op onderwijssysteemniveau onbestaande of verwaarloosbaar klein is ($eta^2 \leq 0.01$).

Tabel 46
Verwachte onwenselijke effecten van centrale toetsen op onderwijssysteemniveau

	N WN/ NVT	Gem (SD)	T-test Gem SL	L	LO	SO	Anova Gem VG	G	OG

Als de centrale toetsen vorm krijgen zoals gepland, verwacht ik dat ze op onderwijssysteemniveau:												
de invulling van onderwijskwaliteit vernauwen tot cognitieve leerresultaten.	573 5.3%	3.5 (1.2)	3.4	3.6	3.1	◄	3.7	3.6	3.3	3.4		
leerprestaties van leerlingen verengen tot de toetsresultaten van de leerlingen op de centrale toetsen voor wiskunde en Nederlands.	581 4.0%	3.6 (1.2)	3.4	3.6	3.2	◄	3.7	3.6	3.4	3.3		
leiden tot oneerlijke vergelijkingen tussen scholen.	576 4.8%	3.7 (1.2)	3.7	3.8	3.3	◄	3.9	3.8	3.6	3.7		
een bedreiging vormen voor de pedagogische vrijheid van scholen.	573 5.3%	3.4 (1.2)	3.3	3.5	3.0	◄	3.6	3.5	3.3	3.3		
tot grotere segregatie tussen scholen leiden.	553 8.6%	3.7 (1.1)	3.6	3.6	3.2	◄	3.8	3.7	3.6	3.6		
de focus van de onderwijsinspectie te sterk richten op output-data.	549 9.3%	3.8 (1.1)	3.8	3.8	3.5	◄	4.0	3.9	3.6	3.6		
Schaal: Verwachte onwenselijke effecten op onderwijssysteemniveau (α = 0.88)	553 2.6%	3.6 (1.0)	3.5	3.7	3.2	◄	3.8	3.7	3.5	3.5		

5.5 SPECIFIEKE ACTUELE THEMA'S

In de vragenlijst werden aan respondenten eveneens specifieke actuele thema's met betrekking tot het systeem van centrale toetsen voorgelegd (zie Tabel 47). Sommige van de thema's richtten zich enkel tot respondenten uit een specifiek onderwijsniveau en werden bijgevolg enkel beantwoord door respondenten uit hetzij het lager onderwijs, hetzij het secundair onderwijs.

Schoolleiders en leraren zijn van mening dat de feedback gemiddeld eerder geen betrouwbare informatie geeft over de kwaliteit van de school (gem = 2.6, SD = 1.0) en hebben een gemiddeld neutrale houding ten aanzien van de stelling dat centrale toetsen bijdragen tot een kwaliteitsvolle schoolwerking (gem = 3.2, SD = 1.0). Daarnaast maken schoolleiders en leraren een gemiddeld gematigd negatieve inschatting ten aanzien van de stellingen dat de feedback op klasniveau betrouwbare informatie geeft over de kwaliteit van de instructie van de leraar (gem = 2.4, SD = 1.0) en dat de toetsresultaten van de leerlingen de school goede handvaten geeft voor de evaluatie van de individuele leraren (gem = 2.3, SD = 1.0). Verder blijken schoolleiders en leraren de stellingen 'centrale toetsen voegen weinig toe voor wie IDP-of OVSG-toetsen gebruikt' (gem 3.4, SD = 0.9) en 'centrale toetsen vullen de huidige evaluatie van leerlingen op onze school goed aan' (gem 3.2, SD = 1.0) gemiddeld als neutraal te beoordelen. Eveneens maken schoolleiders en leraren een gemiddeld neutrale inschatting van de stellingen dat centrale toetsen een effectievere begeleiding van

scholen door hun pedagogische begeleidingsdienst mogelijk maken (gem = 3.0, SD = 1.0) en voelen ze zich gemiddeld eerder niet door de pedagogische begeleidingsdienst ondersteund bij het interpreteren en analyseren van de feedback (gem = 2.9, SD = 1.2). Vervolgens zijn schoolleiders en leraren het eerder oneens dat centrale toetsen scholen de mogelijkheid geeft zich beter te profileren (gem = 2.9, SD = 1.1) en meer betrouwbare informatie leveren voor ouders om de schoolkeuze op te baseren (gem = 2.5, SD = 1.1). Schoolleiders en leraren vinden het gemiddeld eveneens eerder niet aanvaardbaar dat centrale toetsen slechts een selectie van alle eindtermen van een leergebied meten (gem = 2.9, SD = 1.1) en zijn tegelijkertijd neutraal in hun inschatting over de stellingen dat centrale toetsen uitspraken over alle termen van de getoetste leergebieden mogelijk moeten maken (gem = 3.4, SD = 1.0) en dat centrale toetsen zich beter richten op de leerplannen dan op de eindtermen (gem = 3.1, SD = 1.0). Bovendien blijkt dat schoolleiders en leraren gemiddeld een neutrale houding hebben ten aanzien van de stelling 'alle leerlingen zouden dezelfde centrale toetsen moeten afleggen' (gem = 3.4, SD = 1.4), terwijl ze gemiddeld een matig negatieve inschatting maken van de stelling 'om de klas- en schoolpraktijk beter te informeren is het zinvol verschillende toetsen aan te bieden die slechts door een deel van de leerlingen ingevuld worden' (gem = 2.5, SD = 1.0). Schoolleiders en leraren beoordelen de stelling 'de inhoud en moeilijkheidsgraad van de centrale toets zou zich moeten aanpassen aan het (vaardigheids)niveau van de leerling die de toets aflegt (gem = 3.1, SD = 1.2) eveneens gemiddeld als eerder negatief en de stelling 'deelnemen aan centrale toetsen zou voor de school op vrijwillige basis moeten gebeuren' (gem = 3.3, SD = 1.2) gemiddeld als neutraal. Daarenboven zijn schoolleiders en leraren uit respectievelijk het lager onderwijs en het secundair onderwijs gemiddeld neutraal in hun inschatting dat er meer dan de huidige meetmomenten in het basisonderwijs (gem = 3.4, SD = 1.1) en in het secundair onderwijs (gem = 3.3., SD = 1.2) nodig zijn om centrale toetsen zinvol te maken. Tot slot is het opvallend dat schoolleiders en leraren gemiddeld gematigd positief zijn over de stellingen 'de overheid zou ook bij aanvang van het lager onderwijs een centrale toets voor Nederlands en wiskunde moeten organiseren' (gem = 3.5, SD = 1.3), 'centrale toetsen mogen zich niet beperken tot Nederlands en wiskunde' (gem = 3.7, SD = 1.0) en 'een centrale toets voor een praktijkvak is even relevant voor wiskunde dan voor Nederland' (gem = 3.6, SD = 1.2).

Met betrekking tot verschillen in de inschatting van deze specifieke actuele stellingen naar functie stellen we kleine statistisch significante verschillen vast. Zo zijn schoolleiders het gemiddeld eerder wel eens, terwijl leraren een neutrale houding hebben ten aanzien van de stelling dat centrale toetsen bijdragen aan een kwaliteitsvolle schoolwerking [gem L = 3.1, gem SL = 3.5 ; $t(544.0) = 3.6$, $p < 0.01$, $d = 0.36$]. Hoewel schoolleiders en leraren de stelling dat de feedback op schoolniveau betrouwbare info over de kwaliteit van de school geeft als gemiddeld neutraal beoordelen, maken schoolleiders toch in kleine mate een positievere

inschatting [gem SL = 3.0 gem L = 2.5 ; $t(554.0) = 4.4$, $p < 0.01$, $d = 0.43$]. Daarnaast maken schoolleiders gemiddeld een minder negatieve inschatting dan leraren van de stellingen 'de feedback op klasniveau geeft betrouwbare informatie over de kwaliteit van de instructie van de leraar' [gem SL = 2.7 gem L = 2.3 ; $t(559.0) = 4.7$, $p < 0.01$, $d = 0.46$] en 'om de klas- en schoolpraktijk beter te informeren is het zinvol verschillende toetsen aan te bieden die slechts door een deel van de leerlingen worden ingevuld' [gem SL = 2.7 gem L = 2.4 ; $t(524.0) = 2.5$, $p < 0.01$, $d = 0.25$]. Verder beoordelen schoolleiders de stellingen 'ik voel me door de pedagogische begeleidingsdienst ondersteund bij het interpreteren en analyseren van de feedback' [gem SL = 3.1 gem L = 2.8 ; $t(574.0) = 2.2$, $p = 0.03$, $d = 0.22$] en 'ik vind het aanvaardbaar dat centrale toetsen slechts een selectie van alle eindtermen van een leergebied meten' [gem SL = 3.3 gem L = 2.8; $t(484.0) = 4.2$, $p < 0.01$, $d = 0.43$] als neutraal, terwijl leraren het er gemiddeld eerder oneens mee zijn. Tevens zijn schoolleiders het er gemiddeld in meerdere mate dan leraren mee eens dat de inhoud en moeilijkheidsgraad van de centrale toets zich moeten aanpassen aan het (vaardigheids)niveau van de leerling die de toets aflegt [gem SL = 3.3 gem L = 3.0 ; $t(560.0) = 2.1$, $p = 0.03$, $d = 0.21$] en zijn ze het gemiddeld meer eens dan leraren dat centrale toetsen zich niet mogen beperken tot Nederlands en wiskunde [gem SL = 3.8 ,gem L = 3.6 ; $t(565) = 2.1$, $p = 0.04$, $d = 0.21$]. Leraren op hun beurt maken gemiddeld een minder negatieve inschatting dan schoolleiders van de stelling dat centrale toetsen meer betrouwbare informatie opleveren voor ouders om de schoolkeuze op te baseren [gem SL = 2.3 gem L = 2.5 ; $t(570.0) = 2.1$, $p = 0.03$, $d = 0.21$]. Verder zijn leraren gemiddeld neutraal en schoolleiders gemiddeld gematigd negatief ten aanzien van de stellingen dat centrale toetsen scholen de mogelijkheid geven zich beter te profileren [gem SL = 2.7 gem L = 3.0 ; $t(555.0) = -2.0$, $p = 0.04$, $d = -0.20$] en 'centrale toetsen richten zich beter op leerplandoelen dan op eindtermen' [gem SL = 2.9, gem L = 3.2 ; $t(196.2) = -2.0$, $p = 0.05$, $d = -0.21$]. Daarnaast zijn leraren gemiddeld gematigd positief en schoolleiders gemiddeld neutraal ten aanzien van de stellingen 'alle leerlingen zouden dezelfde toetsen moeten afleggen' [gem SL = 3.1 gem L = 3.5 ; $t(209.6) = -2.9$, $p < 0.01$, $d = -0.29$]. Bovendien zijn leraren uit het lager onderwijs het gemiddeld in beperkte mate meer eens dan schoolleiders uit het lager onderwijs dat de overheid ook bij aanvang van het lager onderwijs een centrale toets voor Nederlands en wiskunde zou moeten organiseren' [gem SL = 3.1, gem L = 3.6 ; $t(87.7.0) = -2.2$, $p = 0.03$, $d = 0.38$].

Ook naar onderwijsniveau merken we enkele kleine statistisch significante verschillen vast. Zo schatten respondenten uit het lager onderwijs de stelling dat de feedback uit centrale toetsen op schoolniveau betrouwbare info geeft over de kwaliteit van de school gemiddeld minder negatief in dan respondenten uit het secundair onderwijs [gem LO = 2.8, gem SO = 2.5 ; $t(558.0) = 3.3$, $p < 0.01$, $d = 0.31$]. Daarnaast zijn respondenten uit het lager onderwijs gemiddeld gematigd positief en respondenten uit het secundair onderwijs gemiddeld neutraal ten aanzien van de stellingen 'centrale toetsen dragen bij tot een

kwaliteitsvolle schoolwerking' [gem LO = 3.5, gem SO = 3.1 ; $t(390.2) = 4.9$, $p < 0.01$, $d = 0.44$] en 'alle leerlingen zouden dezelfde centrale toetsen moeten afleggen' [gem LO = 3.6, gem SO = 3.3 ; $t(387.6) = 2.4$, $p < 0.02$, $d = 0.21$]. Hiertegenover staat dat respondenten uit het secundair onderwijs gemiddeld positiever zijn dan respondenten uit het lager onderwijs met betrekking tot de stelling dat centrale toetsen zich beter richten op leerplandoelen dan op eindtermen [gem LO = 2.9, gem SO = 3.2 ; $t(534.0) = -2.5$, $p = 0.01$, $d = -0.24$].

Naar onderwijsnet stellen we geen statistisch significante verschillen met minstens een klein effect naar de beoordeling van specifieke actuele stellingen met betrekking tot het systeem van centrale toetsen vast.

Tabel 47
Specifieke actuele stellingen met betrekking tot het systeem van centrale toetsen

	N WN/ NVT	Gem (SD)	T-test Gem SL		L	LO		SO	Anova Gem VG	G	OG
Centrale toetsen dragen bij tot een kwaliteitsvolle schoolwerking.	550 8.5%	3.2 (1.0)	3.5	►	3.1	3.5	►	3.1	3.2	3.0	3.4
De feedback uit centrale toetsen op schoolniveau geeft betrouwbare informatie over de kwaliteit van de school.	560 6.8%	2.6 (1.0)	3.0	►	2.5	2.8	►	2.5	2.5	2.7	2.7
De feedback uit centrale toetsen op klasniveau geeft betrouwbare informatie over de kwaliteit van de instructie van de leraar.	565 6.0%	2.4 (1.0)	2.7	►	2.3	2.5		2.3	2.3	2.3	2.5
De toetsresultaten van leerlingen op centrale toetsen geven mijn school goede handvaten voor de evaluatie van individuele leraren.	562 6.5%	2.3 (1.0)	2.4		2.2	2.2		2.3	2.3	2.2	2.1
Centrale toetsen voegen weinig toe voor wie IDP-proeven of OVSG-toetsen gebruikt.	139 21.0%	3.4 (0.9)	3.5		3.4				3.4	3.3	3.2
Centrale toetsen vullen de huidige evaluatie van leerlingen op onze school goed aan.	517 14.0%	3.2 (1.0)	3.3		3.2	3.1		3.3	3.2	3.3	3.4
Centrale toetsen maken een effectievere begeleiding van scholen door hun pedagogische begeleidingsdienst mogelijk	544 9.5%	3.0 (1.0)	3.0		3.0	3.1		3.0	3.0	3.0	3.0
Ik voel me door de pedagogische begeleidingsdienst ondersteund bij het interpreteren en analyseren van de feedback uit centrale toetsen.	490 21.6%	2.9 (1.2)	3.3	►	2.8	3.1		2.9	3.0	2.6	2.6
Centrale toetsen geven scholen de mogelijkheid zich beter te profileren.	561 6.7%	2.9 (1.1)	2.7	◄	3.0	2.9		2.9	2.9	2.9	2.8
Centrale toetsen leveren meer betrouwbare informatie op voor ouders om de schoolkeuze op te baseren.	576 4.2%	2.5 (1.1)	2.3	◄	2.5	2.5		2.5	2.5	2.6	2.6

Centrale toetsen moeten uitspraken over alle eindtermen van de getoetste leergebieden mogelijk maken.	566 6.1%	3.4 (1.0)	3.4		3.4	3.6		3.4	3.4	3.4	3.7
Ik vind het aanvaardbaar dat centrale toetsen slechts een selectie van alle eindtermen van een leergebied meten.	580 3.8%	2.9 (1.1)	3.1	►	2.8	3.0		2.9	2.9	3.0	2.9
Centrale toetsen richten zich beter op leerplandoelen dan op eindtermen.	536 11.1%	3.1 (1.0)	2.9	◄	3.2	2.9	◄	3.2	3.2	3.0	3.0
Om de klas- en schoolpraktijk beter te informeren is het zinvol verschillende toetsen aan te bieden die slechts door een deel van de leerlingen ingevuld worden.	530 11.8%	2.5 (1.0)	2.7	►	2.4	2.6		2.4	2.5	2.3	2.3
Alle leerlingen zouden dezelfde centrale toetsen moeten afleggen.	581 3.3%	3.4 (1.3)	3.1	◄	3.5	3.6	►	3.3	3.4	3.5	3.6
De inhoud en moeilijkheidsgraad van de centrale toets zou zich moeten aanpassen aan het (vaardigheids)niveau van de leerling die de toets aflegt.	566 5.8%	3.1 (1.2)	3.3	►	3.0	3.0		3.1	3.1	2.9	2.9
Deelnemen aan centrale toetsen zou voor de school op vrijwillige basis moeten gebeuren.	585 3.0%	3.3 (1.2)	3.3		3.3	3.2		3.3	3.3	3.4	3.0
Er zijn meer dan de huidige meetmomenten (in het vierde en zesde leerjaar) nodig om centrale toetsen in het basisonderwijs zinvol te maken.	168 4.5%	3.4 (1.1)	3.5		3.4				3.4	3.6	3.1
Er zijn meer dan de huidige meetmomenten (aan het einde van de eerste en einde derde graad) nodig om centrale toetsen in het secundair onderwijs zinvol te maken.	384 10.1%	3.3 (1.2)	3.4		3.3				3.3	3.2	3.6
De overheid zou ook bij aanvang van het lager onderwijs een centrale toets voor Nederlands en wiskunde moeten organiseren.	173 1.7%	3.5 (1.3)	3.1	◄	3.6				3.5	3.6	3.6
Centrale toetsen mogen zich niet beperken tot Nederlands en wiskunde.	571 5.3%	3.7 (1.0)	3.8	►	3.6	3.6		3.7	3.7	3.6	3.5
Een centrale toets voor een praktijkvak is even relevant dan voor wiskunde en Nederlands.	365 14.5	3.6 (1.2)	3.7		3.6				3.6	3.7	3.7

We stelden respondenten tevens de open vraag 'Voor welke ander(e) vak(ken)/ leergebied(en)/ sleutelcompetentie(s) zouden centrale toetsen volgens u nog nuttig zijn?' (zie Tabel 48).

Uit de analyse van de antwoorden op deze vraag blijkt dat 27.2% van de leraren uit het lager onderwijs en 47.8% van de schoolleiders uit het lager onderwijs ook het centraal toetsen van wereldoriëntatie als nuttig beschouwen. Daarnaast geven 21.6% van de leraren en 35.8% van de schoolleiders uit het lager

onderwijs tevens aan het nuttig te vinden Frans centraal te toetsen. 13.4 % van de schoolleiders uit het lager onderwijs zien daarenboven ook het nut van het centraal toetsen van leerinhouden met betrekking tot wetenschap en techniek.

Ook respondenten uit het secundair onderwijs vinden het centraal toetsen van andere vakken (dan Nederlands en wiskunde) nuttig. Zo menen 22.6% van de leraren en 19.2% van de schoolleiders uit het secundair onderwijs dat het centraal toetsen van moderne vreemde talen (Frans, Engels en/of Duits) nuttig zou kunnen zijn. Ook het centraal toetsen van exacte wetenschappen (o.a. chemie, biologie, techniek, natuurwetenschappen, fysica) en humane wetenschappen (o.a. economie, geschiedenis, gedragswetenschappen, cultuurwetenschappen) wordt gesuggereerd. Zo verwachten 14.1% en 7.6% van de leraren en 19.2% en 9.1% van de schoolleiders uit het secundair dat het centraal toetsen van respectievelijk exacte wetenschappen en humane wetenschappen als nuttig. Daarnaast zien ook 8.9% van de leraren en 19.2% van de schoolleiders uit het secundair onderwijs het nut van het centraal toetsen van specifieke richtingsvakken.

Tabel 48
Andere vakken (dan Nederlands en wiskunde) waarvoor centrale toetsen nuttig zouden zijn

	L		SL	
	n	%	*n*	%
Lager Onderwijs				
Wereldoriëntatie	44	27.2%	32	47.8%
Frans	35	21.6%	22	35.8%
Wetenschap & Techniek	0	0%	9	13.4%
ICT	2	1.2%	4	6.0%
Andere	17	5.6%	21	32.8%
Secundair Onderwijs				
Alle vakken	15	3.3%	3	3.0%
Alle vakken met eindtermen	23	5.0%	5	5.1%
Richtingspecifieke vakken	41	8.9%	19	19.2%
Moderne vreemde talen	104	22.6%	19	19.2%
Exacte wetenschappen	65	14.1%	19	19.2%
Humane wetenschappen	35	7..6%	9	9.1%

Andere	9	2.0%	3	3.0%

6 CONCLUSIE EN AANDACHTSPUNTEN VOOR BELEID EN STEUNPUNT

Op basis van de analyse van de resultaten van dit surveyonderzoek naar verwachtingen en randvoorwaarden voor feedbackgebruik uit centrale toetsen bij Vlaamse schoolleiders en leraren formuleren we aandachtspunten. Deze aandachtspunten informeren alle betrokken samenwerkingspartners met betrekking tot een effectieve implementatie van (het gebruik van) centrale toetsen.

Dit surveyonderzoek werd gevoerd in een periode (i.c. december 2021 – februari 2022) waarin informatie (van o.a. Steunpunt) over het opzet en het doel van het Vlaams systeem van centrale toetsen eerder onbestaande was. Respondenten baseerden zich voor het invullen van de survey met andere woorden op (te) beperkte informatie van o.a. collega's of uit de media (Molenberghs, Van Gasse, De Maeyer, & Vanhoof, 2022). Om de resultaten en conclusies van dit surveyonderzoek recht aan te doen, is het van belang ze binnen deze context te interpreteren. Daarenboven is het omwille van de heterogeniteit in de antwoorden van respondent niet mogelijk deze resultaten te veralgemenen. We duiden dan ook met klem op het belang om deze resultaten en conclusie met de nodige voorzichtigheid en correcte nuancering te behandelen.

6.1 TOETSRESULTATEN OP LEERLING-, KLAS- EN SCHOOLNIVEAU MAKEN DE FEEDBACK RELEVANT

Dit gebruikersonderzoek leert dat de bevraagde schoolleiders en leraren een behoeften ervaren aan feedback die ze meteen kunnen relateren aan hun dagelijkse praktijk. Zo adviseren deze schoolleiders en leraren dat de feedback uit centrale toetsen zowel informatie op leerlingniveau als op klasniveau en schoolniveau dient te bevatten. Deze schoolleiders en leraren zijn vragende partij voor het ontvangen van de individuele toetsresultaten van de leerlingen en van de gemiddelde toetsresultaten voor de verschillende klassen in de school. Vooral de schoolleiders geven aan geïnformeerd te willen worden over de gemiddelde toetsresultaten van al de leerlingen in de school. Om aan (aspecten van) onderwijskwaliteit te kunnen werken is het dan ook een belangrijk aandachtspunt om scholen feedback op leerling-, klas- en schoolniveau aan te bieden.

Behalve het rapporteren over toetsresultaten tonen de bevraagde schoolleiders en leraren zich bijkomend (gematigd) positief over het ontvangen van het soort toetsvragen dat door de leerlingen al dan niet correct beantwoord werd. De vraag naar deze informatie op klasniveau en schoolniveau blijkt bij de schoolleiders het grootst. Hieruit leiden we af dat helder communiceren en inzetten op 'verwachtingsmanagement' een aandachtspunt is voor het beleid en Steunpunt. Een heldere communicatie en verwachtingsmanagement kunnen immers zorgen voor verwachtingen ten aanzien van feedback die realiseerbaar zijn en bijgevolg valse verwachtingen vermijden.

6.2 HET AFTOETSEN VAN TOETSRESULTATEN AAN EINDTERMEN, HET RAPPORTEREN VAN LEERWINST EN 'FAIRE (SCHOOL)VERGELIJKINGEN' MAKEN DE FEEDBACK BETEKENISVOL

In het licht van relevantie van feedback is het noodzakelijk dat informatie ontsloten wordt die handvaten biedt voor betekenisgeving. De bevraagde schoolleiders en leraren verwachten dat feedback uit centrale toetsen betekenisvol kan worden door toetsresultaten gecombineerd af te toetsen aan een groepscriterium, een standaardcriterium en een zelfcriterium. De schoolleiders beoordeelden de wenselijkheid tot het gebruik van deze criteria op zowel klas- als schoolniveau steeds iets positiever dan de leraren.

De schoolleiders en leraren geven aan dat het belangrijk is dat de feedback uit centrale toetsen toetsresultaten aftoetst aan een te bereiken standaard, i.c. de getoetste eindtermen. Hierbij vinden de schoolleiders en leraren het niet enkel belangrijk te worden geïnformeerd over de mate waarin individuele leerlingen en klassen van de school de getoetste eindtermen bereiken, maar ook over het aantal leerlingen dat in de school de getoetste eindtermen bereikt. Daarnaast blijken de leraren en schoolleiders in grote mate vragende partij voor informatie over de leerwinst die individuele leerlingen boeken. De schoolleiders wensen daarenboven in grote mate bijkomend geïnformeerd te worden over de leerwinst van de leerlingen op klasniveau en schoolniveau en over de mate waarin de eigen school meer of minder bijdraagt aan de leerwinst van leerlingen dan andere scholen.

Daarenboven tonen de schoolleiders en leraren zich eveneens gematigd positief over de verdeling van de toetsresultaten in vaardigheidsniveaus. Ook de mogelijkheid tot het vergelijken van toetsresultaten op leerling-, op school- en in iets mindere mate op klasniveau wordt zowel door de schoolleiders als door de

leraren eerder als positief beschouwd. De schoolleiders en leraren menen bovendien dat dergelijke vergelijkingen nog meer betekenis krijgen als deze vorm krijgen vanuit een eerlijk vergelijkingskader, waarbij vergelijkingen tot stand komen tussen vergelijkbare leerlingen, vergelijkbare klassen en vergelijkbare scholen.

De leraren en – in nog meerdere mate de schoolleiders – geven aan het ook wenselijk te vinden om te weten hoe de toetsresultaten van de school zich verhouden tot de toetsresultaten van vergelijkbare Vlaamse scholen. Om een effectief feedbackgebruik vorm te geven, is het daarom ook zaak vanuit een gebruikersperspectief voldoende oog te hebben voor de validiteit van de feedback. Om gepercipieerd faire en valide schoolvergelijkingen tot stand te kunnen brengen, bevroegen we de schoolleiders over de mate waarin er volgens hen rekening dient gehouden te worden met leerling- en schoolkenmerken wanneer de toetsresultaten van de school op centrale toetsen vergeleken worden met die van andere scholen. Bij het vergelijken van de toetsresultaten van de school adviseren deze schoolleiders rekening te houden met de thuistaal van de leerlingen en de aanwezigheid van leerstoornissen bij de leerlingen. De schoolleiders uit het officieel gesubsidieerd onderwijs tonen zich ten opzichte van de schoolleiders uit het vrij gesubsidieerd onderwijs echter duidelijk minder positief ten aanzien van het rekening houden met de thuistaal in dergelijke vergelijkingen. Daarnaast blijken ook de schoolleiders uit het lager onderwijs een sterkere positieve inschatting te maken dan de schoolleiders uit het secundair onderwijs ten aanzien van het rekening houden met de aanwezigheid van leerstoornissen bij de leerlingen in schoolvergelijkingen. Hierbij stelt zich de vraag in welke mate een verschil in verdeling van leerlingen met (een) leerstoornis(sen) een rol speelt in dergelijke vergelijkingen. Het lijkt dan ook zowel voor methodologen als voor feedback-ontwikkelaars wenselijk hieromtrent bijkomende inzichten te verzamelen.

Verder adviseren al de bevraagde schoolleiders om ook bij de vergelijkingen van de toetsresultaten van de school op centrale toetsen rekening te houden met de kansarme/kansrijke buurt van de school, de leermotivatie van de leerlingen, de aanwezigheid van ouderlijke steun voor de leerlingen, het opleidingsniveau van de ouder(s) van de leerlingen, de stedelijk/landelijke context van de school, de eerdere toetsresultaten van de leerlingen, de mate waarin het schoolteam zich professionaliseert, het onderwijsaanbod van de (secundaire) school, de mate waarin het schoolteam samenwerkt, de ervaring van het schoolteam, het intelligentieniveau van de leerlingen en het personeelsverloop binnen het schoolteam. Hun intuïtieve invulling van leerling- en schoolkenmerken waarmee in het vergelijken van toetsresultaten van scholen rekening dient gehouden te worden, ligt duidelijk anders dan in actuele gangbare praktijken.

Net zoals blijkt uit het kwalitatief gebruikersonderzoek bij schoolleiders en leraren (Molenberghs et al., 2022) zijn ijkingspunten volgens onze respondenten noodzakelijk in de betekenisgeving van de feedback voor het presteren van de school. Het gebruik van een standaardcriterium (zoals het al dan niet behalen van de eindtermen) en een zelfgericht criterium (zoals trends doorheen de tijd) zijn opstappen die voor schoolleiders en leraren duidelijk zijn. Wanneer het echter een groepsgericht criterium betreft (zoals de vergelijking van de toetsresultaten van scholen) blijkt het niet alleen noodzakelijk om vanuit een psychometrisch standpunt toetsresultaten met een hoog betrouwbaarheidsniveau te kunnen terugkoppelen, maar blijkt het ook van groot belang dat deze als 'fair' gepercipieerd worden. Om een effectief feedbackgebruik te (kunnen) stimuleren blijkt het dan ook cruciaal verder in kaart te brengen hoe 'faire' schoolvergelijkingen geoperationaliseerd en gerapporteerd kunnen worden.

6.3 VOORAL LERAREN VERWACHTEN FEEDBACKGEBRUIK OP LEERLINGNIVEAU EN VOORAL SCHOOLLEIDERS VERWACHTEN FEEDBACKGEBRUIK OP SCHOOLNIVEAU

Een belangrijke vraag om inzicht te verwerven in feedbackgebruik uit centrale toetsen binnen scholen is het 'waarom' van feedbackgebruik. We onderscheiden hierin doelen op leerling-, leraar- en schoolniveau. Op leerlingniveau geven de schoolleiders en leraren aan gebruik te willen maken van de feedback uit centrale toetsen als een bron om een zicht te krijgen op sterktes en zwaktes van leerlingen, om remediëring voor leerlingen die getoetste eindtermen nog niet bereiken vorm te geven, om lesonderwerpen te identificeren die extra uitleg behoeven, als (aanvullende) informatie tijdens de klassenraad en om het studieadvies voor leerlingen te motiveren. Tegelijkertijd geven zowel de schoolleiders als leraren aan eerder geen gebruik van de feedback uit centrale toetsen te maken om te beslissen of een leerling mag overgaan naar het volgende (leer)jaar. Hierin tonen respondenten uit het vrij gesubsidieerd onderwijs zich overigens negatiever dan respondenten uit het gemeenschapsonderwijs.

Met betrekking tot ontwikkelingsgerichte doelen op leraarniveau verwachten vooral de leraren dat de feedback uit centrale toetsen bruikbaar is om te reflecteren over het eigen functioneren. Daarnaast zien de schoolleiders op hun beurt mogelijkheden om de feedback te gebruiken om instructie-aanpak te evalueren. Verder blijken zowel de schoolleiders als leraren van mening de feedback niet te gebruiken om keuzes met betrekking tot de eigen werkzaamheden op de school te verantwoorden.

Tot slot verwachten de schoolleiders in de eerste plaats om op schoolniveau met de feedback uit de centrale toetsen aan de slag te gaan als een bijkomende bron van kwaliteitszorg en als een bijkomende manier om aan de kwaliteitsverwachtingen van het Referentiekader onderwijskwaliteit te werken. Zo verwachten de schoolleiders bijvoorbeeld eerder wel met de feedback aan de slag te gaan om concrete actiepunten te formuleren voor het pedagogische beleid, om te bespreken wat het schoolteam verstaat onder goed onderwijs, om aan visie-ontwikkeling te doen binnen de school en om de leer- en onderwijsprestaties te verantwoorden aan de onderwijsinspectie. De leraren spreken in mindere mate de intentie uit om voor deze doelen met de feedback aan de slag te gaan. Hiertegenover zijn de schoolleiders en leraren het eens de feedback eerder niet te gebruiken om nieuwe leerlingen aan te trekken en om naar de buitenwereld toe een positief imago van de school te creëren. Tot slot verwachten de schoolleiders en leraren de feedback niet intensief te gebruiken om de leer- en onderwijsprestaties van de school te verantwoorden aan het schoolbestuur.

In het algemeen kunnen we hieruit concluderen dat de leraren eerder op leerlingniveau en de schoolleiders eerder op schoolniveau verwachten met de feedback uit centrale toetsen aan de slag te gaan. De bijhorende doelen worden veelal slechts in bescheiden mate naar voren geschoven. Het is dan ook mogelijk dat onze respondenten nog geen scherp beeld hebben van mogelijk feedbackgebruik. De implementatie van het systeem van centrale toetsen zal dan ook een proces van lange adem zijn, waarin het belangrijk is goede voorbeelden van feedbackgebruik te laten ontstaan en inspirerend te laten zijn, om op die manier gewenst feedbackgebruik tot op de Vlaamse klasvloeren te kunnen brengen.

6.4 TOEGANG VOOR SPECIFIEKE GEBRUIKERSGROEPEN TOT GERICHTE FEEDBACK

Toegang tot de feedback uit centrale toetsen is een evidente basisvoorwaarde om met de feedback aan de slag te kunnen gaan. Zowel mogelijke onwenselijke effecten als het nastreven van wenselijk doelen voor feedbackgebruik hangen er mee samen.

Om feedback zinvol te kunnen inzetten zouden verschillende teamleden volgens de respondenten toegang moeten hebben tot de niet-geanonimiseerde toetsresultaten van de individuele leerlingen. Dat geldt bijvoorbeeld voor de betrokken vakleraar, de zorgcoördinator/leerlingbegeleider, de schoolleiding en de

ruimere groep van leraren die aan de leerling lesgeven. Vooral de schoolleiders voegen daar ook inzage door de leerlingen zelf en hun ouders aan toe. Die oproep klinkt in het secundair onderwijs sterker dan in het lager onderwijs. Globaal lijkt men in het lager onderwijs minder te pleiten voor het breed delen van feedback over leerlingen binnen de school dan in het secundair onderwijs. Dat leert dat de vraag naar 'wie tot welke feedback toegang geven' mogelijks in het secundair onderwijs een ander antwoord behoort te krijgen dan in het lager onderwijs. Daarentegen zijn zowel de schoolleiders als leraren vragende partij dat zowel externen, zoals leden van de schoolgemeenschap, het schoolbestuur, de pedagogische begeleidingsdienst en de onderwijsinspectie, als al de leraren van het lerarenteam geen (directe) toegang zouden hebben tot de niet-geanonimiseerde toetsresultaten van een individuele leerling.

Ook op klasniveau zijn onze respondenten van mening dat het belangrijk is dat teamleden zoals de betrokken vakleraar, de zorgcoördinator/leerlingbegeleider, de schoolleiding en al de leraren die aan de klas lesgeven toegang tot de toetsresultaten van de deelnemende klasgroepen hebben. De schoolleiders uit het secundair onderwijs vinden het eveneens opportuun dat alle leraren van de betrokken vakgroep toegang tot deze toetsresultaten zouden moeten krijgen. Opmerkelijk hierbij is dat de leraren zelf vragende partij zijn voor een rechtstreekse toegang tot de feedback zonder dat hun schoolleiding daarbij tussenpersoon is, terwijl de schoolleiders hier meer beducht voor blijken. Hiertegenover spreken de schoolleiders en leraren de uitdrukkelijke wens uit dat er absoluut geen toegang mag zijn voor het grote publiek tot de toetsresultaten van de deelnemende klassen en blijken ze eveneens geen voorstaander te zijn van (directe) toegang tot deze feedback voor de schoolgemeenschap, het schoolbestuur, de overheid, de onderwijsinspectie en de ouders van de betrokken leerlingen. Daarenboven vinden voornamelijk de schoolleiders het ook niet nodig dat de betrokken leerlingen en de pedagogische begeleidingsdienst (rechtstreeks) toegang hebben tot deze vorm van feedback. Onze respondenten uit het lager onderwijs blijken verder geen voorstaander te zijn voor rechtstreekse toegang tot deze toetsresultaten voor de leerling zelf en voor de ouders van deze leerling.

Om ook op schoolniveau effectief met de feedback aan de slag te kunnen gaan, schuiven zowel de schoolleiders als de leraren de toegang tot de toetsresultaten van de school voor de betrokken vakleraar, al de leraren van de betrokken vakgroep (voor het secundair onderwijs), de zorgcoördinator/leerlingbegeleider en de schoolleiding naar voor. Voor de schoolleiders en de respondenten uit het lager onderwijs is het bovendien ook sterk wenselijk dat al de leraren van het lerarenteam toegang mogen hebben tot deze informatie. Daar staat tegenover dat de schoolleiders en leraren te kennen geven geen vragende partij te zijn voor de toegang tot toetsresultaten van de school voor het grote publiek.

Verder vinden ze het eveneens niet opportuun dat al de leerlingen van de school en hun ouders toegang hebben tot deze feedback. De respondenten uit het lager onderwijs zijn daarenboven ook van mening dat er geen toegang tot deze feedback wenselijk is voor de betrokken leerlingen en hun ouders.

Om een effectief feedbackgebruik te stimuleren, gewenste effecten te kunnen realiseren en ongewenste effecten te vermijden is het een belangrijk aandachtspunt goed te reflecteren over en doordachte beslissingen te nemen omtrent de – al dan niet – toegang van specifieke gebruikersgroepen tot gerichte vormen van feedback uit de centrale toetsen.

6.5 (SLECHTS) EEN BEPERKT ACTUEEL (INTENSIEF) GEBRUIK VAN (DE RESULTATEN VAN) GESTANDAARDISEERDE TOETSEN

In functie van een zicht op het actueel gebruik van gestandaardiseerde toetsen en de mogelijke ervaringen hiermee, bevroegen we respondenten zowel over het huidige gebruik van gestandaardiseerde toetsen op school als over de mate waarin op de school momenteel intensief aan de slag wordt gegaan met de feedback/resultaten uit die gestandaardiseerde toetsen.

In het algemeen kunnen we concluderen dat scholen in zekere mate gebruikmaken van gestandaardiseerde toetsen. Hierbij geven de bevraagde schoolleiders en leraren uit het lager onderwijs aan – in grote getalen – gebruik te maken van de koepeltoetsen, zoals de IDP- of OVSG-toetsen en van de toetsen in een leerlingvolgsysteem. Het gebruik van paralleltoetsen en gestandaardiseerde diagnose-instrumenten blijkt in het lager onderwijs daarentegen zeer beperkt te zijn. In het secundair onderwijs blijkt voornamelijk – weze het in beperkte mate – gebruik gemaakt te worden van de paralleltoetsen. Van toetsen in een leerlingvolgsysteem en van gestandaardiseerde diagnose-instrumenten wordt in nog mindere mate gebruik gemaakt. Opmerkelijk hierbij is echter dat blijkt dat met de feedback/resultaten van die gestandaardiseerde toetsen slechts in zeer beperkte mate intensief aan de slag wordt gegaan. Dit geldt zowel voor de schoolleiders als leraren, als voor de respondenten uit het lager onderwijs en het secundair onderwijs.

Hoewel zowel in het lager onderwijs als in het secundair onderwijs gestandaardiseerde toetsen voorhanden zijn en er in zekere mate gebruik van gemaakt wordt, blijken er koepelgebonden verschillen ten aanzien van een intensief gebruik van de feedback/resultaten van die gestandaardiseerde toetsen. Desalniettemin is het in het algemeen zorgwekkend vast te stellen dat onze respondenten gemiddeld aangeven dat de

resultaten/feedback uit de beschikbare gestandaardiseerde toetsen amper ingezet worden in kwaliteitszorgprocessen. In het kader van schoolontwikkeling, het primaire doel van het systeem van centrale toetsen, dringt een continue gebruikersmonitoring zich dan ook op. Op die manier is het mogelijk een blijvend venster te creëren op het denken en vooral het handelen van het Vlaamse onderwijsveld bij het gebruik van (de feedback uit) de centrale toetsen én kunnen eventuele gebruikersnoden en professionaliseringsbehoeften gedetecteerd worden.

6.6 SCHOOLINTERNE VERWACHTINGEN INZAKE FEEDBACKGEBRUIK ZIJN EERDER AFWEZIG

Hoewel leraren in eerder gebruikersonderzoek aangaven vragende partij te zijn voor duidelijke (schoolinterne) verwachtingen ten aanzien van feedbackgebruik uit centrale toetsen (Molenberghs et al., 2022), geven de respondenten in dit surveyonderzoek aan dat dergelijke schoolinterne verwachtingen afwezig zijn, en dit in grotere mate in het secundair onderwijs dan in het lager onderwijs.

Van zodra scholen met de voorbereiding op het gebruik van feedback uit centrale toetsen aan de slag gaan, is het voor teamleden van belang over wederzijdse verwachtingen inzake feedbackgebruik in dialoog te gaan en deze duidelijk en helder te expliciteren. Volgens onze respondenten is het hierbij cruciaal dat deze verwachtingen vorm krijgen vanuit een (school)ontwikkelingsperspectief.

6.7 BELEID INZAKE FEEDBACKGEBRUIK IN SCHOLEN WORDT EERDER NIET GEVOERD

‘Gezamenlijke doelgerichtheid’ en ‘ondersteunende relaties en samenwerking’ vormen doorgaans belangrijke pijlers voor het beleidsvoerend vermogen van scholen inzake feedbackgebruik. Toch wordt er op basis van dit gebruikersonderzoek in Vlaamse scholen in het algemeen en in secundaire scholen in het bijzonder nog geen gezamenlijke doelgerichtheid gerapporteerd ten aanzien van feedbackgebruik uit centrale toetsen. Wat betreft ondersteunende relaties en samenwerking inzake feedbackgebruik uit centrale toetsen geven de schoolleiders en leraren in het algemeen aan dat deze in zekere mate aanwezig zijn. In het lager onderwijs blijken collega’s hierbij iets meer ondersteuning en samenwerking te ervaren dan in het secundair onderwijs. Op basis van beide pijlers kunnen we besluiten dat er – momenteel – in Vlaamse scholen amper een beleid inzake feedbackgebruik uit centrale toetsen aanwezig is of gevoerd wordt.

Nochtans slagen scholen met een sterkere gezamenlijke doelgerichtheid en ondersteunende relaties en samenwerking inzake feedbackgebruik er over het algemeen beter in hun schoolbeleid en -praktijk vorm te geven op basis van die feedback. Voor scholen is het dus niet alleen zaak om een gezamenlijk doel inzake feedbackgebruik na te streven, maar ook om dat doel gezamenlijk na te streven. Een gezamenlijke visie en bereidheid ten aanzien van feedbackgebruik in functie van beter lesgeven, betere leerlingresultaten en beter school maken, waarbij feedbackgebruik niet de verantwoordelijkheid is van enkelingen, maar van het ganse schoolteam, zal scholen in staat stellen tot schoolontwikkeling te komen. Het is dan ook een belangrijk aandachtspunt dat zowel scholen hiervoor in hun beleid voldoende oog hebben, als ondersteunings- en professionaliseringsinitiatieven hiervoor voldoende houvast bieden.

6.8 KENMERKEN VAN TRANSFORMATIONEEL LEIDERSCHAP INZAKE FEEDBACKGEBRUIK BLIJKEN SLECHTS IN BEPERKTE MATE IN SCHOLEN AANWEZIG

Schoolleiders kunnen als transformationeel leider belangrijke betekenis-gevers zijn om met feedback aan de slag te gaan en in die hoedanigheid ook de katalysator zijn in het opbouwen van een cultuur van feedbackgebruik. Op basis van dit gebruikersonderzoek kunnen we in het algemeen echter stellen dat er op scholen slechts in beperkte mate kenmerken van zulk transformationeel leiderschap aanwezig zijn. Opvallend hierbij is dat de schoolleiders het eigen transformationeel leiderschap positiever inschatten dan de leraren. In het lager onderwijs is transformationeel leiderschap meer uitgesproken aanwezig dan in het secundair onderwijs.

De vaststelling dat er in het algemeen geen sterk transformationeel leiderschap inzake feedbackgebruik in Vlaamse scholen aanwezig is, doet concluderen dat ook bij schoolleiders ondersteuning en professionalisering hieromtrent van belang zal zijn. Het streven naar een effectief feedbackgebruik heeft met andere woorden niet alleen oog voor de ondersteuningsbehoeften en professionaliseringsnoden van leraren, maar ook voor die van schoolleiders.

6.9 SCHOOLEXTERNE VERWACHTINGEN INZAKE FEEDBACKGEBRUIK ZIJN EERDER AFWEZIG

Net zoals schoolinterne verwachtingen, zijn schoolexterne verwachtingen op dit moment in de beleving van de respondenten nog eerder afwezig. Voornamelijk voor de leraren en voor de respondenten uit het secundair onderwijs is het niet duidelijk wat de overheid, de onderwijsinspectie, de pedagogische begeleidingsdienst, de scholengemeenschap en het schoolbestuur verwachten inzake feedbackgebruik. Communicatie en dialoog tussen feedbackgebruikers en betreffende instanties en organisaties kan duidelijkheid scheppen in schoolexterne verwachtingen. Het is dan ook van belang om in de verdere vormgeving van het systeem van centrale toetsen hier nodige aandacht aan te schenken. Onderzoek leert immers dat heldere en expliciete verwachtingen essentieel zijn om tot feedbackgebruik te komen (Van Gasse, Vanhoof, & de Vos, 2014). Verwachtingen expliciteren de verantwoordelijkheid die scholen hebben ten aanzien van de leerresultaten van hun leerlingen en dat blijkt essentieel in het kader van schoolontwikkeling (Schildkamp et al., 2016).

6.10 SCHOOLLEIDERS EN LERAREN STAAN GLOBAAL NEUTRAAL TEGENOVER (FEEDBACKGEBRUIK UIT) CENTRALE TOETSEN

De schoolleiders en leraren rapporteren op het moment van bevragen slechts in beperkte mate een positieve affectieve attitude. Hoewel de komst van centrale toetsen hen (voorlopig) niet helemaal enthousiast maakt, schrikt de komst ervan hen ook niet af. Negatieve gevoelens ten aanzien van de komst van centrale toeten laten zich sterker optekenen bij de respondenten uit het secundair onderwijs dan bij de respondenten uit het lager onderwijs. Wellicht kan het reeds meer uitgesproken actueel gebruik van gestandaardiseerde toetsen en de daarmee gepaard gaande ervaringen in het lager onderwijs deze vaststelling verklaren. Ook de cognitieve houding ten aanzien van feedbackgebruik is in het algemeen eerder neutraal. Zo blijkt dat de schoolleiders en de leraren slechts in beperkte mate aangeven een uitgesproken voorstander te zijn van de invoering van het systeem van centrale toetsen en feedbackgebruik niet meteen als bijzonder waardevol te percipiëren. Daar staat tegenover dat zij wel van mening zijn met de feedback aan de slag te willen gaan. Daarnaast blijken de schoolleiders iets positiever dan de leraren ten opzichte van de meerwaarde van feedbackgebruik uit centrale toetsen.

Hoe schoolleiders en leraren staan ten opzichte van feedbackgebruik bepaalt in sterke mate in hoeverre de school tot effectief feedbackgebruik komt. Aangezien de attitude van de hier bevraagde toekomstige gebruikers (i.c. schoolleiders en leraren) niet onverdeeld positief blijkt, is een belangrijk element in de

professionele ontwikkeling van schoolleiders en leraren dus de ondersteuning in hun denken ten aanzien van feedbackgebruik, in hun attitude ten aanzien van feedbackgebruik en in hoe ze reageren op (het gebruik) van centrale toetsen.

6.11 SCHOOLLEIDERS EN LERAREN TONEN ZICH (IN ZEKERE MATE) WELWILLENDHEID INZAKE FEEDBACKGEBRUIK

De zelfdeterminatietheorie (zie Theoretische uitgangspunten) onderscheidt vijf types van motivatie, waarbij de beschrijving van deze types vertrekt vanuit het onderscheid tussen autonome en gecontroleerde motivatie enerzijds en amotivatie anderzijds. Onze respondenten blijken in het algemeen eerder autonoom gemotiveerd voor feedbackgebruik. Ze tonen met andere woorden een zeker mate van welwillendheid om met de feedback uit centrale toetsen aan de slag te gaan. Hierbij blijken de schoolleiders sterker autonoom gemotiveerd dan de leraren, en de respondenten uit het lager onderwijs sterker autonoom gemotiveerd dan de respondenten uit het secundair onderwijs. Deze autonome motivatie voor feedbackgebruik wordt eerder vanuit een geïdentificeerde regulatie gevoed dan vanuit een intrinsieke regulatie. Onze respondenten geven hiermee aan dat een toekomstig feedbackgebruik een volgens hen professioneel relevant doel kan dienen. Daarnaast geven de schoolleiders en leraren in het algemeen aan eerder niet gecontroleerd gemotiveerd te zijn voor feedbackgebruik. Zowel de schoolleiders als de leraren blijken noch geïntrojecteerd gereguleerd, noch extern gereguleerd tot feedbackgebruik. De schoolleiders en leraren ervaren met andere woorden momenteel noch interne, noch externe druk om met de feedback aan de slag te gaan. Verder blijken de schoolleiders en leraren eerder niet geamotiveerd voor feedbackgebruik. Op basis van beider percepties kunnen we dus motivatie voor feedbackgebruik vaststellen.

Op basis van deze bevindingen blijft het een permanent aandachtspunt autonomie-ondersteunend te handelen en te communiceren met toekomstige feedbackgebruikers. Wanneer feedbackgebruikers zich autonoom kunnen voelen, ontwikkelen ze doorgaans ook de eigen autonome motivatie sterker. Het is net dit type motivatie dat doorgaans tot meer positieve uitkomsten (i.c. feedbackgebruik) aanleiding geeft (Vansteenkiste et al., 2007).

6.12 SCHOOLLEIDERS EN LERAREN VOELEN ZICH (EERDER WEL) BEKWAAM INZAKE FEEDBACKGEBRUIK

Op basis van de verzamelde data kunnen we stellen dat de schoolleiders en de leraren zich (eerder wel) bekwaam voelen ten aanzien van het interpreteren van feedback, het diagnosticeren van feedback en het vervolgens bedenken van gepaste acties. De persoonlijke doelmatigheid inzake verkennen van en overleggen over de feedback beoordelen de schoolleiders en leraren eerder als neutraal. Hierbij menen ze echter zich eerder wel bekwaam te voelen om met leerlingen over de feedback te communiceren en met collega's de feedback te bespreken, maar geven ze tegelijkertijd aan geen goed zicht te hebben op de bruikbaarheid van de feedback en op wat uit de feedback geleerd kan worden.

Een dergelijke positieve inschatting van de respondenten ten aanzien van het geloof in eigen kunnen met betrekking tot feedbackgebruik is opvallend. Uit de meer diepgaande wetenschappelijke literatuur komt immers naar voor dat schoolleiders en leraren in vele gevallen twijfelen aan hun eigen bekwaamheid om met feedback aan de slag te gaan binnen de school (zie o.a. Kowalski & Lasley, 2009). Ondanks deze positieve zelfinschatting lijkt het ons aangewezen de persoonlijke doelmatigheidsbeleving van deze feedbackgebruikers te blijven monitoren en hen te blijven ondersteunen in hun feedbackgebruik. Dergelijke ondersteuning kan zowel gericht zijn op het ontwikkelen van kennis met betrekking tot centrale toetsen en feedbackgebruik, als op het verwerven van vaardigheden met betrekking tot feedbackgebruik.

6.13 SCHOOLLEIDERS EN LERAREN ZIJN BEKOMMERD OVER HET SYSTEEM VAN CENTRALE TOETSEN EN VRAGEN NAAR INFORMATIE OVER FEEDBACKGEBRUIK

Om de bereidheid tot feedbackgebruik in kaart te brengen, brachten we in dit gebruikersonderzoek ook de bekommernissen die bij de respondenten leven in kaart. Op basis van de survey kunnen we in het algemeen vaststellen dat zelf-bekommernissen sterker aanwezig zijn dan bekommernissen ten aanzien van de taak of andere-bekommernissen (zie Theoretische uitgangspunten). De schoolleiders en leraren vragen zich met andere woorden vooral af wat feedbackgebruik voor hen zal betekenen. Een grotere aanwezigheid van zelf-bekommernissen kan volgens dit theoretisch kader wijzen op weerstand bij de respondenten of op een gebrek aan voorbereiding en informatie.

Wat betreft zelf-bekommernissen worden bekommernissen op het niveau van informatie gemiddeld het meest vaak gerapporteerd. De bevraagde schoolleiders en leraren geven blijk van interesse in bijkomende informatie over feedbackgebruik. Zo geven ze aan meer informatie te willen hebben over hoe met de feedback aan de slag te kunnen en de mogelijkheden van feedback te willen verkennen. Naast bekommernissen ten aanzien van (gebrek aan) informatie, leven er bij de schoolleiders en leraren ook in zekere mate bekommernissen over hoeveel tijd feedbackgebruik zal vergen (bekommernissen op het niveau van persoonlijke betrokkenheid). Op het niveau van taak-bekommernissen kunnen we, naar analogie met de conclusie met betrekking tot de persoonlijke doelmatigheid van de schoolleiders en leraren ten aanzien van feedbackgebruik, vaststellen dat onze respondenten niet vrezen dat ze niet in staat zijn om met de feedback aan de slag te gaan. Daarnaast blijken het vooral de schoolleiders die zich ook bekommeren over de 'andere(n)'. Hierbij geven de schoolleiders aan te willen verkennen hoe anderen met de feedback aan de slag gaan en ook binnen en buiten de school te willen samenwerken voor feedbackgebruik. Bekommernissen over de gevolgen van feedbackgebruik voor leerlingen zijn zowel bij de schoolleiders als bij de leraren slechts in geringe mate aanwezig.

Hoewel het succes van veranderingen (i.c. feedbackgebruik uit centrale toetsen) van vele factoren afhangt, is er eveneens een aanzienlijke hoeveelheid onderzoek die wijst op het belang van rekening te houden met de bekommernissen van degenen die direct bij de verandering betrokken zijn (o.a. Hall & Hord, 2020). Hierbij blijkt vooral de manier waarop schoolleiders en leraren betekenis geven aan de verandering van belang (März & Kelchtermans, 2013). Het is dan ook van belang de vraag naar informatie vanuit het werkveld te beantwoorden en adequate informatie te verschaffen. Daarnaast is het ook aan te bevelen deze bekommernissen blijvend te monitoren om kort op de bal te kunnen spelen en optimalisaties door te kunnen voeren.

6.14 OOK VERWACHTE ONWENSELIJKE EFFECTEN OP LERAAR-, SCHOOL- EN SYSTEEMNIVEAU

Verwachte onwenselijke effecten van centrale toetsen werden zowel bevraagd op leerling- en leraarniveau, als op school- en onderwijssysteemniveau. In het algemeen kunnen we op basis van dit gebruikersonderzoek besluiten dat de schoolleiders en leraren (slechts) in beperkte mate onwenselijke effecten op leerlingenniveau verwachten, en in meerdere mate op leraar-, school- en onderwijssysteemniveau. Het gaat hier telkens om globale conclusies, gebaseerd op gemiddelde surveyscores. Dat doet echter niets af aan de vaststelling dat er in beide richtingen telkens best wat variatie is in de antwoorden van respondenten.

De bevraagde schoolleiders en leraren zijn globaal eerder neutraal in hun verwachtingen inzake onwenselijke effecten van feedbackgebruik op leerlingniveau. Wel verwachten vooral de respondenten uit het secundair onderwijs en uit het vrij gesubsidieerd onderwijs dat centrale toetsen een mogelijke bedreiging kunnen vormen voor de brede evaluatie van leerlingen. Op leraarniveau zijn de respondenten bekommerd dat centrale toetsen voor extra planlast kunnen zorgen, leraren kunnen aanzetten om leerlingen extra voor te bereiden op de centrale toetsen, tot minder aandacht kunnen leiden voor leerinhouden die niet getoetst worden in centrale toetsen en voor minder ruimte kunnen zorgen voor een persoonlijke inbreng van leraren in de klaspraktijk. Daarnaast zijn de leraren bezorgd dat de centrale toetsen kunnen zorgen voor onnodige twijfel over de evaluatiecompetenties van leraren. Vooral respondenten uit het secundair onderwijs zijn er daarenboven voor beducht dat centrale toetsen een mogelijke bedreiging voor de pedagogische vrijheid van leraren kunnen vormen. Opvallend hierbij is dat de respondenten uit het secundair onderwijs steevast in meerdere mate onwenselijke effecten verwachten van centrale toetsen voor leraren dan de respondenten uit het lager onderwijs. Op schoolniveau verwachten de schoolleiders en leraren naast wenselijke ook mogelijke onwenselijke effecten van centrale toetsen. Hierbij vrezen ze dat centrale toetsen tot openbare *rankings* van scholen kunnen leiden, tot competitie tussen scholen kunnen leiden en voor te grote druk kunnen zorgen op het schoolteam om goede prestaties te realiseren op de centrale toetsen. Ook onwenselijke effecten van centrale toetsen op schoolniveau worden door de respondenten uit het secundair onderwijs in meerdere mate verwacht dan door de respondenten uit het lager onderwijs. Zowel de schoolleiders als de leraren verwachten tot slot ook in zeker mate onwenselijke effecten van centrale toetsen op onderwijssysteemniveau. Hierbij verwachten beide respondentengroepen dat de centrale toetsen de focus van de onderwijsinspectie te sterk op outputdata zullen richten en tot grotere segregatie tussen scholen zullen leiden. Daarnaast zijn de schoolleiders en leraren bezorgd dat centrale toetsen tot oneerlijke vergelijkingen tussen scholen leiden. Verder vrezen ze dat centrale toetsen leerprestaties van leerlingen kunnen verengen tot de toetsresultaten van de leerlingen op de centrale toetsen voor wiskunde en Nederlands en de invulling van onderwijskwaliteit bijgevolg vernauwt tot cognitieve leerresultaten. Bovendien zijn de respondenten uit het secundair onderwijs ongerust dat centrale toetsen een bedreiging voor de pedagogische vrijheid van scholen kunnen vormen. Ook nu weer blijken de respondenten uit het secundair onderwijs in sterkste mate onwenselijke effecten van centrale toetsen te verwachten.

Ook ten aanzien van verwachte onwenselijke effecten van centrale toetsen is het bijgevolg van belang toekomstige feedbackgebruikers duidelijk te informeren over de doelen van het systeem van centrale

toetsen en het wettelijk kader waarbinnen het systeem uitgetekend wordt, zodat primaire doelen voor feedbackgebruik helder worden en de verwachtingen van onwenselijke effecten getemperd kunnen worden.

6.15 AFRONDEND: BELANG VAN INFORMATIE, ONDERSTEUNING EN PROFESSIONALISERING

Onze ambitie met dit gebruikersonderzoek was in kaart brengen in welke mate randvoorwaarden voor een effectief feedbackgebruik uit centrale toetsen in deze fase van het implementatieproces (begin 2022) in Vlaamse scholen, door de ogen van schoolleiders en leraren, (reeds) aanwezig zijn. Het is belangrijk bij de beschrijving van de resultaten mee te geven dat de meerderheid van de bevraagde respondenten ongetwijfeld nog botsten op het hypothetisch karakter van zowel het systeem van (werken met) centrale toetsen als van de aangeboden stellingen en items. Wellicht zullen de inschattingen van deze respondentengroepen nog wijzigen, naarmate de invulling omtrent het systeem van centrale toetsen gedetailleerder wordt en feedback uit centrale toetsen beschikbaar zal zijn. Dit surveyonderzoek fungeert daarom als een nulmeting om in een latere fase evoluties en trends in de percepties van schoolleiders en leraren hieromtrent te kunnen capteren. De standaardafwijkingen wijzen in deze fase alvast op aanzienlijke verschillen tussen de respondenten. Er is met andere woorden niet één gemiddelde realiteit, maar er zijn diverse realiteiten in het Vlaamse onderwijsveld.

Centrale toetsen voor het onderwijsveld vergen gedragenheid door het onderwijsveld. Het is dan ook van cruciaal belang de vinger (blijvend) aan de pols te houden. Respondenten blijken er momenteel een behoorlijke bereidwillige houding op na te houden ten aanzien van een ontwikkelingsgericht gebruik van centrale toetsen. Tegelijkertijd blijkt er een grote vraag en behoefte bij de respondenten aan bijkomende informatie van het beleid over het doel en de impact van (het gebruik van) centrale toetsen, en aan communicatie over het proces van, de mogelijkheden van en verwachtingen bij feedbackgebruik uit centrale toetsen. Daarnaast blijken ook enkele stimulerende randvoorwaarden in Vlaamse scholen nog niet optimaal aanwezig. Naast een doorgedreven informatiecampagne zal het ook van belang zijn toekomstige feedbackgebruikers individueel, collectief, specifiek en gericht te ondersteunen en te professionaliseren.

BIBLIOGRAFIE

Abelmann, C., Elmore, R., Even, J., Kenyon, S., & Marshall, J. (1999). When Accountability Knocks, Will Anyone Answer?

Acquah, D. (2013). *Examining School Accountability*. Manchester, UK: Centre for Education Research and Policy.

Atkinson, R. C., & Geiser, S. (2009). Reflections on a century of college admissions tests. *Educational researcher, 38*(9), 665-676.

Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational researcher, 36*(5), 258-267.

Bandura, A. (1997). *Self Eflicacy. The Exercise of Control*. New York: Freeman.

Bandura, A., & Walters, R. H. (1977). *Social learning theory*. NY, Prentice Hall: Englewood Cliffs

Barrezeele, G. (2012). *Informatiemanagement: een nieuw tijdperk, een nieuwe aanpak: praktijkgids voor geavanceerd informatiebeheer*. Leuven: Acco.

Bishop, J. H., & Wößmann, L. (2004). Institutional effects in a simple model of educational production. *Education Economics, 12*(1), 17-38.

Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework 1. *European law journal, 13*(4), 447-468.

Breiter, A., & Light, D. (2006). Data for school improvement: Factors for designing effective information systems to support decision-making in schools. *Journal of Educational Technology & Society, 9*(3), 206-217.

Brunner, C., Fasca, C., Heinze, J., Honey, M., Light, D., Mardinach, E., & Wexler, D. (2005). Linking data and learning: The Grow Network study. *Journal of Education for Students placed at Risk, 10*(3), 241-267.

Campbell, C., & Levin, B. (2009). Using data to support educational improvement. *Educational Assessment Evaluation and Accountability, 21*(1), 47-65. doi:10.1007/s11092-008-9063-x

Castellano, K. E., & Ho, A. D. (2015). Practical differences among aggregate-level conditional status metrics: From median student growth percentiles to value-added models. *Journal of Educational and Behavioral Statistics, 40*(1), 35-68.

Charalambous, C. Y., & Philippou, G. N. (2010). Teachers' concerns and efficacy beliefs about implementing a mathematics curriculum reform: integrating two lines of inquiry. *Educational studies in Mathematics, 75*(1), 1-21.

Cizek, G. J. (2005). High-stakes testing: Contexts, characteristics, critiques, and consequences. In R. P. Phelps (Ed.), *Defending standardized testing* (pp. 23-54). Mahaj, NJ/London: Lawrence Erlbaum Associates Publishers.

Coburn, C. E., & Talbert, J. E. (2006). Conceptions of evidence use in school districts: Mapping the terrain. *American journal of Education, 112*(4), 469-495.

Coburn, C. E., Touré, J., & Yamashita, M. (2009). Evidence, interpretation, and persuasion: Instructional decision making at the district central office. *Teachers College Record, 111*(4), 1115-1161.

Coburn, C. E., & Turner, E. O. (2011). Research on data use: A framework and analysis. *Measurement: Interdisciplinary Research & Perspective, 9*(4), 173-206.

Coghlan, M., & Desurmont, A. (2007). *School Autonomy in Europe: Policies and Measures*. Brussel: Eurydice.

Collins, S., Reiss, M., & Stobart, G. (2010). What happens when high-stakes testing stops? Teachers' perceptions of the impact of compulsory national testing in science of 11-year-olds in England and its abolition in Wales. *Assessment in Education: Principles, Policy & Practice, 17*(3), 273-286.

Cosner, S. (2011). Supporting the initiation and early development of evidence-based grade-level collaboration in urban elementary schools: Key roles and strategies of principals and literacy coordinators. *Urban Education, 46*(4), 786-827.

Datnow, A., Park, V., & Kennedy-Lewis, B. (2012). High school teachers' use of data to inform instruction. *Journal of Education for Students placed at Risk 17*(4), 247-265.

Datnow, A., Park, V., & Wohlstetter, P. (2007). *Achieving with data: How high-performance school systems use data to improve instructions for elementary students*. Los Angeles: University of Southern California, Center on Educational Governance.

De Meyer, I., Janssens, R., Warlop, N., Valcke, M., De Wever, B., & Van Keer, H. (2020). *Leesvaardigheid van 15-jarigen. Vlaams rapport PISA 2018*. Gent: Universiteit Gent.

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry, 11*(4), 227-268.

Denis, J., Carpentier, N., Laenen, I., Willem, L., Janssen, R., & Aesaert, K. (2018). *Peiling Frans in het basisonderwijs - 2017. Eindrapport.*

Diamond, J. B., & Spillane, J. P. (2004). High-stakes accountability in urban elementary schools: Challenging or reproducing inequality? *Teachers College Record, 106*(6), 1145-1176.

Dockx, J., Van Landeghem, G., Aesaert, K., Van Damme, J., & De Fraine, B. (2019). *Begrijpend lezen van het vierde naar het zesde leerjaar. Herhalingsmeting van PIRLS in 2018 vergeleken met PIRLS 2016*. Leuven: KU Leuven, Centrum voor Onderwijsefffectiviteit en -evaluatie.

Ehren, M. C. M., & Swanborn, M. S. L. (2012). Strategic data use of schools in accountability systems. *School Effectiveness and School Improvement, 23*(2), 257-280. doi:10.1080/09243453.2011.652127

Elmore, R. F. (2005a). *Accountable leadership.* Paper presented at the The Educational Forum.

Elmore, R. F. (2005b). Accountable leadership. *The Educational Forum, 69*(2), 134-142.

Elmore, R. F., Abelmann, C. H., & Fuhrman, S. H. (1996). The new accountability in state education reform: From process to performance. *Holding schools accountable: Performance-based reform in education*, 65-98.

Faddar, J., Appels, L., Merckx, B., Boeve-de Pauw, J., Delrue, K., De Maeyer, S., & Van Petegem, P. (2020). *Vlaanderen in TIMSS 2019. Wiskunde- en wetenschapsprestaties van het vierde leerjaar in internationaal perspectief en doorheen de tijd*. Antwerpen: Universiteit Antwerpen.

Figlio, D. N., & Page, M. E. (2003). Can school choice and school accountability successfully coexist? In *The economics of school choice* (pp. 49-66). Chicago: University of Chicago Press.

Firestone, W. A., Monfils, L., Camilli, G., Schorr, R. Y., Hicks, J., & Mayrowetz, D. (2002). The ambiguity of test preparation. A multimethod analysis in one state. *Teachers College Record, 104*, 1485-1523.

Fryer, R. G. (2013). Teacher incentives and student achievement: Evidence from New York City public schools. *Journal of Labor Economics, 31*(2), 373-407.

Geijsel, F. P., Krüger, M. L., & Sleegers, P. J. (2010). Data feedback for school improvement: The role of researchers and school leaders. *The Australian Educational Researcher, 37*(2), 59-75.

George, A. A., Hall, G. E., & Stiegelbauer, S. M. (2006). *Measuring implementation in schools: The Stages of Concern*

Questionnaire. Austin, TX: : Southwest Educational Development Laboratory.

Gong, B., Perie, M., & Dunn, J. (2006). *Using student longitudinal growth measures for school accountability under No Child Left Behind: An update to inform design decisions*. Center for Assessment–NCLB growth update.

Hall, G. E., George, A. A., & Rutherford, W. L. (1979). *Measuring stages of concerns about the innovation: A manual for use of the SoC Questionnaire*. Austin: University of Texas.

Hall, G. E., & Hord, S. M. (1979). *Change in schools: Facilitating the process*. New York: State University of New York Press.

Hall, G. E., & Hord, S. M. (1987). *Change in schools: Facilitating the process*. New York: State University of New York Press.

Hall, G. E., & Hord, S. M. (2006). *Implementing change: Patterns, principles, and potholes* (2nd ed.). Boston, MA: Pearson Education Inc.

Hall, G. E., & Hord, S. M. (2020). *Implementing change. Patterns, principles, and potholes* (5 ed.). Boston: Allyn & Bacon.

Hamilton, L. S., Stecher, B. M., & Klein, S. P. (2002). *Making sense of test-based accountability in education*. Santa Monica, CA: Rand Corporation.

Harris, D. N. (2011). Value-added measures and the future of educational accountability. *Science, 333*(6044), 826-827.

Hofman, R. H., Dijkstra, N. J., & Hofman, W. H. A. (2009). School self-evaluation and student achievement. *School Effectiveness and School Improvement, 20*(1), 47-68. doi:10.1080/09243450802664115

Hord, S. M., Rutherford, W. L., Austin, L. H., & Hall, G. E. (1987). *Taking charge of change*. Alexandria, Va: Association for Supervision and Curriculum Development.

Hubbard, L., Datnow, A., & Pruyn, L. (2014). Multiple initiatives, multiple challenges: The promise and pitfalls of implementing data. *Studies in Educational Evaluation, 42*, 54-62. doi:10.1016/j.stueduc.2013.10.003

Ikemoto, G., & Marsh, J. (2007). Cutting through the "data driven" mantra: Different conceptions of data-driven decision making. In P. A. Moss (Ed.), *Evidence and Decision Making* (pp. 105-131). Chicago: National Society for the Study of Education.

Janssens, F. J. G., Rekers-Mombarg, L., & Lacor, E. (2014). *Leerwinst en toegevoegde waarde in het primair onderwijs*. Ministerie van Onderwijs, Cultuur en Wetenschappen.

Jimerson, J. B. (2016). How are we approaching data-informed practice? Development of the Survey of Data Use and Professional Learning. *Educational Assessment Evaluation and Accountability, 28*(1), 61-87. doi:10.1007/s11092-015-9222-9

Johnson, R. B. (1998). Toward a theoretical model of evaluation utilization. *Evaluation and Program Planning, 21*(1), 93-110.

Kelly, A. (2004). *Benchmarking for school improvement. A practical guide for comparing and achieving effectiveness*. London/New York: Routledge Falmer.

Kerr, K. A., Marsh, J. A., Ikemoto, G. S., Darilek, H., & Barney, H. (2006). Strategies to promote data use for instructional improvement: Actions, outcomes, and lessons fromthree urban districts. *American journal of Education, 112*(4), 496-520.

Kowalski, T., & Lasley, T. (2009). Part I: Theoretical and Practical Perspectives. In T. Kowalski & T. Lasley (Eds.), *Handbook of data-based decision making in education* (pp. 3-86). New York: Routledge.

Krathwohl, D., Bloom, B., & Masia, B. (1971). *The classification of educational goals (Handbook II, Affective domain)*. New York: Mc Kay.

Lai, M. K., & Schildkamp, K. (2013). Data-based decision making: An overview. *Data-based decision making in education*, 9-21.

Lee, M., Louis, K. S., & Anderson, S. (2012). Local education authorities and student learning: The effects of policies and practices. *School Effectiveness and School Improvement, 23*(2), 133-158.

Levin, J. A., & Datnow, A. (2012). The principal role in data-driven decision making: Using case-study data to develop multi-mediator models of educational reform. *School Effectiveness and School Improvement, 23*(2), 179-201.

Loeb, S. (2013). How can value-added measures be used for teacher improvement. *What we know series: Value-added methods and applications. Stanford, CA: Carnegie Knowledge Network*.

Lysne, A. (1984). Grading of student's attainment: Purposes and functions. *Scandinavian Journal of Educational Research, 28*(3), 149-165.

Maier, U. (2010). Accountability policies and teachers' acceptance and usage of school performance feedback–a comparative study. *School Effectiveness and School Improvement, 21*(2), 145-165.

Mandinach, E. B. (2012). A Perfect Time for Data Use: Using Data-Driven Decision Making to Inform Practice. *Educational Psychologist, 47*(2), 71-85. doi:10.1080/00461520.2012.667064

Mandinach, E. B., Honey, M., Light, D., & Brunner, C. (2008). A Conceptual Framework for Data-Driven Decision-Making. In E. Mandinach & M. Honey (Eds.), *Data-Driven School Improvement: Linking Data and Learning*. New York, Y: Teachers College Press.

Marlow, R., Norwich, B., Ukoumunne, O. C., Hansford, L., Sharkey, S., & Ford, T. (2014). A comparison of teacher assessment (APP) with standardised tests in primary literacy and numeracy (WIAT-II). *Assessment in Education: Principles, Policy & Practice, 21*(4), 412-426.

Marsh, J. A., Pane, J. F., & Hamilton, L. S. (2005). *Making sense of data-driven decision making in education*. Santa Monica, CA: Rand Education.

Marsh, J. A., Pane, J. F., & Hamilton, L. S. (2006). *Making Sense of Data-Driven Decision Making in Education. Evidence from Recent Rand Research. Occasional Paper*. Santa Monica, CA: Rand Corporation.

März, V., & Kelchtermans, G. (2013). Sense-making and structure in teachers' reception of educational reform. A case study on statistics in the mathematics curriculum. *Teaching and Teacher Education, 29*, 13-24.

McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *The journal of educational research, 95*, 203-213.

Molenberghs, G., Van Gasse, R., De Maeyer, S., & Vanhoof, J. (2022). *Feedbackgebruik uit centrale toetsen. Gebruikersonderzoek: diepte-interviews bij schoolleiders en leraren.* . Gent: Steunpunt Centrale Toetsen in Onderwijs.

Neumann, M., Trautwein, U., & Nagy, G. (2011). Do central examinations lead to greater grading comparability? A study of frame-of-reference effects on the University entrance qualification in Germany. *Studies in Educational Evaluation, 37*(4), 206-217.

O'Day, J. A. (2002). Complexity, accountability, and school improvement. *Harvard Educational Review, 72*(3), 293-329. doi:10.17763/haer.72.3.021q742t8182h238

OECD. (2013). *Synergies for better learning. An international perspective on evaluation and assessment*. Paris: OECD.

Park, V., Daly, A. J., & Guerra, A. W. (2013). Strategic framing: How leaders craft the meaning of data use for equity and learning. *Educational Policy, 27*(4), 645-675.

Parkes, J., & Stevens, J. J. (2003). Legal issues in school accountability systems. *Applied measurement in education, 16*(2), 141-158.

Penninckx, M., Vanhoof, J., Quintelier, A., De Maeyer, S., & Van Petegem, P. (2017). *Zicht op Leerwinst*. Leuven: Acco.

Pierce, R., & Chick, H. (2011). Teachers' intentions to use national literacy and numeracy assessment data: A pilot study. *The Australian Educational Researcher, 38*(4), 433-447.

Pierce, R., Chick, H., & Gordon, I. (2013). Teachers' perceptions of the factors influencing their engagement with statistical reports on student achievement data. *Australian Journal of Education, 57*(3), 237-255.

Pollock, K., & Winton, S. (2016). Juggling multiple accountability systems: how three principals manage these tensions in Ontario, Canada. *Educational Assessment, Evaluation and Accountability, 28*(4), 323-345.

Popham, W. J. (1999). Why standardized tests don't measure educational quality. *Educational leadership, 56*, 8-16.

Ray, A. (2006). *School value added measures in England: A paper for the OECD project on the development of value-added models in education systems.* Department for Education and Skills, London.

Rogers, S., Barblett, L., & Robinson, K. (2018). Parent and teacher perceptions of NAPLAN in a sample of Independent schools in Western Australia. *The Australian Educational Researcher, 45*(4), 493-513.

Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and social Psychology*(57), 749-761.

Sach, E. (2013). An exploration of teachers' narratives: what are the facilitators and constraints which promote or inhibit 'good'formative assessment practices in schools? *Education 3-13, 43*(3), 322-335.

Sanbonmatsu, D. M., & Fazio, R. H. (1990). The role of attitudes in memory-based decision making. *Journal of Personality and social Psychology, 59*(4), 614.

Saunders, L. (1999). A Brief History of Educational'Value Added': How did we get to where we are? *School Effectiveness and School Improvement, 10*(2), 233-256.

Schildkamp, K. (2019). Data-based decision-making for school improvement: Research insights and gaps. *Educational Research, 61*(3), 257-273. doi:10.1080/00131881.2019.1625716

Schildkamp, K., Ehren, M., & Lai, M. K. (2012). Editorial article for the special issue on data-based decision making around the world: from policy to practice to results. *School Effectiveness and School Improvement, 23*(2), 123-131. doi:10.1080/09243453.2011.652122

Schildkamp, K., Karbautzki, L., & Vanhoof, J. (2014). Exploring data use practices around Europe: Identifying enablers and barriers. *Studies in Educational Evaluation, 42*, 15-24. doi:10.1016/j.stueduc.2013.10.007

Schildkamp, K., & Kuiper, W. (2010). Data-informed curriculum reform: Which data, what purposes, and promoting and hindering factors. *Teaching and Teacher Education, 26*(3), 482-496. doi:10.1016/j.tate.2009.06.007

Schildkamp, K., & Poortman, C. (2015). Factors influencing the functioning of data teams. *Teachers College Record, 117*, 1-42.

Schildkamp, K., Poortman, C. L., & Handelzalts, A. (2016). Data teams for school improvement. *School Effectiveness and School Improvement, 27*(2), 228-254. doi:10.1080/09243453.2015.1056192

Schildkamp, K., Rekers-Mombarg, L. T. M., & Harms, T. J. (2012). Student group differences in examination results and utilization for policy and school development. *School Effectiveness and School Improvement, 23*(2), 229-255. Retrieved from https://www.tandfonline.com/doi/pdf/10.1080/09243453.2011.652123?needAccess=true

Schildkamp, K., & Teddlie, C. (2008). School performance feedback systems in the USA and in The Netherlands: A comparison. *Educational Research and Evaluation, 14*(3), 255-282.

Schildkamp, K., & Visscher, A. (2010). The use of performance feedback in school improvement in Louisiana. *Teaching and Teacher Education, 26*(7), 1389-1403. doi:10.1016/j.tate.2010.04.004

Schildkamp, K., Visscher, A., & Luyten, H. (2009). The effects of the use of a school self-evaluation instrument. *School Effectiveness and School Improvement, 20*(1), 69-88. doi:10.1080/09243450802605506

Sireci, S. G. (2005). The most frequently unasked questions about testing. *Defending standardized testing*, 111-121.

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

Steunpunt. (2022a). *Leerprestaties en leerwinst.* Steunpunt Centrale Toetsen in Onderwijs.

Steunpunt. (2022b). *Toetsen voor onderwijsontwikkeling. Krachtlijnen voor de centrale toetsen in Vlaanderen.* Steunpunt Centrale Toetsen in Onderwijs.

Sutherland, S. (2004). Creating a culture of data use for continuous improvement: A case study of an Edison Project school. *American Journal of Evaluation, 25*(3), 277-293.

van Ackeren, I., Block, R., Klein, E. D., & Kühn, S. M. (2012). The impact of statewide exit exams: A descriptive case study of three German states with differing low stakes exam regimes. *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas, 20*, 1-28.

Van Gasse, R., Vanhoof, J., & de Vos, W. (2014). Informatiegebruik in verantwoordingsgerichte en schoolontwikkelingsgerichte onderwijssystemen: een Nederlands-Vlaams perspectief. *Pedagogiek, 34*(2), 84-106.

Van Gasse, R., Vanhoof, J., Mahieu, P., & Van Petegem, P. (2015). *Informatiegebruik door schoolleiders en leerkrachten*. Antwerpen/Apeldoorn: Garant.

Van Petegem, P., & Vanhoof, J. (2002). *Evaluatie op de testbank. Een handboek voor het ontwikkelen.* Mechelen: Wolters Plantyn.

Van Petegem, P., & Vanhoof, J. (2007). Towards a model of effective school feedback: School heads' points of view. *Educational Research and Evaluation, 13*(4), 311-325.

Van Petegem, P., Vanhoof, J., Daems, F., & Mahieu, P. (2005). Publishing information on individual schools? *Educational Research and Evaluation, 11*(1), 45-60.

Vandenberghe, R. (1983). *Studying Change in Primary and Secondary Schools in Belgium and the Netherlands*.

Vanhoof, J., Mahieu, P., & Van Petegem, P. (2009). Geïnformeerde schoolontwikkeling: van een nieuw gegeven naar een beleidsinstrument. *Kwaliteitszorg in het onderwijs, 22*, 17-51.

Vanhoof, J., & Van Petegem, P. (2007). Matching internal and external evaluation in an era of accountability and school development: Lessons from a Flemish perspective. *Studies in Educational Evaluation, 33*(2), 101-119.

Vanhoof, J., & Van Petegem, P. (2017). *Doeltreffend schoolbeleid: praktijkboek beleidsvoerend vermogen in scholen*. Leuven/Den Haag: Acco.

Vanhoof, J., Vanlommel, K., Thijs, S., & Vanderlocht, H. (2014). Data use by Flemish school principals: impact of attitude, self-efficacy and external expectations. *Educational Studies, 40*(1), 48-62. doi:10.1080/03055698.2013.830245

Vanhoof, J., Verhaeghe, G., & Van Petegem, P. (2009). Verschilllen in het gebruik van schoolfeedback: een verkenning van verklaringsgronden. *Tijdschrift voor onderwijsrecht en onderwijsbeleid*(4), 306-322.

Vanhoof, J., Verhaeghe, G., Van Petegem, P., & Valcke, M. (2012). Flemish primary teachers' use of school performance feedback and the relationship with school characteristics. *Educational Research, 54*(4), 431-449. doi:10.1080/00131881.2012.734726

Vanlommel, K., Vanhoof, J., & Van Petegem, P. (2016). Data use by teachers: the impact of motivation, decision-making style, supportive relationships and reflective capacity. *Educational Studies, 42*(1), 36-53. doi:10.1080/03055698.2016.1148582

Vansteenkiste, M., Lens, W., & Deci, E. L. (2007). Intrinsic versus extrinsic goal contents in self-determination theory: Another look at the quality of academic motivation. *Educational Psychologist, 41*(1), 19-31.

Verger, A., & Parcerisa, L. (2018a). Test-based accountability and the rise of regulatory governance in education: A review of global drivers. *Education governance and social theory: Interdisciplinary approaches to research*, 139-158.

Verger, A., & Parcerisa, L. (2018b). Test-based accountability and the rise of regulatory governance in education: A review of global drivers. In A. Wilkins & A. Olmedo (Eds.), *Education governance and social theory: Interdisciplinary approaches to research* (pp. 139-158). London: Bloomsbury.

Verhaeghe, G., Vanhoof, J., Valcke, M., & Van Petegem, P. (2010a). Effecten van ondersteuning bij schoolfeedbackgebruik. *Pedagogische Studien, 88*, 90-106.

Verhaeghe, G., Vanhoof, J., Valcke, M., & Van Petegem, P. (2010b). Using school performance feedback: perceptions of primary school principals. *School Effectiveness and School Improvement, 21*(2), 167-188. doi:10.1080/09243450903396005

Verhaeghe, G., Vanhoof, J., Van Petegem, P., Verhaeghe, J. P., & Van Damme, J. (2010). *Het gebruik van outputgegevens in basisscholen: concretiseringen en illustraties uit het Schoolfeedbackproject*.

Visscher, A. (2002). School improvement through performance feedback. In A. J. Visscher & R. Coe (Eds.), *School Effectiveness and School Improvement*. Lisse: Swets & Zeitlinger.

Visscher, A., & Coe, R. (2003). School performance feedback systems: Conceptualisation, analysis, and reflection. *School Effectiveness and School Improvement, 14*(3), 321-349. doi:10.1076/sesi.14.3.321.15842

Visscher, A., & Ehren, M. (2011). *De eenvoud en complexiteit van Opbrengstgericht Werken*. Universiteit Twente, Vakgroep Onderwijsorganisatie en-management.

Wang, L., Beckett, G. H., & Brown, L. (2006). Controversies of standardized assessment in school accountability reform: A critical synthesis of multidisciplinary research evidence. *Applied measurement in education, 19*(4), 305-328.

Wayman, J. C., Wilkerson, S. B., Cho, V., Mandinach, E. B., & Supovitz, J. A. (2016). *Guide to Using the Teacher Data Use Survey (REL 2017-166)*. Paper presented at the American Educational Research Association, US, San Antonio.

Willingham, W. W., Pollack, J. M., & Lewis, C. (2002). Grades and test scores: Accounting for observed differences. *Journal of Educational Measurement, 39*(1), 1-37.

Winke, P. (2011). Evaluating the validity of a high-stakes ESL test: Why teachers' perceptions matter. *Tesol Quarterly, 45*(4), 628-660.

Woessmann, L. (2002). *How Central Exams Affect Educational Achievement: International Evidence from TIMSS and TIMSS-Repeat*. Retrieved from https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED468717&site=ehost-live

Woessmann, L. (2007). International evidence on school competition, autonomy, and accountability: A review. *Peabody Journal of Education, 82*(2-3), 473-497.

Wohlstetter, P., Datnow, A., & Park, V. (2008). Creating a system for data-driven decision-making: Applying the principal-agent framework. *School Effectiveness and School Improvement, 19*(3), 239-259.

Woolfolk, A. (2008). *Educational Psychology. Active Learning Edition* (2nd ed.). Boston, MA: Allyn & Bacon.

Young, V. M. (2006). Teachers' use of data: Loose coupling, agenda setting, and team norms. *American journal of Education, 112*(4), 521-548.