Zelfevaluatie binnenstebuiten

Onderzoek naar zelfevaluaties in scholen

Zelfevaluatie binnenstebuiten

Onderzoek naar zelfevaluaties in scholen

Jan Vanhoof

Universiteit Antwerpen

Wolters Plantyn

Wolters Plantyn
een Wolters Kluwer-bedrijf

Plantyn ontwikkelt en verspreidt leermiddelen voor alle onderwijsniveaus. Daarnaast geeft Plantyn ook publicaties uit over schoolmanagement, leerlingenbegeleiding, personeelsbeleid voor het onderwijs en didactische ondersteuning van leerkrachten en educatief materiaal voor thuis.

Plantyn NV
Adres: Motstraat 32, 2800 Mechelen
Telefoon: 015 36 36 36
Fax: 015 36 36 37
E-mail: klantendienst@woltersplantyn.be
Website: www.woltersplantyn.be

Ontwerp omslag: Pieter Lesage

Omslagillustratie: Peter Van Hoof

NUR 841

© Plantyn NV, Mechelen, België

Alle rechten voorbehouden. Behoudens de uitdrukkelijk bij wet bepaalde uitzonderingen mag niets uit deze uitgave worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand of openbaar gemaakt, op welke wijze dan ook, zonder de uitdrukkelijke voorafgaande en schriftelijke toestemming van de uitgever.

ISBN 978-90-301-9567-2 D/2007/0032/1120

't leren is mooi

WOORD VOORAF

Zelfevaluatie is 'in'. Geen schoolleider of beleidsmaker die er de mond niet van vol heeft. De verwachtingen zijn groot, maar ook terecht? In dit doctoraatsonderzoek ging ik na hoe kwaliteitsvol zelfevaluaties in scholen verlopen en waaraan verschillen in die kwaliteit kunnen toegeschreven worden. Ik stel u hierbij graag het resultaat van dit onderzoek voor.

Reeds sinds mijn studentenjaren zijn er twee zaken die mij bijzonder boeien: onderwijs en kwaliteitszorg. Ik beschouw het dan ook als een voorrecht dat ik de voorbije jaren tijd mocht besteden aan onderzoek naar kwaliteitszorg in het onderwijs. Deze ervaring doet me – wederom - beseffen dat ik een zondagskind ben. Het is een privilege om de kans en tijd te krijgen om je professioneel in je eigen interesses te kunnen verdiepen. Zeker wanneer dat kan in de omstandigheden waarin ik mocht vertoeven. De voorbije periode is voor mij in veel opzichten dan ook erg waardevol geweest.

Het moet gezegd: zonder de steun die ik mocht ervaren, had dit proefschrift niet de huidige vorm aan kunnen nemen. Peter, het overleg met jou als promotor was leerzaam, inspirerend, en motiverend. De vrijheid en het vertrouwen dat je me de voorbije jaren gaf, bleken voor mij een prima context om in te groeien. Ik hoop dat dit werk in dezelfde mate een leven mag gaan leiden als de andere zaken die we voorheen samen gedaan hebben. Sven, jij bent het levende bewijs dat 'statistics can be fun'. Jouw suggesties hebben de kwaliteit van dit werk sterk opgekrikt (en ik spreek bewust niet alleen over de gebruikte kwantitatieve analyses). Naast Peter Van Petegem en Sven De Maeyer maakten ook Geert Devos en Eric Verbiest deel uit van de Begeleidingscommissie van dit proefschrift. Ik heb onder meer hun kritische inbreng en deskundigheid sterk geapprecieerd. Ook mijn naaste collega's van het Instituut voor Onderwijs- en Informatiewetenschappen (IOIW) en in het bijzonder van de onderzoeksgroep EduBROn ben ik zeer dankbaar, voor alle hulp en de gezellige tijd die we samen hadden. Alexia, jou moet ik zeker ook met naam noemen. Op 1001 momenten vroeg en kreeg ik jouw advies en hulp.

Ook de vele personen die doorgaans niet op de voorgrond treden wil ik bedanken. Daarmee bedoel ik de meer dan 3000 respondenten die bereid waren de vragenlijsten in te vullen, mij te woord te staan of deel te nemen aan de gevalsstudies. Ondanks de soms wat minder positieve teneur doorheen het beschrijven van de onderzoeksresultaten, wil ik benadrukken dat mij van de vele contacten in het veld voornamelijk de warme menselijkheid, de gedrevenheid en de daadkracht zijn bijgebleven.

Mijn fierheid over het voorliggende resultaat, wil ik graag met mijn familie delen. In de eerste plaats met mijn ouders. Niet in het minst wil ik uiteraard ook Ann bedankten, voor haar solidariteit en ruggesteun. Bedankt voor de vele niet-vanzelfsprekende vanzelfsprekendheden. Neel en To, het laatste jaar waren jullie bij wijze van spreken *zondags*kinde-

ren. Nu ik in alle rust op het hele doctoraatsproces kan terugkijken, weet ik welke lessen ik er uit kan trekken. Ik draag dit proefschrift daarom graag aan jullie op.

Jan Vanhoof 10 oktober 2007

INHOUDSOPGAVE

Deel 1: Probleemstelling en onderzoeksvragen **11**

1 Het ruimere kader: zelfevaluatie in de hedendaagse onderwijscontext 11
1.1 (Onderwijs)kwaliteit *12*
1.2 De rol van zelfevaluatie in een kwaliteitszorgsysteem *13*
1.3 Zelfevaluaties in het hedendaagse Vlaamse onderwijs *19*

2 Vanuit de nood aan empirische evidentie naar meta-evaluatie 21

3 Zelfevaluatie binnenstebuiten: Focus, orienterende onderzoeksvragen en doelen van het onderzoek 21
3.1 Focus van deze studie: de 'geïnstrumentaliseerde zelfevaluatie' *22*
3.2 Oriënterende onderzoeksvragen en doelen van deze studie *23*

4 Afrondend 23

Deel 2: Multimethode onderzoek **25**

1 Waarom een multimethode onderzoeksdesign? 26

2 Het verloop van het multimethode onderzoek 27

3 Afrondend 29

Deel 3: Voorbereidende studies **31**

Invalshoek 1: Een delphi-studie naar de visie van Vlaamse onderwijsexperten

1 Waarom een delphi-studie? 32

2 Opzet van de Delphi-studie 33
2.1 Het verloop van de Delphi-studie *33*
2.2 De keuze van informanten en respons *34*
2.3 De inhoudelijke focus van beide Delphi-rondes *34*
2.4 Analyse van de gegevensbronnen *35*

3 Resultaten: (Komen tot) Kwaliteitsvolle zelfevaluaties volgens Vlaamse deskundigen 36
3.1 Het omschrijven van (kwaliteitsvolle) zelfevaluatie *36*

3.2 De voorspellers van kwaliteitsvolle zelfevaluatie *40*
3.3 Stand van zaken inzake zelfevaluatie in Vlaamse scholen *50*
3.4 (De doeltreffendheid van de) beschikbare ondersteuning *55*

4 Algemene conclusies 57
4.1 Indicatoren voor kwaliteitsvolle zelfevaluatie *57*
4.2 Hoe kwaliteitsvol zijn zelfevaluatie in Vlaamse scholen? *58*
4.3 Hoe kunnen de verschillen in kwaliteit verklaard worden? *59*

5 Afrondend 60

Invalshoek 2: Literatuurstudie 60

1 Het proces: de zoektocht naar en selectie van relevante bronnen 61

2 Het resultaat: Een variatie aan geraadpleegde bronnen 62

3 Afrondend: Verfijnde onderzoeksvragen 62

Deel 4: Ontwerp en operationalisatie van het theoretisch kader 65

1 Meta-evaluatie gecontextualiseerd: focus op schoolgebonden kenmerken 67
1.1 De organisatie-effectiviteit van scholen *67*
1.2 De school als lerende organisatie *79*

2 Meta-evaluatie als dusdanig: Focus op zelfevaluatiegebonden kenmerken 91
2.1 Een profiel om zelfevaluaties in scholen te beschrijven *91*
2.2 De 'kwaliteit' van een zelfevaluatie *93*
2.3 Voorspellers van de kwaliteit van zelfevaluaties *95*

3 Omtrent de ondersteuning 112
3.1 Het beschrijven van de ondersteuning *112*
3.2 Het beoordelen van de ondersteuning *112*

4 Afrondend: Te toetsen onderzoekshypothesen 114

Deel 5: Voorbereiding van de kwantitatieve analyses 119

1 (Waarom) verschillende vormen van kwantitatieve analyses 119

2 Toelichting bij de beschikbare data 120
2.1 De GOK-zelfevaluaties *120*
2.2 De GOK-inspectie *121*
2.3 Zelfevaluatie binnenstebuiten *123*
2.4 Afrondend: een vergelijking van de 3 gegevensverzamelingen *130*

3 Voorbereidende analyses 131

3.1 Schaalconstructie *131*
3.2 Clusteranalyses *149*

4 Afrondend 156

Deel 6: Zelfevaluaties in de bestudeerde scholen – een beschrijving 157

1 Initiatief, functies en verwachtingen 157

2 Inhoud en instrumentkeuze 159

3 Betrokkenheid bij de zelfevaluatie 161

4 Ondersteuning 164

5 De kernkenmerken van zelfevaluaties 165

6 Afrondend 167

Deel 7: Zelfevaluaties in de bestudeerde scholen – via hypothesetoetsing naar modelvorming 169

1 Toetsen van de onderzoekshypothesen 169
1.1 De houding ten aanzien van zelfevaluatie *169*
1.2 Het verloop van zelfevaluaties: zelfevaluatie als beleidsdaad *175*
1.3 Het verloop van zelfevaluaties: zelfevaluatie als onderzoeksdaad *186*
1.4 Relaties tussen de zelfevaluatiegebonden kenmerken *192*
1.5 De kwaliteit van zelfevaluaties *194*
1.6 De beoordeling en de impact van ondersteuning *212*

2 Een passend verklaringsmodel voor de kwaliteit van zelfevaluaties 215
2.1 Het volledige verklaringsmodel *215*
2.2 Het meest spaarzame model *219*
2.3 Conclusies bij de vooropgestelde onderzoekshypothesen *223*

3 Algemene conclusies en discussie 224
3.1 Aan de hand van welke indicatoren kan men de kwaliteit van een zelfevaluatie operationaliseren? *224*
3.2 Hoe kwaliteitsvol zijn zelfevaluaties in de bestudeerde scholen? *232*
3.3 Hoe kunnen de verschillen in het verloop van zelfevaluaties verklaard worden? *237*
3.4 Hoe kunnen de verschillen in de resultaten van zelfevaluaties verklaard worden? *239*

4 Afrondend 241

Deel 8: Gevalstudies **243**

1 Waarom gevalstudies? 243

2 Het verloop van de gevalstudies 244
2.1 De selectie van casussen *245*
2.2 De natuurlijke omgeving *245*
2.3 De onderzochte periode *246*
2.4 De databronnen *246*
2.5 De onderzochte variabelen *247*
2.6 Analyse *247*

3 Bevindingen 249

4 Verticale analyse 250
4.1 Uitgebreide beschrijving van de zelfevaluatie in het Technisch Instituut Voldegem *251*
4.2 Synthese van de zelfevaluatie in het Technisch Instituut Baldegem en het Sint-Odrada Instituut *266*

5 Horizontale analyse 271
5.1 Zelfevaluatie als beleidsdaad *271*
5.2 Zelfevaluatie als onderzoeksdaad *299*
5.3 Resultaten van de zelfevaluatie *310*

6 Conclusies en theoretische reflecties 316
6.1 Een antwoord op de onderzoeksvragen *317*
6.2 Theoretische reflecties naar aanleiding van de gevalstudies *335*

7 Afrondend 336

Deel 9: Synthese en eindbeschouwingen **339**

1 Achtergrond, onderzoeksvragen en methodologie 339

2 Empirische evidentie: antwoorden op de onderzoeksvragen 340

3 Voorbij de empirische evidentie: (beleids)aanbevelingen naar aanleiding van deze studie 346

4 Reflecties bij het onderzoek en suggesties voor toekomstig onderzoek 354

5 Afrondend 357

Literatuur **359**

DEEL 1

PROBLEEMSTELLING EN ONDERZOEKSVRAGEN

Van scholen wordt (steeds meer) verwacht dat ze mee verantwoordelijkheid dragen voor het ontwikkelen en garanderen van onderwijskwaliteit. Ze dienen aan zelfevaluatie te doen. Dat houdt in dat zij zicht dienen te krijgen op hun huidige functioneren (sterktes en zwaktes) als vertrekpunt voor een plan of visie voor de toekomst (Devos, Verhoeven, Van den Broeck, Gheysen, Opbrouck & Verbeeck, 1999, p. 26). Zelfevaluatie is een door de school geïnitieerde en uitgevoerde procedure om het eigen functioneren te beschrijven en te beoordelen (Blok, Sleegers & Karsten, 2005, p.3). Vormen van zelfevaluatie zijn het voorbije decennium in vele Vlaamse scholen gemeengoed geworden of aan het worden. Onderwijskoepels, universiteiten, steunpunten en scholen zelf ontwikkelden instrumenten en methodieken die scholen in staat stellen om zichzelf te evalueren. Deze studie handelt over het gebruik van dergelijke zelfevaluatie-instrumenten.

Een analyse van literatuur inzake zelfevaluaties leert dat de verwachte resultaten in sterke mate positief zijn, terwijl de evidentie voor dit positieve beeld vooralsnog beperkt is. Ondanks de vaststelling dat velen het bestaan van zelfevaluaties toejuichen, kunnen – zoals zal blijken - bij de kwaliteit van deze zelfevaluaties op dit moment vraagtekens geplaatst worden. Daardoor dringt zich de vraag op in welke mate zelfevaluaties momenteel kwaliteitsvol verlopen en hoe verschillen in de kwaliteit van zelfevaluaties verklaard kunnen worden.

Een goed begrip van deze vragen vergt inzicht in de functies die zelfevaluaties momenteel in het Vlaamse onderwijs vervullen. We starten dit eerste deel daarom met het schetsen van een ruimer kader waarin de focus van het onderzoek gesitueerd wordt in de actuele onderwijscontext (zie paragraaf 1). Dit kader brengt ons bij de noodzaak van onderzoek naar de kwaliteit van zelfevaluaties (zie paragraaf 2). Aan die noodzaak wil dit proefschrift tegemoet komen. De derde paragraaf beschrijft de uitgangspunten die we daarbij hanteerden. Er is aandacht voor de onderzoeksvragen en de doelen van deze studie.

1 HET RUIMERE KADER: ZELFEVALUATIE IN DE HEDENDAAGSE ONDERWIJSCONTEXT

Twee recente beleidsinitiatieven illustreren dat scholen zelfstandig geacht worden de professionaliteit aan de dag te leggen om zichzelf te evalueren.

> In het Gelijke Onderwijskansendecreet (X, 2002) werd gestipuleerd dat scholen vanuit een analyse van hun beginsituatie moesten aangeven op welke manier zij hun schoolspecifieke invulling van het gelijke onderwijskansenbeleid (GOK-beleid) zouden evalueren. Daarnaast werd van scholen verwacht dat ze aspecten van het eigen schoolfunctioneren kritisch onder de loep namen, beschreven en beoordeelden. Dat gebeurde onder meer in de vorm van een zelfevaluatie. De manier waarop scholen daarbij dienden te werk gaan werd door de overheid niet a priori geëxpliciteerd. Die verantwoordelijkheid werd aan de scholen zelf toegekend.
>
> Daarnaast is er de trend om de combinatie interne evaluatie (de zelfevaluatie door scholen) en externe evaluatie (de doorlichting) te benadrukken. Deze beleidsintentie van de onderwijsinspectie vindt haar oorsprong in de nadruk die gelegd wordt op de het ter verantwoording roepen van scholen en op het realiseren van schoolinterne kwaliteitsontwikkeling door middel van het uitvoeren van zelfevaluaties.

Om deze illustraties en het toenemende belang van zelfevaluaties te doorgronden kaderen we het geheel in de actuele onderwijscontext. Daartoe besteden we eerst aandacht aan de concepten '(onderwijs)kwaliteit' en 'kwaliteitszorg' en aan de rol die zelfevaluatie daar vandaag de dag in speelt.

1.1 (Onderwijs)kwaliteit

(Onderwijs)kwaliteit, daar draait het allemaal om. Daarmee is meteen alles en niets gezegd. Het begrip is tegelijkertijd allesomvattend en leeg. Het is allesomvattend in die zin dat het al het goede omvat omtrent waartoe onderwijs zou moeten bijdragen en omtrent hoe onderwijs er zou moeten uitzien. Het begrip 'kwaliteit' is daardoor uitgegroeid tot een (blijvend) modewoord dat door iedereen als een nastrevenswaardig iets aanzien wordt. Dat maakt het precies tot een leeg begrip. De gezamenlijke aanhang van 'onderwijskwaliteit' brokkelt immers meteen in sterke mate af wanneer concretere invullingen gegeven worden aan het begrip. Wanneer het begrip 'onderwijskwaliteit' geoperationaliseerd wordt, blijkt het plots subjectief, evolutief en zelfs manipulatief (Van Petegem, 2005, p. 12-13).

Het bovenstaande impliceert dat we ons niet op de oppervlakte mogen houden omtrent de invulling die 'onderwijskwaliteit' in deze studie zal krijgen. (Onderzoek naar) kwaliteitszorg impliceert immers een notie over wat kwaliteit inhoudt (Standaert, 1999). Ons vertrekpunt is dat *(onderwijs)kwaliteit betekent: het op een gepaste manier realiseren van de door belanghebbenden gedeelde verwachtingen* (Harrington & Harrington, 1994). Vragen naar kwaliteit (van zelfevaluaties) zullen in deze studie vanuit die invalshoek benaderd worden.

De (gedeelde) verwachtingen zijn echter velerlei van aard (Hendriks, Doolaard & Bosker, 2002) en met het oog op de invulling van kwaliteitszorg is het van belang een duidelijk

onderscheid te maken naar de herkomst ervan. Wij onderscheiden schoolinterne en schoolexterne verwachtingen. Binnen de groep van schoolexterne verwachtingen maken we verder een onderscheid tussen decretaal verankerde verwachtingen en niet decretaal verankerde verwachtingen. Decretaal verankerde verwachtingen zijn die verwachtingen die de overheid (in naam van de samenleving) aan scholen oplegt en die afgedwongen kunnen worden (het realiseren van eindtermen, leerlingenparticipatie, het in acht nemen van de minimumlessenroosters, toepassen van de taalwetgeving). Wanneer scholen niet aan deze verwachtingen voldoen kan en moet de overheid sanctionerend optreden. Niet decretaal verankerde schoolexterne verwachtingen zijn afkomstig van belanghebbenden wiens verwachtingen niet in een juridisch kader zijn vastgelegd. Deze groep belanghebbenden is zeer uitgebreid en groeiende (bvb. scholen waar leerlingen naar doorstromen, afnemers van gediplomeerde leerlingen, ouders (en hun verenigingen), pastorale diensten en allerlei (belangen)organisaties zoals Sensoa of de Fietsersbond)). Deze belanghebbenden doen weliswaar een beroep op scholen, bepalen mee de invulling van onderwijskwaliteit, maar kunnen niet sanctionerend optreden zoals een overheid dat kan. Tenslotte zijn er de schoolinterne verwachtingen. Hier gaat het om eigen accenten die scholen - bovenop de minimumeisen die vanuit de samenleving verwacht worden – naar voren schuiven (Van Petegem, 1999). Deze verwachtingen kunnen bijvoorbeeld van levensbeschouwelijke aard zijn of kunnen betrekking hebben op de te hanteren didactische methoden. Er zijn dus verschillende kwaliteiten. Elke type verwachtingen vergt een eigen invulling van kwaliteitszorg en van de rol van zelfevaluatie in dat geheel.

1.2 De rol van zelfevaluatie in een kwaliteitszorgsysteem

Kwaliteitszorg is een overkoepelend begrip dat alle initiatieven en activiteiten omvat die ondernomen worden om de kwaliteit van scholen te onderzoeken, te bewaken, te verbeteren, te borgen en eventueel zelfs openbaar te maken. Op basis van de vraag wie de verantwoordelijkheid draagt voor de kwaliteitszorg wordt een onderscheid gemaakt tussen interne en externe kwaliteitszorg (Scriven, 1991; Blanquaert, Maes & Ver Eecke, 2001; Nevo, 2001). Er is sprake van interne kwaliteitszorg wanneer de activiteiten worden ondernomen op initiatief van en door de school zelf. Met *interne kwaliteitszorg* wordt dus de bewaking, ontwikkeling en verbetering van de onderwijskwaliteit binnen de schoolorganisatie bedoeld. Bij *externe kwaliteitszorg* ligt het initiatief voor de kwaliteitszorgactiviteiten en de uitvoering ervan bij personen of instellingen buiten de school (bvb. de onderwijsinspectie of een accrediteringsinstituut) (Nevo, 2001). Externe evaluaties door de inspectie richten zich doorgaans vooral op het beleid, de regelgeving en het onderwijsrendement (c.q. de decretale verwachtingen). Interne evaluatie kan zich in principe toeleggen op gelijk welk thema dat de school de moeite waard vindt (c.q. de drie groepen verwachtingen).

De invulling van kwaliteitszorg is de voorbije decennia sterk geëvolueerd. Traditioneel lag de verantwoordelijkheid zeer sterk bij de overheid. Dat leidde tot een omvangrijk en complex geheel van regelgeving dat aan scholen werd opgelegd en tot een sterke nadruk op externe evaluatie. Hoewel de noodzaak daarvan algemeen aanvaard was, werd de

concrete uitwerking ervan steeds sterk bekritiseerd (Norton Grubb, 1999; Nevo, 2001). In de jaren '90 is men gaan inzien dat de sterke regulering een belemmering is gaan vormen voor de beweegruimte van scholen (Güthe, 1997). Als reactie daarop is men scholen ruimte gaan geven voor het voeren van een eigen schoolbeleid. Ook de herkomst van de aandacht voor zelfevaluatie moet gezocht worden in de groeiende autonomie die scholen in Vlaanderen en in de ons omringende landen genieten. Deregulering en decentralisatie maken ook in Vlaanderen voortdurend deel uit van het gevoerde onderwijsbeleid (Vanderpoorten, 2000; Vandenbroucke, 2004). Scholen worden – tenminste afgaande op de beleidsretoriek - steeds meer autonoom en komen tot een verhoogde mate van zelfsturing. Ook de invulling van kwaliteitszorg is in het Vlaamse onderwijssysteem bijgevolg in ontwikkeling. Het hedendaagse onderwijsbeleid is gekenmerkt 'door een bijna driftig zoeken naar nieuwe beleidsinstrumenten' (Verstegen, 2005, p. 137).

Grofweg zijn er binnen kwaliteitszorg momenteel twee belangrijke perspectieven te onderscheiden: het 'verantwoordingsperspectief' en het 'schoolontwikkelingsperspectief' (Nisbet, 1988; Cheng, 1996). Het onderscheid tussen beide perspectieven berust op verschillende antwoorden op de vraag of (1) het bij kwaliteitszorg in de eerste plaats gaat om bewaking en verantwoording dan wel om ontwikkeling en verbetering en (2) op de vraag wie de 'kwaliteit van onderwijs' omschrijft, c.q. de overheid dan wel de school zelf. De verschillende antwoorden van beide perspectieven worden in Tabel 1 schematisch samengebracht (De Wolff, 2003, p. 18).

Tabel 1: De invulling van verantwoording en schoolontwikkeling

Verantwoordingsperspectief	Schoolontwikkelingsperspectief
Kwaliteit van onderwijs	
• bereiken van vooraf bepaalde, meetbare doelen • meetbaar (kwantificeerbaar) • landelijk/wettelijk vast te stellen	• gedeelde visie op gewenste kwaliteit: commitment bij ieder om dat waar te maken • niet volledig te vatten; zorg voor een zo rijk mogelijke weergave • door de school zelf te bepalen
Kwaliteitszorg	
• uniformiteit • behalen van minimumdoelen • bedrijfsvoering: dmv model en procedures: heldere processen, afstemming • meten en vergelijken (t.o.v. normen) • externe sturing • externe evaluatie (bvb. schooldoorlichting)	• inzetten van verschillen • behalen van het (zelfgekozen) maximale • voortgaande dialoog met betrokkenen over gewenste en gerealiseerde kwaliteit • evaluatiegegevens als input voor dialoog • zelfsturing • interne evaluatie (bvb. zelfevaluatie)

Binnen het *verantwoordingsperspectief* staat de beheersbaarheid centraal. Van scholen wordt verwacht dat ze vooraf bepaalde doelen realiseren. Die doelen worden door de overheid (c.q. de samenleving) opgelegd en zijn in een wetgevend kader vastgelegd. Merk op dat die doelen niet noodzakelijk enkel betrekking hebben op de output van scholen, ook omtrent processen kan de overheid verwachtingen vooropstellen (bvb. de aandacht

voor het participatiedenken). Kwaliteitszorg is binnen dit perspectief extern gestuurd en gericht op uniformiteit, het behalen van minimumdoelen en een efficiënte bedrijfsvoering. Het vindt in eerste instantie plaats op initiatief van schoolexterne actoren. Bij het *schoolontwikkelingsperspectief* (of schoolverbeteringsperspectief) daarentegen gaat men uit van een gedeelde visie op de gewenste kwaliteit; het gaat om door de school zelf te bepalen doelen omtrent de output en processen. Kwaliteitszorg richt zich hier op de verschillen tussen scholen, leerkrachten en leerlingen en wil een dialoog op gang brengen. Die dialoog moet duidelijkheid scheppen over waar men met de school naartoe wil en moet de leden van de schoolgemeenschap motiveren om de vooropgestelde doelen ook effectief te gaan realiseren. Kwaliteitszorg is er in het schoolontwikkelingsperspectief in eerste instantie op initiatief van schoolinterne actoren en factoren (Clift, Nuttall & McCormick, 1989).

Het verantwoordingsperspectief was tot voor kort het exclusieve terrein van externe evaluatie en het schoolontwikkelingsperspectief dat van interne evaluatie. De tweedeling is evenwel aan het vervagen. Zelfevaluatie speelt momenteel een rol in de twee perspectieven van kwaliteitszorg. We duiden de relevantie van dit onderzoek dan ook vanuit beide perspectieven.

1.2.1 Zelfevaluatie als schoolontwikkelingsinstrument

Sociaal-culturele, economische en technologische ontwikkelingen maken dat scholen organisaties zijn die voortdurend veranderen. De werk- en leersituaties in scholen zijn dan ook in volle ontwikkeling, of worden althans geacht in volle ontwikkeling te zijn om tegemoet te komen aan nieuwe inzichten en uitdagingen. De tendenzen van deregulering en autonomievergroting maken dat van scholen verwacht wordt dat ze meer vorm gaan geven aan hun eigen vermogen tot veranderen (Scheerens, 1996). Dat brengt ons bij het concept 'schoolontwikkeling'. Op basis van het werk van Lagerweij (Lagerweij & Voogt, 1997; Lagerweij, Lagerweij & Voogt, 2004) omschreven wij schoolontwikkeling voorheen als *'het voortdurend proces van het toevallig, autonoom en/of bewust veranderen van het organisatorische en onderwijskundige functioneren van de school met als doel het verbeteren van de onderwijskwaliteit'* (Van Petegem & Vanhoof, 2004). Deze omschrijving impliceert dat het wellicht onmogelijk is een algemene consensus te bereiken over het concrete doel van schoolontwikkeling. Uit onze invulling van (onderwijs)kwaliteit volgt immers dat het doel van schoolontwikkeling een gecontextualiseerde en dus schooleigen invulling vergt. Zelfevaluaties in het kader van schoolontwikkeling kunnen zich dan ook op gelijk welke verwachting richten (intern en - al dan niet decretaal verankerd - extern).

Merk op dat er in de beschrijving van schoolontwikkeling zowel sprake is van toevallig als van autonoom en bewust veranderen. Vanuit de chaostheorie wordt weliswaar gewezen op het belang van toevalligheden bij veranderingsprocessen (Zuijderhoudt, Wobben & Ten Have, 2002). Gegeven hun maatschappelijke rol dienen scholen van schoolontwikkeling echter ook een systematisch proces te maken. Dat betekent dat met een gezamenlijk doel voor ogen een strategisch plan wordt uitgewerkt om het klas- en schoolbeleid te

verbeteren (Hulpia, 2004). Het essentiële onderdeel van dat proces bestaat uit een kritische reflectie op het eigen functioneren - doorgaans in de vorm van een algemene of partiële zelfevaluatie - die resulteert in concrete handelingsplannen (Van Petegem, 1998). Zelfevaluatie is in dat perspectief een gebeuren dat schoolintern een (langzame) dynamiek van groei en ontwikkeling, bijsturing en vernieuwing moet helpen grondvesten (Fullan, 1991; Gombeir, 1998). Ook Van Petegem stelt zelfevaluatie in zijn definitie in het teken van schoolontwikkeling. Zelfevaluatie is volgens deze definitie 'een proces, in hoofdzaak geïnitieerd door de school zelf, waarbij welgekozen participanten op een systematische wijze het functioneren van de school beschrijven en beoordelen met het oog op het nemen van beslissingen c.q. initiatieven *in het kader van (aspecten van) de algehele school(beleids)-ontwikkeling'* (Van Petegem, 1998, p. 104). Schoolontwikkeling vergt dan ook een gewaarborgde kwaliteit van de uitgevoerde zelfevaluaties.

1.2.2 Zelfevaluatie als verantwoordingsinstrument

Uit een Delphi-studie naar de opinies van Vlaamse onderwijsexperts komen verschillende motieven naar voor die scholen nopen zich te verantwoorden voor de keuzes die ze maken (Vanhoof, Van Petegem, Mahieu & Daems, 2004). Zij dienen onder meer rekenschap af te leggen voor het besteden van de gemeenschapsgelden. De overheid, en dus de burger, heeft het recht om te eisen dat haar middelen op een efficiënte manier worden besteed. Daarenboven maken ook de verantwoordelijkheid die scholen dragen voor jongeren en de invloed die ze hebben op de maatschappelijke integratie van jongeren dat scholen verantwoording dienen af te leggen. Scholen hebben een mandaat van de overheid om op basis van kwaliteitsvol onderwijs diploma's en getuigschriften uit te reiken. Scholen moeten volgens Vlaamse onderwijsexperts dan ook kunnen aantonen dat hun onderwijs inderdaad voldoende kwaliteit heeft zodat de waarde van de uitgereikte getuigschriften en diploma's verzekerd wordt. In de beleidsbrief van Vandenbroucke *(2004, p. 48)* lezen we hieromtrent: '*In de mate dat scholen meer mogelijkheden tot zelfsturing krijgen, mag van die scholen ook meer worden verwacht dat ze zich verantwoorden voor de beslissingen die ze nemen'*.

Zelfevaluatie kan binnen dit verantwoordingsperspectief op twee manieren een rol spelen: op zichzelf staand en in combinatie met externe evaluatie. Hieronder beschrijven we welke rol zelfevaluatie in beide benaderingen zou kunnen vervullen. We spreken ons op deze plaats evenwel niet uit over de wenselijkheid van mogelijke scenario's. Hoe en aan wie scholen verantwoording dienen af te leggen zijn vragen die in Vlaanderen immers nog moeten uitgeklaard worden. Deze maatschappelijke discussie moet nog op kruissnelheid komen. Dat maakt het bijzonder moeilijk om nu al vast te leggen welke rol zelfevaluatie kan krijgen in dit geheel. Zelfevaluaties zijn in deze invalshoek een 'middel tot' dat maar kan ingezet worden wanneer er sprake is van een duidelijk doel.

a. VERANTWOORDING OP BASIS VAN ZELFEVALUATIES ALS DUSDANIG

In principe zou men kunnen stellen dat een zelfevaluatie die valide en betrouwbare resultaten oplevert de basis kan zijn voor de gemeenschapsinspectie om de decretaal vastge-

legde verwachtingen ten aanzien van scholen te bewaken. Als je staat kan maken op de resultaten, klinkt het aannemelijk dat die resultaten de grond kunnen zijn voor beslissingen omtrent het al dan niet voldoen aan de voorwaarden voor subsidiëring. Dat is het principe van meta-evaluatie. Eenzelfde logica geldt voor de andere verwachtingen die het concept 'onderwijskwaliteit' met zich meebrengt. Over externe verwachtingen die niet decretaal verankerd zijn, hoeven scholen in principe geen verantwoording af te leggen aan de inspecterende overheid (Michielsens, 1999). Ze doen dit (on)rechtsreeks aan de instanties en (groepen) personen die deze verwachtingen vooropstellen. Hetzelfde geldt voor schoolinterne verwachtingen. Het komt er in de verschillende domeinen van kwaliteitszorg op aan een communicatie op gang te brengen tussen de school en de betrokken personen, groepen en instanties. De mate waarin scholen tegemoet komen aan verschillende verwachtingen kan in deze logica ook door scholen zelf onderworpen worden aan een evaluatie. De resultaten van dit zelfevaluatieproces kunnen dan de input zijn voor deze communicatie.

Het bovenstaande veronderstelt uiteraard een voldoende kwaliteit van de uitgevoerde zelfevaluaties. Die veronderstelling wordt zowel omwille van (school)beleid-strategische overwegingen als omwille van de capaciteiten van scholen om zelfevaluaties uit te voeren niet algemeen ondersteund (Rudd & Davies, 2000; Van Petegem, Verhoeven, Buvens & Vanhoof, 2005b). Wanneer scholen zich bewust zijn van het verantwoordingskarakter van een (zelf)evaluatie treden er immers neveneffecten op die meteen de kans op het verzamelen van valide en betrouwbare gegevens in de kiem smoren. Zo zal er in scholen die een negatieve beoordeling vrezen bijvoorbeeld makkelijk sprake zijn van schijngedrag, windowdressing en mooipraterij (Scheerens, Van Amelsfoort & Donoughue, 1999). De kans op artefacten is er niet denkbeeldig (Watling & Arlow, 2002; Maes, Ver Eecke & Zaman, 2004, p. 220). Van Petegem (1999, p. 30) spreekt van een strategisch gebruik van de zelfevaluatie. Het doel van scholen in een verantwoordingsperspectief is immers in de eerste plaats een zo goed mogelijke beurt te maken. Kyriakides & Campbell (2004) bestempelen geldigheid als het allergrootste probleem van zelfevaluaties. Bovendien worden momenteel ook vraagtekens geplaatst bij de kwaliteit van zelfevaluaties in (Vlaamse) scholen (Forss, Cracknell & Samset, 1994; Cousins & Earl, 1995; MVG, 2006). De rol van zelfevaluaties in het verantwoordingsperspectief wordt dan ook eerder gezien in combinatie met een externe evaluatie.

b. VERANTWOORDING OP BASIS VAN ZELFEVALUATIES ALS COMPLEMENT VAN EEN EXTERNE EVALUATIE

De relatie tussen interne en externe evaluatie van onderwijsinstellingen is al sinds lang een veel besproken thema in het denken over kwaliteitszorg in het onderwijs. Externe evaluaties in de vorm van schoolinspecties lijken immers altijd al een enigszins eigenaardig bestaan te hebben gekend; hoewel iedereen hun noodzaak onderschrijft, is de invulling ervan (niet in het minst door het onderwijsveld zelf) steeds sterk bekritiseerd geworden. In vele landen werden om die reden evaluatiemethoden ontwikkeld die een meer participatieve en zelfgestuurde vorm van evaluatie mogelijk maakten (Nevo, 2001). In het afleggen van verantwoording worden zelfevaluaties – in Vlaanderen maar ook in andere landen - dan als complement van een externe evaluatie aangezien (Scheerens e.a., 1999).

In de algemene conclusies van het project 'zelfevaluatie' (Van Hoyweghen, 2002, p. 16) lezen we bijvoorbeeld dat '*de zelfevaluatieactiviteiten van scholen een goede basis blijken te kunnen vormen voor externe schooldoorlichtingen. In het toezicht op kwaliteit kunnen interne en externe evaluatie complementair zijn*'.

Er worden drie rollen onderscheiden die het wenselijk zouden kunnen maken om aan zelfevaluatie een rol toe te kennen binnen het verantwoordingsperspectief (Nevo, 2001). Zelfevaluatie heeft ten aanzien van externe evaluatie een verruimende rol, een interpretatiebevorderende rol en een implementerende rol. Zelfevaluaties kunnen in eerste instantie een verruimende rol spelen (Maes e.a., 2004). Een externe evaluatie blijkt immers met meer diepgang en kwaliteitsvoller te kunnen gebeuren wanneer de zelfevaluatie van een school in de doorlichting betrokken wordt (Van Hoyweghen, 2002). De resultaten van een zelfevaluatie kunnen bovendien een bevorderende rol spelen in het interpreteren van gegevens die tijdens een doorlichting verzameld werden. In het verantwoordingsperspectief wordt doorgaans immers uitgegaan van een criterium- en normgericht referentiekader om de kwaliteit van scholen te beoordelen. Dat zijn op zich belangrijke en legitieme invalshoeken maar ze hebben als nadeel dat ze het lokale perspectief uit het oog verliezen. Een zelfevaluatie kan dat perspectief aan een externe evaluatie toevoegen (Christie, Ross & Klein, 2004). Tenslotte is er de implementerende rol. Scholen met ervaring inzake zelfevaluatie hebben een grotere kans om een constructieve houding aan te nemen ten opzichte van (school)doorlichtingen en om op een zinvolle manier om te gaan met de resultaten van een externe evaluatie (Peeters, Van Cauwenberghe & Schollaert, 1996; Nevo, 2002). Bovendien roept een interne evaluator minder weerstand op en blijft deze na de evaluatie ter plaatse om van de vooropgestelde actiepunten werk te maken (Love, 1991; Christie e.a., 2004).

Standaert (1999) stelt dat naarmate de kwaliteit van de interne evaluatie toeneemt, het aandeel van externe evaluatie kan verminderen. Los van de vraag of men deze stelling ondersteunt, dient opnieuw te worden benadrukt dat zelfevaluaties de geformuleerde rollen maar kunnen waarmaken wanneer deze met een gewaarborgde kwaliteit plaatsvinden. 'The question is whether schools can meet these high expectations' (Devos & Verhoeven, 2000, p. 405). Bovendien moet men in staat zijn om over de kwaliteit van de zelfevaluaties gefundeerde uitspraken te doen.

1.2.3 Uitdagingen voor de toekomst: de rol van zelfevaluatie in het volledige plaatje

Het uitgangspunt om de rol van zelfevaluatie in een Vlaams kwaliteitszorgsysteem vorm te geven, is het onderscheid tussen de verschillende verwachtingen die we onderscheidden bij het definiëren van het concept 'kwaliteit'. We maakten daar een onderscheid tussen twee groepen schoolexterne verwachtingen (al dan niet in wetgeving verankerd) en schoolinterne verwachtingen. We maakten daarnaast ook een onderscheid tussen het afleggen van verantwoording en schoolontwikkeling. Als men tot een goed functionerend kwaliteitszorgsysteem wil komen, zullen beide ons inziens duidelijk van elkaar moeten worden onderscheiden. In totaal zijn er dus zes domeinen waarin een vorm van kwaliteitszorg moet uitgewerkt worden (Tabel 2). Daarbij zal telkens een gepaste invulling

moeten gegeven worden aan de rol van externe en interne evaluatie. Het inhoudelijk invullen van de tabel is één van de belangrijke uitdagingen waarmee het onderwijsbeleid in Vlaanderen de komende jaren geconfronteerd zal worden.

Tabel 2: De 6 domeinen van kwaliteitszorg

	Decretaal verankerde verwachtingen	Niet decretaal verankerde externe verwachtingen	Schoolinterne verwachtingen
Verantwoording	Domein 1	Domein 3	Domein 5
Schoolontwikkeling	Domein 2	Domein 4	Domein 6

Het is niet onze bedoeling dit schema hier uit te werken. We hebben het in deze probleemstelling opgenomen om aan te tonen dat de rol van zelfevaluaties in het Vlaamse kwaliteitszorgsysteem nog niet eenduidig is vastgelegd. Bovendien geeft het aan dat zelfevaluaties in de toekomst wellicht nog aan belang zullen winnen. De theoretische en retorische verwachtingen ten aanzien van de effecten van zelfevaluatie op de effectiviteit van scholen zijn dan ook groot. We hebben evenwel herhaaldelijk benadrukt dat de logica omtrent de rol van zelfevaluatie in zowel het schoolontwikkelingsperspectief als het verantwoordingsperspectief maar kan opgaan, gegeven een gewaarborgde kwaliteit van de zelfevaluaties door scholen. Daarover is op dit moment weinig geweten. Het beleid wordt op dit moment in sterke mate uitgezet op basis van retoriek omdat systematisch verzamelde gegevens (in Vlaanderen) schaars zijn. Deze studie wil bijdragen tot het vergroten van de kennisbasis over het verloop en de resultaten van zelfevaluaties door (Vlaamse) scholen.

1.3 Zelfevaluaties in het hedendaagse Vlaamse onderwijs

Zowel scholen als de overheid nemen initiatieven om tot systematische vormen van zelfbevraging van scholen te komen. De voorbije jaren werden kaders ontwikkeld om dergelijke zelfevaluatie in goede banen te leiden en werden zelfevaluatie-instrumenten met zeer uiteenlopende focussen ontwikkeld. Nu scholen meer en meer werk maken van het uitvoeren van zelfevaluaties dringt zich de vraag op hoe kwaliteitsvol dit gebeurt en hoe verschillen in kwaliteit van zelfevaluaties tussen scholen verklaard kunnen worden. Ondanks de toegenomen aandacht zijn zelfevaluaties immers niet populair in het Vlaamse onderwijs. De doorlichtingservaring van de Vlaamse onderwijsinspectie leert bovendien dat de meeste scholen wel vertrouwd zijn met het begrip zelfevaluatie, maar dat ze een eigen, doorgaans beperkte invulling geven aan dat begrip (Van Hoyweghen, 2002). Scholen geven verschillende invullingen aan zelfevaluatie en beoordelen zichzelf vaak positief op dat vlak terwijl ze in realiteit nog maar aan het begin staan van een lange ontwikkeling (Standaert, 1999; Michielsens, 2002). Devos e.a. (1999) concluderen uit de getuigenissen van deskundigen dat er weinig secundaire scholen zijn die systematisch aan zelfe-

valuatie doen. De professionaliteit die het uitvoeren van een zelfevaluatie vergt, mag dus niet zomaar verondersteld worden (Gombeir, 1998; Blok e.a., 2005). Rydant (in Van Hoyweghen, 2001, p. 183) stelt dat de verwachting dat scholen zich autonoom kunnen opstellen om hun pedagogisch beleid vorm te geven vaak fictie is. Niet alle scholen beschikken volgens Rydant over het ontwikkelingsniveau en de bereidwilligheid om zich autonoom en onafhankelijk op te stellen. Bovendien zouden ze niet over de middelen beschikken om methodologisch en inhoudelijk tot goede zelfevaluaties te komen.

Meer recent komen de Vlaamse onderwijsinspectie en onderwijsonderzoekers naar aanleiding van een studie van de GOK-zelfevaluaties tot opmerkelijk gelijkaardige conclusies (Van Petegem e.a., 2005b; MVG, 2006). De doelstellingen en het nut van de GOK-zelfevaluaties zouden in scholen vaak onvoldoende gekend en/of verkeerd zijn. Het feit dat scholen vooral ervaring hebben met zelfevaluaties die opgelegd werden vanuit de overheid, zou er bovendien toe bijdragen dat het maken van een zelfevaluatie eerder als een verplichting beschouwd wordt met als voornaamste doelstelling 'in orde zijn' (beantwoorden aan de verplichtingen) dan als een instrument ter verbetering van de eigen schoolwerking. De doelen die scholen vooropstellen voor de zelfevaluatie stroken dus niet noodzakelijk met de grote verwachtingen die de overheid met het verplichten van de zelfevaluatie beoogt. Verder wordt evidentie gerapporteerd voor een gebrek aan openheid binnen het team en voor een tekort aan kritische ingesteldheid ten opzichte van het eigen functioneren. Het luidt dat de geesten vaak niet rijp zijn voor het uitvoeren van een zelfevaluatie. In veel scholen blijkt het niet vanzelfsprekend problemen, vragen, twijfels e.d. aan te kaarten en deze openlijk te bespreken. Evaluatie en zelfevaluatie worden vaak nog als bedreigend ervaren. Leraren staan nog steeds op hun autonomie binnen de klasmuren en voelen (zelf)evaluatie als een vorm van (sociale) controle. De mentale condities die een gunstige voedingsbodem moeten vormen voor het uitvoeren van zelfevaluaties zijn dus vaak niet gerealiseerd. Van Petegem et al. (2005) stellen verder ook vast dat de voorbereiding van een zelfevaluatie in heel wat scholen onvoldoende grondig aangepakt wordt. Het blijkt te ontbreken aan een kader waarin de zelfevaluatie op een transparante manier vorm kan krijgen. Tenslotte speelt ook de negatieve houding omwille van de gepercipieerde taakbelasting een innovatiebelemmerende rol. Schoolleiders en leerkrachten delen dezelfde weerstand ten aanzien van de papierlast die zelfevaluaties met zich meebrengen. Nochtans staan schoolleiders volgens deze onderzoekers doorgaans positiever tegenover zelfevaluatie en zijn ze meer overtuigd van het nut van zelfevaluatieactiviteiten dan hun leerkrachten.

Het blijkt dus niet evident voor scholen om in de eigen vertrouwde context het systematische en cyclische proces van de zelfevaluatie te initiëren en te realiseren. Het uitvoeren van zelfevaluaties is zoals beschreven nochtans een steeds belangrijker onderdeel van het kwaliteitszorgsysteem in het onderwijs. Dat impliceert dat aanzienlijk geïnvesteerd dient te worden in het vergroten van de kennisbasis omtrent het uitvoeren van zelfevaluaties en in de professionalisering van scholen inzake zelfevaluatie.

2 VANUIT DE NOOD AAN EMPIRISCHE EVIDENTIE NAAR META-EVALUATIE

Het groeiende belang van zelfevaluaties – als instrument voor schoolontwikkeling én verantwoording – maakt het nodig empirische evidentie te verzamelen over de processen en de resultaten van zelfevaluaties. Het opwaarderen van zelfevaluatie vereist immers een minimaal aanvaardbare kwaliteit van de zelfevaluatie en instrumenten om de kwaliteit van de zelfevaluatie in kaart te brengen. Er is nood aan een zogenaamde meta-evaluatie, waarbij de kwaliteit van zelfevaluaties geëvalueerd wordt (Scriven, 1969). Meta-evaluatie betekent dat een evaluatie op haar beurt wordt verheven tot het onderwerp van een evaluatie.

De empirische evidentie omtrent de kwaliteit van zelfevaluaties is vooralsnog beperkt. Kyriakydes & Campbell (2004, p. 32) stellen: '*The field of school self evaluation is at an early stage of development*'. Dat geldt bijvoorbeeld voor de kennisbasis over het verloop van zelfevaluaties en over kwaliteitsbevorderende kenmerken. De vraag welke kwaliteitsaspecten in rekening moeten gebracht worden of welke criteria aangewend kunnen worden om de kwaliteit van zelfevaluaties te beoordelen, is tot op heden in grote mate onbeantwoord. Op dit moment is het ook onduidelijk in welke mate scholen beschikken over voldoende professionaliteit om zelf strategieën, methodieken en criteria voor een kwaliteitsvolle zelfevaluatie uit te werken (Rogers & Hough, 1995). Ook de vraag naar de verschillen die er op dit moment bestaan tussen scholen is tot op heden onbeantwoord. Dat maakt het moeilijk om in te schatten welke verwachtingen er op dit punt aan scholen kunnen gesteld worden.

Uit het bovenstaande blijkt dat verscheidene aspecten van het uitvoeren van zelfevaluaties de aandacht van onderzoekers verdienen. In wezen komt het er op neer dat de determinanten van zowel het proces als van de resultaten van een zelfevaluatie onder de loep worden genomen. Vandaar: zelfevaluatie binnenstebuiten.

3 ZELFEVALUATIE BINNENSTEBUITEN: FOCUS, ORIENTERENDE ONDERZOEKSVRAGEN EN DOELEN VAN HET ONDERZOEK

De mate waarin scholen autonoom vorm kunnen geven aan kwaliteitsvolle zelfevaluatie is van bepalend belang voor het kwaliteitsbeleid van de school en voor de rol die interne evaluatie kan innemen in de kwaliteitscontrole door de overheid. Dit onderzoek wil bijdragen tot het vergroten van de kennisbasis over het succes waarmee scholen erin slagen om tot kwaliteitsvolle zelfevaluaties te komen. We willen de kwaliteit van zelfevaluaties beschrijven, beoordelen en verklaren. Het concept 'zelfevaluatie' wordt daartoe in paragraaf 3.1 eerst duidelijk afgebakend. Daarna zetten we in paragraaf 3.2 achtereenvolgens de onderzoeksvragen en de doelen van deze studie op een rij.

3.1 Focus van deze studie: de 'geïnstrumentaliseerde zelfevaluatie'

In de huidige literatuur en praktijkinvullingen van kwaliteitszorg in onderwijs speelt zelfevaluatie een groeiende rol. Onder de term 'zelfevaluatie' kan men evenwel verscheidene zaken verstaan; het dekt verschillende ladingen. Een eerste stap is dan ook het omschrijven van de invulling die zelfevaluatie in het volgende zal krijgen. De focus van deze studie is de 'geïnstrumentaliseerde zelfevaluatie'. Zowel 'zelfevaluatie' als 'geïnstrumentaliseerd' vergen een duidelijke omschrijving.

De term 'zelfevaluatie', als verzamelbegrip voor activiteiten gericht op het onderzoeken en beoordelen van het eigen functioneren, is al een tijd ingeburgerd (van Aanholt & Buis, 1990). De beschrijvingen van de term zijn talrijk. Wij baseren ons op de beschrijving die van Petegem (1998, p. 104) op grond van een studie van de evolutie in het denken en handelen met betrekking tot zelfevaluatie voorstelt: Zelfevaluatie is *'het proces, in hoofdzaak geïnitieerd door de school zelf, waarbij welgekozen participanten op een systematische wijze het functioneren van de school beschrijven en beoordelen met het oog op het nemen van beslissingen c.q. initiatieven in het kader van (aspecten van) de algehele school(beleids)ontwikkeling'*. Deze beschrijving maakt het immers op een eenvoudige manier mogelijk om variatie in zelfevaluaties te beschrijven (zie theoretisch kader).

Er zijn – in lijn met de bovenstaande omschrijving - talrijke vormen van zelfevaluatie die tot schoolontwikkeling kunnen bijdragen (denk aan interviews, reflectietanken, focusgroepen, themagerichte werkgroepen, fotoreportages, ...) en die scholen aanwenden in hun beleidsvoeringsproces. Wij hanteren een duidelijke afbakening in deze studie. We focussen op vormen van zelfevaluatie waarbij scholen gebruik maken van schriftelijke vragenlijsten/instrumenten die door (een selectie van) de teamleden ingevuld worden. We spreken daarom van geïnstrumentaliseerde zelfevaluaties. Zelfevaluaties die enkel gebaseerd zijn op het analyseren van uitstroomgegevens, op interviews of op klasbezoeken, behoren dus niet tot het domein van deze studie. De keuze voor geïnstrumentaliseerde vormen van zelfevaluatie is ingegeven door de noodzaak om voldoende uniformiteit in te bouwen in de bestudeerde verschijnselen. Bovendien gaat het om een uitvoerige en tevens veel gebruikte vorm van zelfevaluatie door scholen. Niet zelden wordt in het pleidooi voor zelfevaluatie in een ontwikkelings- en verantwoordingsperspectief een geïnstrumentaliseerde zelfevaluatie voor ogen gehouden. Het is een vorm van zelfevaluatie die hoge eisen stelt aan scholen. De keuze is echter geenszins normatief. We zijn er ons van bewust dat specifieke vormen van zelfevaluatie van onderzoek uitgesloten worden (zie bijvoorbeeld Van Petegem & Cautreels (2006)). Hoewel dus ook niet-geïnstrumentaliseerde vormen van zelfevaluatie waardevol zijn in het ontwikkelen en bewaken van onderwijskwaliteit, richten onze onderzoeksactiviteiten zich toch op zelfevaluaties waarbij wel sprake is van het gebruik van vragenlijsten. Wanneer er in het vervolg van dit boek sprake is van 'zelfevaluatie' gaat het dus om een geïnstrumentaliseerde vorm.

3.2 Oriënterende onderzoeksvragen en doelen van deze studie

Ook de onderzoeksvragen worden duidelijk afgebakend. Het ruimere kader reikte vanuit verschillende invalshoeken verscheidene legitieme vragen aan. Zonet werd de focus van deze studie duidelijk afgebakend. Vanuit het hiervoor behandelde denkkader komen we met die focus in het achterhoofd tot drie oriënterende onderzoeksvragen. Deze onderzoeksvragen zijn van verschillende aard. We onderscheiden beschrijvings-, verklarings- en ontwerpproblemen (Swanborn, 2000). De volgende oriënterende onderzoeksvragen zullen gaandeweg deze studie verfijnd worden.

1. Aan de hand van welke indicatoren kan men de kwaliteit van een zelfevaluatie operationaliseren?
2. Hoe kwaliteitsvol verlopen zelfevaluaties in Vlaamse scholen en wat zijn de verschillen tussen scholen terzake?
3. Hoe kunnen verschillen in de kwaliteit van zelfevaluatie verklaard worden?

De doelen van deze studie zijn niet alleen fundamenteel wetenschappelijk geïnspireerd. In de eerste plaats willen we met het uitvoeren van deze studie weliswaar bijdragen tot de kennisbasis omtrent zelfevaluaties in (Vlaamse) scholen, maar toch gaan onze betrachtingen ook verder. Zo willen we – bijvoorbeeld - op basis van de onderzoeksresultaten:

- bij scholen de aandacht vestigen op condities in de school die een gunstige voedingsbodem (moeten) vormen opdat zelfevaluaties op een kwaliteitsvolle wijze kunnen plaatsvinden.
- de sleutelelementen in kaart brengen waarop de ondersteuning die scholen genieten bij het uitvoeren van een zelfevaluatie zich best zou richten.
- beleidsmakers - door de stand van zaken inzake zelfevaluatie te beschrijven - informeren over de verwachtingen die ze ten aanzien van Vlaamse scholen mogen koesteren.
- een bijdrage leveren tot een instrumentarium dat het mogelijk maakt om de sterktes en zwaktes van scholen inzake zelfevaluatie te identificeren. Ook de stap naar meer gerichte ondersteuning kan dan op maat van de individuele school invulling krijgen.
- een bijdrage leveren tot een kader voor meta-evaluatie waarbij de gebruiker op een eenvoudige manier een beoordeling kan uitspreken over de kwaliteit van een zelfevaluatie die heeft plaatsgevonden. Dergelijke kaders zijn momenteel immers niet voorhanden (Blok e.a., 2005).

4 AFRONDEND

In het eerste deel van dit proefschrift zijn wij vanuit een ruime situering van zelfevaluaties in het bestaande, én evoluerende kwaliteitszorgsysteem gekomen tot drie oriënterende onderzoeksvragen. De volgende delen van dit proefschrift focussen op deze onder-

zoeksvragen. De ruimere context zoals gepresenteerd in dit deel wordt doelbewust pas in het laatste deel met conclusies weer opgenomen. De oriënterende onderzoeksvragen zetten de inhoudelijke bakens uit voor het verdere verloop van deze studie; niet in het minst voor de methodologische onderbouwing van het geheel. In deel 2 beargumenteren we dat de onderzoeksvragen een multimethode benadering vergen en geven we een overzicht van de verschillende onderzoeksactiviteiten die hiertoe zullen worden uitgevoerd.

DEEL 2

MULTIMETHODE ONDERZOEK

Om het verloop en de resultaten van zelfevaluaties in scholen te beschrijven en te verklaren, doen wij een beroep op verschillende vormen van gegevensverzameling, zowel kwalitatief als kwantitatief. Men spreekt dan van 'mixed method onderzoek' (Tashakkori & Teddlie, 2002) of van een 'multimethode onderzoeksdesign' (Brewer & Hunter, 1989). Wij kiezen voor deze laatste term, omdat onderzoekmethoden in deze studie niet 'gemixed' worden, maar elk hun eigen bijdrage leveren tot het beantwoorden van de onderzoeksvragen (Bergman, 2007). Het multimethode onderzoeksdesign is een procedure voor het verzamelen, analyseren en combineren van kwantitatieve en kwalitatieve gegevens in één studie.

Onze veronderstelling is dat kwantitatieve en kwalitatieve vormen van onderzoek elkaar niet uitsluiten. Hoewel ook onderzoekers de werkelijkheid door hun eigen lens benaderen (cf. een constructivistisch paradigma), is het toch mogelijk om voldoende objectieve uitspraken te doen over bijvoorbeeld het functioneren van scholen (cf. een positivistisch paradigma). In dit onderzoek is een combinatie van een inductieve én een deductieve logica aangewezen. Onze instelling in deze is pragmatisch (Creswell, 2005, p. 512). We erkennen dat de ene onderzoeksmethode (eerder) aansluit bij de ene onderzoeksbenadering (paradigma) en de andere bij een andere. We spreken ons echter niet uit in het voordeel van deze of gene benadering. Verschillende onderzoeksmethoden zijn hier nodig om een volledig beeld te krijgen van de hier bestudeerde sociale verschijnselen. Wij zien verschillende onderzoeksmethoden niet als een onderdeel van mutueel exclusieve benaderingen. Onderzoeksmethoden kunnen gesitueerd worden op een continuüm dat loopt van kwalitatief tot kwantitatief (Reichardt & Cook, 1979). Bij het bestuderen van zelfevaluaties willen wij verschillende posities op dit continuüm bestrijken. Dit deel van het proefschrift beschrijft hoe wij dit in deze studie trachten te realiseren.

In de volgende alinea's expliciteren we eerst de argumenten voor onze keuze voor een multimethode design. Daarna geven we in paragraaf 2 de grote lijnen van de onderzoeksopzet weer.

1 WAAROM EEN MULTIMETHODE ONDERZOEKSDESIGN?

De keuze voor het gebruik van een multimethode onderzoeksdesign is ingegeven door de meerwaarde van een gepaste sequentiële én concurrente combinatie van verschillende onderzoeksmethoden (Creswell, Clark, Gutmann & Hanson, 2003). We lichten deze vormen van combineren kort toe. Beide zijn de wezenlijke argumenten voor onze keuze voor een multimethode onderzoeksdesign.

Het eerste argument is de mogelijkheid om onderzoeksmethoden op mekaar verder te laten bouwen (cf. sequentieel). Het verloop van zelfevaluaties kan bijvoorbeeld eerst kwalitatief verkend worden om te verzekeren dat de juiste variabelen in het theoretisch kader opgenomen worden (dit is de verkennende benadering). Of nog: de in een survey-onderzoek vastgestelde relaties kunnen door middel van gevalstudies in een gecontextualiseerde setting 'betekenis' krijgen (dit is de verklarende benadering). De openheid naar verschillende onderzoeksmethodieken maakt het met andere woorden mogelijk om een *strategische sequens* in de onderzoeksbenadering in te bouwen.

Het tweede argument voor de keuze voor deze methodologische invulling is dat het beantwoorden van de onderzoeksvragen beter kan gebeuren door beide types onderzoek ook *concurrent* te verbinden. Dat betekent dat zij elkaar niet opvolgen maar gelijktijdig plaatsvinden. Op die manier kan men een beroep doen op de sterke punten van de verschillende manieren van werken (Brewer & Hunter, 1989; Johnson & Turner, 2003). Kwantitatieve gegevens kunnen door statistische analyse bijvoorbeeld een indicatie geven voor de relatieve impact van een bepaalde variabele. Kwalitatieve gegevens maken het op hun beurt mogelijk te achterhalen hoe men, bijvoorbeeld in scholen, over het bestudeerde fenomeen spreekt. De ene werkwijze kan bovendien de tekorten van een andere werkwijze neutraliseren. Het wezenlijke doel hiervan is te komen tot data- en methodentriangulatie. Dit houdt (onder meer (Erzberger & Kelle, 2003)) in dat verschillende onderzoeksmethoden hetzelfde fenomeen bestuderen en dat hun bevindingen in de analyse geïntegreerd worden (Miles & Huberman, 1994). De (generaliseerbare) bevindingen uit survey-onderzoek kunnen dan bijvoorbeeld getoetst worden aan de resultaten van contextinclusieve gevalstudies (Creswell, 2005).

Het bovenstaande impliceert niet dat een verregaande sequentiële én concurrente combinatie van verschillende onderzoeksmethoden per definitie wenselijk is. Morse (2003) waarschuwt immers dat men ook de efficiëntie van de ingezette middelen als een bijkomend praktisch criterium moet hanteren. Multimethode onderzoek is immers geen doel op zich maar zoals we beargumenteerden een middel om kwaliteitsgaranties in te bouwen. Het pleidooi situeert zich dus binnen de grenzen van het mogelijke/verantwoorde.

Uit de vorige alinea's blijkt dat de verschillende onderzoeksactiviteiten in een multimethode studie op drie verschillende manieren met elkaar in verbinding kunnen staan. Hun relatie is verkennend, verklarend of gericht op datatriangulatie. In onze invulling van het onderzoeksdesign komen de drie manieren voor en blijkt dat ook combinaties van deze relaties kunnen voorkomen.

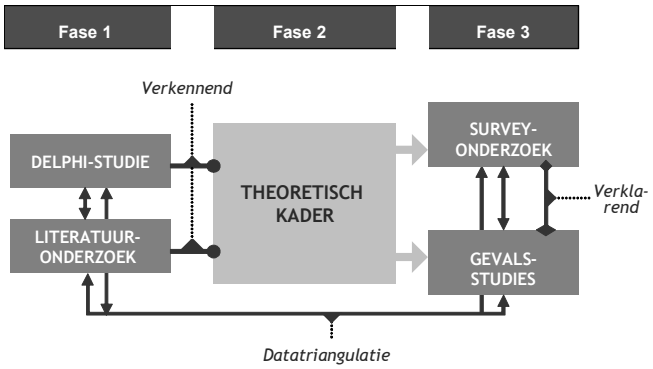
2 HET VERLOOP VAN HET MULTIMETHODE ONDERZOEK

De onderzoeksactiviteiten in deze studie omvatten een literatuuronderzoek, een Delphi-studie, survey-onderzoek en gevalstudies. Voor een goed begrip kenschetsen we elke onderzoeksactiviteit hier aan de hand van één zin. Voor een meer uitvoerige beschrijving van de verschillende activiteiten verwijzen we naar de beschrijving van de afzonderlijke studies. Vooralsnog zetten we enkel de grote methodologische lijnen van dit onderzoek uit.

- Het *literatuuronderzoek* brengt in kaart wat er over (het verklaren en beoordelen van) de kwaliteit van zelfevaluatie in de literatuur reeds beschreven is.
- De *Delphi-studie* omvat een systematische kwalitatieve bevraging van bevoorrechte getuigen die bij het uitvoeren van een zelfevaluatie betrokken (kunnen) zijn.
- Het *survey-onderzoek* bestaat (voornamelijk) uit een schriftelijke bevraging van ruim 2700 respondenten (schoolleiders, leerkrachten en ondersteunend personeel) in 96 scholen over hoe zij het verloop en de kwaliteit van de zelfevaluatie die in hun school plaats vond percipiëren.
- De *gevalstudies* omvatten een diepgaande studie van 3 scholen doorheen het uitvoeren van een zelfevaluatie door middel van diepte-interviews met verschillende teamleden.

Figuur 1 visualiseert de volgorde van en de relaties tussen de verschillende onderzoeksactiviteiten. De relaties worden weergegeven door middel van drie verschillende pijlvormen. De start bestaat uit een literatuuronderzoek en een Delphi-studie. We noemen deze in dit rapport de 'voorbereidende onderzoeksactiviteiten'. Beide gaan aan het ontwikkelen van het theoretisch kader vooraf. Het gebruik van een Delphi-studie past mooi in de hoger gepresenteerde verkennende sequentiële opbouw van onderzoeksactiviteiten in een multimethode aanpak. In zulk een verkennende benadering verzamelt men eerst kwalitatieve gegevens om het onderzoeksthema te verkennen. Vervolgens worden dan gegevens (onder meer kwantitatief) verzameld om de gevonden verbanden te toetsen. Dit laatste gebeurt hier in het survey-onderzoek en de gevalstudies. Creswell (2005, p. 516) geeft aan dat deze strategie het meest aangewezen is wanneer onduidelijk is welke variabelen ertoe doen en/of wanneer de instrumenten om de variabelen te meten nog niet beschikbaar zijn. Uit de probleemstelling bleek dat deze situatie in sterke mate opgaat voor deze studie. Deze aanpak heeft als voordeel dat men met een open houding aan de studie kan beginnen en niet start met een vooraf vastgelegde set van variabelen. Dit zijn ook meteen de verwachtingen die aan de grondslag lagen van onze keuze om een verkennende Delphi-studie uit te voeren. Gezien de eerder beperkte input uit het literatuuronderzoek is dus ook de Delphi-studie sterk richtinggevend bij het vastleggen van het theoretisch kader voor het verdere onderzoek. De twee voorbereidende studies brengen ons tot hypothesen die in de latere fasen van het onderzoek getoetst worden.

Figuur 1: Relaties tussen en timing van de onderzoeksactiviteiten



Het theoretisch kader wordt met name in twee studies getoetst, in het survey-onderzoek en in de gevalstudies. Toch wordt ook in het toetsen van hypothesen de link bewaard met de voorbereidende onderzoeksactiviteiten. Dat blijkt uit de talrijke pijlen die datatriangulatie symboliseren. In de conclusies over de verschillende onderdelen heen kijken we in welke mate bevindingen convergeren, consistent zijn en elkaar aanvullen. Wanneer dit niet of onvoldoende het geval is, gaan we op zoek naar mogelijke verklaringen. We spreken dus van triangulatie van de data verzameld gedurende vier onderzoeksactiviteiten. Dat neemt uiteraard niet weg dat ook van de afzonderlijke onderzoeksactiviteiten waardevolle resultaten kunnen verwacht worden.

De gevalstudies werden echter niet alleen met het oog op datatriangulatie geconcipieerd. Zij hebben namelijk ook een verklarende relatie met het survey-onderzoek. Dit verklarend verband tussen verschillende onderzoeksbenaderingen is wellicht de meest gekende vorm van multimethode onderzoek. De kwalitatieve gegevensverzameling zijn dan bedoeld als ondersteuning bij het verklaren en elaboreren van de kwantitatieve gegevens. De veronderstelling hierbij is dat de kwalitatieve onderzoeksbevindingen uit de gevalstudies het beeld dat verkregen wordt in het surveyonderzoek zowel kunnen verbreden als verdiepen (Creswell, 2005).

Door te kiezen voor het bovenstaande multimethode onderzoeksdesign komen we tegemoet aan de beperkingen die inherent zijn aan het geïsoleerde gebruik van deze of gene onderzoeksmethode (mono methode bias). Op deze manier worden garanties ingebouwd voor het bewaken van de validiteit van het onderzoek. In geval van convergente onderzoeksresultaten zal voor de validiteit van de resultaten meer bewijskracht ter beschikking zijn. In het geval van divergente en complementaire onderzoeksresultaten kunnen we komen tot belangrijke nuanceringen (Donche, 2005).

3 AFRONDEND

Dit deel zette de grote methodologische lijnen uit voor deze studie. Bij het bespreken van de verschillende onderzoeksactiviteiten wordt telkens meer diepgaand ingegaan op de methodologische verantwoording. Dat blijkt bijvoorbeeld in de voorbereidende studies.

DEEL 3

VOORBEREIDENDE STUDIES

Bij het beantwoorden van de vooropgestelde onderzoeksvragen past - gezien we ons onderzoeksmatig grotendeels op braakliggend terrein begeven - een verkennende benadering. In eerste instantie willen we het terrein van de meta-evaluatie dan ook exploreren. Dat doen we door middel van twee voorbereidende studies. Het doel van deze studies is te komen tot een gefundeerd referentiekader voor de volgende fases van deze studie. Het ontwerp van het referentiekader voor meta-evaluatie is dus gebaseerd op twee invalshoeken. Hoewel deze invalshoeken initieel bedoeld zijn om de initiële onderzoeksvragen te verdiepen, zullen uiteraard ook al aanzetten tot antwoorden opduiken. In eerste instantie willen we echter achterhalen welke aspecten van het uitvoeren van zelfevaluaties enerzijds en van het functioneren van scholen anderzijds in het kader voor meta-evaluatie opgenomen moeten worden. De twee invalshoeken die we hanteren om het referentiekader te informeren zijn:

> Invalshoek 1: Een verkennende Delphi-studie naar de visie van Vlaamse onderwijsexperten. Bij een groep van experten in het domein van zelfevaluaties brengen we een proces van kwalitatieve bevragingen op gang om verschillende standpunten en argumenten te exploreren.
>
> Invalshoek 2: Een literatuuronderzoek naar het verloop en de kwaliteit van zelfevaluaties en naar (f)actoren die van invloed zijn op deze kwaliteit.

De verantwoording, het verloop en de resultaten van deze twee onderzoeksactiviteiten worden in de volgende paragrafen beschreven. We staan achtereenvolgens stil bij de verkennende Delphi-studie en het literatuuronderzoek.

INVALSHOEK 1: EEN DELPHI-STUDIE NAAR DE VISIE VAN VLAAMSE ONDERWIJSEXPERTEN

Een (voor Vlaanderen) algemeen aanvaarde set van criteria om de kwaliteit van een zelfevaluatie te beschrijven, is bij de start van deze studie niet voorhanden. Het is dan ook van essentieel belang om verschillende argumenten, ervaringen en opinies van bevoorrechte getuigen in kaart te brengen. De eerste stap daartoe is een verkennend communicatieproces tussen verschillende onderwijsexperten op gang te brengen om alzo verschil-

lende standpunten en argumenten te exploreren. Zoals zal blijken is de Delphi-studie daartoe als onderzoeksmethodiek zeer geschikt. In de eerstvolgende paragraaf verantwoorden we de keuze voor de Delphi-studie als onderzoeksmethode en beschrijven we de gehanteerde werkwijze. Daarna worden onder de titel 'Kwaliteitsvolle zelfevaluaties volgens Vlaamse deskundigen' de resultaten van de bevraging gepresenteerd.

Doorheen het rapporteren van de onderzoeksresultaten wordt een (aanzet tot) antwoord gegeven op de drie oriënterende onderzoeksvragen. Deze zijn: (1) Aan de hand van welke indicatoren kan men de kwaliteit van een zelfevaluatie operationaliseren? (2) Hoe kwaliteitsvol verlopen zelfevaluaties in Vlaamse scholen en wat zijn de verschillen tussen scholen terzake? En (3) Hoe kunnen verschillen in de kwaliteit van zelfevaluaties verklaard worden?

1 WAAROM EEN DELPHI-STUDIE?

Het doel van de meeste Delphi-toepassingen is een verkenning van standpunten en de productie van gepaste informatie met het oog op het maken van beleidskeuzes (Adler & Ziglio, 1996). De Delphi-techniek werd dan ook ontwikkeld als een snelle manier om zicht te krijgen op de opinies van verschillende experts zonder hen met elkaar in contact te brengen (Linstone & Turoff, 1975). Kortweg bestaat de techniek uit de raadpleging van een groep vooraf geselecteerde informanten[1].

Oorspronkelijk werd de Delphi-methode alleen gehanteerd in de zin van 'de toekomst voorspellen'; de naam is immers ontleend aan het orakel van Delphi. Momenteel is de toepassing ervan ruimer. Van Tulder et al. (1989) schrijven dat de Delphi-methode een beproefde techniek is voor het leveren van een bijdrage aan het oplossen van een probleem door het belichten van zoveel mogelijk facetten vanuit een veelheid van gezichtspunten. Het is immers zo dat de collectieve wijsheid verder gaat dan individuele inzichten omdat een geheel van perspectieven verzameld en geanalyseerd kan worden. Hierdoor komt een completer beeld van de bestudeerde kwestie naar voren (Mahieu, 1999).

De specifieke aard van het in deze studie onderzochte onderwerp maakt dat wij getracht hebben de verschillende categorieën Vlaamse experts inzake zelfevaluatie in de Delphistudie te betrekken. Doordat iemand tot een bepaalde categorie behoort, kan verwacht worden dat hij of zij oog heeft voor juist die argumenten, zienswijzen, verwachtingen en wensen die in die belangengroep leven. Door het vertegenwoordigd zijn van de verschillende belangengroepen komt dan ook het complete plaatje naar voren (Swanborn, 1991). Werken met experts kan echter representativiteitsproblemen scheppen (Linstone & Turoff, 1975). Het is dan ook van groot belang dat alle standpunten vertegenwoordigd zijn

[1] We spreken in deze Delphi-studie over experts of informanten en niet over respondenten. Daarmee willen we het exploratieve karakter van deze gegevensverzameling in de verf zetten. Bovendien gaat het hier effectief eerder om 'informanten' die informatie geven over het object van dit onderzoek dan om 'respondenten' die informatie geven over zichzelf (Baarda, de Goede & Teunissen, 2001).

in de heterogene groep deskundigen die men samenstelt. Geen enkele invalshoek mag uit het onderzoek verdwijnen en alle visies moeten op tafel komen. In de opzet van de Delphi-studie trachtten wij dan ook de verschillende betrokkenen en belanghebbenden binnen het Vlaamse onderwijsveld op te nemen. We hebben echter niet de ambitie om op basis van de antwoorden van de respondenten generaliserende uitspraken te doen over specifieke doelgroepen in hun geheel. Met de door de respondenten aangedragen argumenten en aandachtspunten beogen we in de eerste plaats theorieontwikkeling en hypothesevorming.

2 OPZET VAN DE DELPHI-STUDIE

In de volgende paragrafen wordt achtereenvolgens stilgestaan bij het verloop van de Delphi-studie, de keuze van de respondenten en de gerealiseerde respons, en de inhoudelijke focus van de bevraging.

2.1 Het verloop van de Delphi-studie

Doorgaans worden in een Delphi-studie meerdere vraagrondes georganiseerd en worden de deelnemers bij elke nieuwe ronde geconfronteerd met de inzichten van de andere geraadpleegde deskundigen. Linstone et al. (1975) geven aan dat respondenten minstens éénmaal de kans moeten krijgen om de antwoorden van de gehele groep te herevalueren. Dit betekent dat een Delphi-studie minstens bestaat uit twee vraagrondes. Zoals uit de beargumentering van de opzet kan blijken, opteerden wij voor een schriftelijke eerste ronde gevolgd door een tweede mondelinge ronde bestaande uit diepte-interviews. Tussen de beide rondes werden de respondenten geïnformeerd over de voorlopige bevindingen.

De manier waarop men de vraagrondes organiseert, kan variëren. De Delphi-varianten hebben betrekking op de vraag of de bevraging mondeling of schriftelijk georganiseerd wordt, of men de respondenten rechtstreeks bij elkaar brengt en of men al dan niet tot een consensus onder de respondenten wil komen. In eerste instantie werden voorgestructureerde vragenlijsten aan een panel van deskundigen voorgelegd. Door deze wijze van bevragen, wilden we de eventuele invloed van dominantie van en het ontzag voor bepaalde respondenten beperken. De anonieme verzameling van de gegevens maakt dat deelnemers gemakkelijker een eigen en eventueel afwijkend of scherp standpunt kunnen verdedigen. Er treedt immers geen openlijk gezichtsverlies op. Bovendien zullen allerhande machtsmechanismen en belangenconflicten een minder grote rol spelen. Goed om weten is verder dat we in deze Delphi-studie niet streefden naar het bereiken van een consensus in de bevraagde expertengroep. Veeleer wilden we verschillende invullingen van kwaliteitsvolle zelfevaluatie aan het licht brengen. Ook in de tweede ronde werden gegevens daarom anoniem verzameld en verwerkt. Om tot diepgang en nuancering te komen enerzijds én om de tijdsinvestering van de respondenten te bewaken anderzijds

werd gekozen voor individuele diepte-interviews. Deze namen gemiddeld 80 minuten in beslag.

2.2 De keuze van informanten en respons

De selectie van informanten mag uiteraard niet afhangen van de persoonlijke voorkeuren van de onderzoeker. De selectieprocedure moet gebaseerd zijn op expliciete criteria (Adler & Ziglio, 1996). In deze Delphi-studie werden de twee volgende criteria vooropgesteld: (1) De verschillende belangengroepen inzake zelfevaluatie in Vlaamse scholen moeten in het onderzoek worden betrokken, en (2) de informanten in de verschillende groepen moeten geselecteerd worden op basis van een combinatie van hun ervaring en expertise enerzijds en de positie die zij in het Vlaamse onderwijsland innemen anderzijds (Vanhoof & Van Petegem, 2006a).

Om tegemoet te komen aan het eerste criterium werden de volgende categorieën informanten onderscheiden: beleidsverantwoordelijken, inspectieleden, netgebonden pedagogische begeleiders, niet-netgebonden begeleiders van zelfevaluatieprocessen, schoolleiders, leerkrachten en onderwijskundigen (onderzoekers). Om tot een verkenning van de onderzoeksvragen te komen, zijn daarmee de verschillende perspectieven en belangen in deze Delphi-studie vertegenwoordigd. Het tweede criterium werd nagestreefd door uit een lijst van mogelijke informanten die informanten te selecteren die het meest aan het criterium tegemoet kwamen. De selectiecriteria sluiten experten inzake zelfevaluatie in een niet schoolse context uit van deelname. Hoewel dit gezien de focus van deze studie goed te beargumenteren is, zijn we ons ervan bewust dat dit ook een nadeel kan zijn. Door enkel 'onderwijsexperten' te betrekken, is het immers mogelijk dat de zogenaamde blinde vlekken van de sector zonder meer overgenomen worden.

De doelstelling bij de aanvang van de Delphi-studie was een 15-tal informanten aan het exploratieve eerste deel van dit onderzoek deel te laten nemen. Hiertoe werd voor elk van de belangengroepen aan twee of drie personen gevraagd aan het onderzoek deel te nemen. In totaal hebben 14 informanten de schriftelijke vragenlijst bij de eerste ronde ingevuld. Daarbij werd verzekerd dat elke categorie vertegenwoordigd zou zijn. Aan de diepte-interviews bij de tweede ronde namen 16 informanten deel. Om de verdeling over de verschillende belangengroepen te bewaken, werd in deze ronde ruimer gerekruteerd onder schoolleiders en leerkrachten. In principe namen de informanten echter aan beide bevragingsrondes deel.

2.3 De inhoudelijke focus van beide Delphi-rondes

Zoals in de meeste Delphi-toepassingen stelde de vragenlijst in de eerste ronde het voorliggende onderwerp in vrij algemene termen aan de orde (Adler & Ziglio, 1996, p. 9). Op die manier wilden we de respondent zo neutraal als mogelijk uitnodigen om antwoorden en opinies te formuleren. Dat gebeurde op basis van vragen zoals: Welke criteria zijn volgens u geschikt om het onderscheid te maken tussen een kwaliteitsvolle en een niet of

minder kwaliteitsvolle zelfevaluatie? De tweede ronde verdiepte de resultaten van de eerste ronde. Aan alle deelnemers werd een samenvatting van de resultaten van de eerste ronde toegestuurd. Tijdens de diepte-interviews bij de tweede ronde van de Delphi-studie kregen de respondenten de kans om hun originele antwoorden aan te vullen en/of te nuanceren in het licht van de eerdere bijdrage van de gehele respondentengroep. Er werd eveneens een aantal elementen bevraagd die op basis van de antwoorden op de eerste vragenlijst onvoldoende bleken uitgewerkt. In grote lijnen was de inhoud van de tweede ronde evenwel gelijk aan deze van de eerste ronde. Omdat we geïntegreerd rapporteren over beide rondes presenteren we ook de vragen tezamen. Dat brengt ons tot het volgende overzicht.

> *Deel 1: Wat te verstaan onder* (kwaliteitsvolle) *zelfevaluatie?* Hoe zou u een (kwaliteitsvolle) zelfevaluatie definiëren? Wat zijn noodzakelijke ingrediënten? Welke criteria zijn volgens u geschikt om het onderscheid te maken tussen een kwaliteitsvolle en een niet of minder kwaliteitsvolle zelfevaluatie? Tot welke resultaten zou een zelfevaluatie moeten komen om als kwaliteitsvol te worden bestempeld? Kunnen die resultaten a priori vastgelegd worden?
>
> *Deel 2: Welke zijn mogelijk voorspellers van kwaliteitsvolle zelfevaluatie*? Welke culturele en organisatorische voorwaarden moeten in een school vervuld zijn om tot een kwaliteitsvolle zelfevaluatie te komen? Uit welke kenmerken van het verloop van een zelfevaluatie kan volgens u worden afgeleid of het resultaat ervan al dan niet kwaliteitsvol zal zijn? Door welke kenmerken van het functioneren van de school als geheel wordt het verloop van een zelfevaluatie bepaald?
>
> *Deel 3: Zijn de voorwaarden om een succesvolle zelfevaluatie uit te voeren vervuld in Vlaamse scholen?* Welke wel/niet? Welke zijn de belangrijkste sterktes van Vlaamse scholen om tot kwaliteitsvolle zelfevaluaties te komen? Ten aanzien van welke voorwaarden schieten Vlaamse scholen volgens u tekort om kwaliteitsvolle zelfevaluaties uit te voeren?
>
> *Deel 4: Omtrent de ondersteuning van scholen.* Op welke manier doen scholen bij uw weten momenteel beroep op ondersteuning bij het uitvoeren van zelfevaluaties? Kunnen scholen momenteel voldoende ondersteuning genieten bij het uitvoeren van zelfevaluaties? Voor welke instanties ziet u daarin welke rol weggelegd? In welke mate komt de beschikbare ondersteuning tegemoet aan de behoeftes van scholen?

2.4 Analyse van de gegevensbronnen

De gegevensbronnen bij deze Delphi-studie bestonden uit de schriftelijke antwoorden van respondenten op de vragenlijst bij de eerste ronde en de getranscribeerde interviews bij de tweede ronde. Voor de analyse van deze gegevens baseerden we ons op de grounded theory van Glaser & Strauss (1967). Gegeven de exploratieve aanpak opteerden wij ervoor om de analyse te baseren op de inbreng van de informanten zonder expliciete verwijzing naar reeds bestaande theoretische kaders. De analyse gebeurde handmatig door één

onderzoeker. In eerste instantie werden de data meermaals grondig doorgelezen. Het doel daarvan was eerst de grote lijnen te zien in plaats van meteen op kleine onderdelen te focussen. Vervolgens werd de inhoud van de tekstsegmenten gelabeld aan de hand van codes. Deze codes werden gekozen op basis van het woordgebruik van de respondenten (cf. in vivo). Alle relevante tekstfragmenten werden ondergebracht onder één of meer zulke codes (zoals 'inhoud van de zelfevaluatie', 'instrumentkeuze', 'behoefte aan ondersteuning', 'communicatie', 'reflectie' en dergelijke). Daarbij werd ook telkens nagegaan voor welke onderzoeksvraag of -vragen een tekstsegment informatief was. Een tekstfragment kan immers gelijktijdig informatie aanbrengen bij de vraag naar relevante voorspellers van kwaliteitsvolle zelfevaluatie en bij de vraag naar de stand van zaken in Vlaamse scholen. Er was op deze manier sprake van een inductief proces van het verengen van de data in een beperkt aantal thema's (Creswell, 2005). Na dit codeerproces werden gelijkaardige codes op elkaar afgestemd, werden complementaire codes in een overkoepelende categorieën gegroepeerd en werden redundante codes verwijderd. Daarbij werd uitgegaan van de sturing door de onderzoeksvragen, van de samenhangen en verschillen die de onderzoeker in de gegevens herkende en de wens om de gegevens op een overzichtelijke manier te presenteren. De rol van de onderzoeker bestond verder uit het categoriseren, samenvatten en stroomlijnen van de antwoorden. Er werd daarbij zowel gezocht naar gelijkenissen als naar verschillen. In de volgende rapportering van de resultaten zal echter blijken dat de meningen sterk convergeren. In de rapportering zijn alle argumenten die door de respondenten aangereikt worden opgenomen, ook al zijn deze uniek of onderling tegenstrijdig. We streven immers geen veralgemening van de onderzoeksresultaten na maar een inventarisatie van relevante standpunten. Het doel de Delphi-studie en de rapportering ervan is in de eerste plaats theorieontwikkeling en hypothesevorming.

3 RESULTATEN: (KOMEN TOT) KWALITEITSVOLLE ZELFEVALUATIES VOLGENS VLAAMSE DESKUNDIGEN

De inhoud van de Delphi-vragenlijst vormt in grote lijnen meteen ook de structuur van de beschrijving van de resultaten. We onderscheiden vier paragrafen: Het omschrijven van (kwaliteitsvolle) zelfevaluatie; de voorspellers van kwaliteitsvolle zelfevaluatie, de stand van zaken inzake zelfevaluaties in Vlaanderen en de (doeltreffendheid van de) ondersteuning waarover scholen kunnen beschikken.

3.1 Het omschrijven van (kwaliteitsvolle) zelfevaluatie

In het omschrijven van kwaliteitsvolle zelfevaluatie bouwen we twee stappen in. Eerst staan we stil bij de invulling van het concept zelfevaluatie. Daarna stellen we de vraag naar kwaliteitsvolle zelfevaluatie.

3.1.1 Wat is zelfevaluatie?

De activiteiten die scholen ondernemen in het kader van kwaliteitszorg zijn zeer verscheiden. Sommige daarvan noemen de informanten zelfevaluatie en anderen niet. Zowel in de literatuur als in het onderwijsveld bestaan dan ook talrijke omschrijvingen van 'zelfevaluatie' en van concepten die naar hetzelfde verwijzen. We gingen allereerst na of de invulling van het concept door onderwijsexperts strookt met de bestaande invullingen. Daartoe inventariseerden we welke de noodzakelijke ingrediënten zijn om van een zelfevaluatie te kunnen spreken.

De informanten omschrijven zelfevaluatie als 'een cyclisch proces waarbij een school, op eigen initiatief en vanuit een overkoepelend kwaliteitszorgconcept, zelf op een systematische manier aspecten van het eigen functioneren beschrijft en beoordeelt met als doel (indien nodig) te komen tot concrete verbeterprocessen'. Deze omschrijving omvat alle bestanddelen die volgens de geraadpleegde Vlaamse experten noodzakelijk zijn om van een zelfevaluatie te kunnen/mogen spreken. Het volgende kader gaat kort in op hun invulling van de verschillende bestanddelen.

> *Cyclisch proces*. De informanten geven aan dat zelfevaluatie geen op zichzelf staand proces is. Zelfevaluatie past in een cyclisch kader en omvat het nastreven van herhaling, om te zien of er evolutie is in de tijd.
>
> *Op eigen initiatief*. Het gaat om een bewuste keuze van de school zelf om het proces te starten. De keuze om het proces te initiëren ligt bij de school zelf.
>
> *Zelf*. Het is de school die de verantwoordelijkheid draagt voor de vormgeving van en uitvoering van de sleutelelementen van het zelfevaluatiegebeuren.
>
> *Vanuit een overkoepelend kwaliteitszorgconcept*. Zelfevaluatie wordt vormgegeven vanuit een duidelijke visie op wat kwaliteitsonderwijs is en waar men met de school naartoe wil.
>
> *Systematisch beschrijven.* Zelfevaluatie verwijst naar het systematisch verzamelen en analyseren van informatie. Alles gebeurt bewust en met duidelijke intenties in het achterhoofd. De verschillende stappen van het proces sluiten goed op elkaar aan.
>
> *Systematisch beoordelen.* De vaststellingen worden door toetsing aan het ruimere *kwaliteitszorgconcept* beoordeeld. De interpretatie en beoordelingsfase wordt daarbij losgekoppeld van de beschrijvingsfase.
>
> *Komen tot concrete verbeterprocessen.* Het doel van zelfevaluaties is te komen tot een actieplan dat tot een verbeterd functioneren bijdraagt. Let wel: er is sprake van 'indien nodig'. Zelfevaluatie kan dus ook leiden tot het borgen van wat goed is.

Zelfevaluatie is op deze manier beschreven in de meest zuivere vorm. Het niet volledig gerealiseerd zijn van één bestanddeel betekent volgens de informanten niet dat er geen sprake meer kan zijn van een zelfevaluatie. De invulling van het begrip 'zelfevaluatie' is

volgens de informanten niet zo zwart/wit. Het is volgens hen aangewezen de verschillende bestanddelen te beschouwen als kenmerken waaraan een zelfevaluatie in meer of mindere mate kan voldoen. Wanneer dat in sterke mate het geval is, gaat het om een zelfevaluatie. Wanneer dit in mindere mate het geval is, gaat het onder een minder zuivere vorm van zelfevaluatie. We illustreren deze gedachte[2]:

- Aan de hand van de eis dat er bij zelfevaluatie steeds uitgegaan wordt van een door de school zelf geïnitieerde activiteit. Een informant stelt hierover: *'Ik denk dat je scholen wel kan verplichten om aan zelfevaluatie te doen. Maar je kan niet zeggen rond welke dingen er zou moeten geëvalueerd worden en op welke manier die zelfevaluatie moet gebeuren. Als de overheid zegt tegen een school, u moet aan zelfevaluatie doen, is dat nog altijd - behalve die verplichting - heel sterk zelfevaluatie. (...) Maar eens dat men heel expliciete regels begint vast te leggen, van wat die zelfevaluatie moet inhouden, zit je volgens mij steeds minder in de zelfevaluatie.'*
- Aan de hand van de eis van systematische gegevensverzameling. *'Ook scholen die zeggen: Oké, we gaan een keer voor ons zelf nadenken over wat dat een goede school moet zijn en we gaan op basis van gewone gesprekken een aantal criteria vast leggen om de werking van de school daar aan te toetsen. Dat is dan geen instrument, maar het is wel een mogelijke vorm van zelfevaluatie. Omdat het ook weer een soort van systematische informatieverzameling is. Die misschien niet zo hoogstaand wetenschappelijk is, maar die wel afstand neemt van de dagelijkse praktijk.'*
- Aan de hand van de eis van het cyclische karakter. *'Is een eenmalige zelfevaluatie, dan geen zelfevaluatie? Ik weet het niet. Het is ook een zelfevaluatie, maar het is niet de ideale manier om aan zelfevaluatie te doen.'*

In het licht van de bovenstaande bemerkingen zijn zelfevaluaties dus geen kwestie van alles of niets, maar eerder van meer en minder. De meest zuivere vorm van zelfevaluatie is dan deze die perfect aan de gepresenteerde beschrijving voldoet. Door de variatie binnen de verschillende bestanddelen, kent zelfevaluatie volgens de informanten dan ook talrijke verschijningsvormen.

3.1.2 Wanneer is zelfevaluatie kwaliteitsvol?

Het blijkt niet eenvoudig om kernachtig te omschrijven wat een kwaliteitsvolle zelfevaluatie is. In deze paragraaf gaan we hier daarom expliciet op in. We stellen de vraag welke volgens de informanten de resultaten van een zelfevaluatie moeten zijn om deze als kwaliteitsvol te bestempelen. In lijn met onze omschrijving van kwaliteit als 'het op een gepaste manier voldoen aan de door de betrokkenen gedeelde verwachtingen', werd in

[2] De verschillende citaten in deze beschrijving van de resultaten zijn afkomstig uit de diepte-interviews die in de tweede ronde van de Delphi-studie plaats vonden.

de Delphi-studie immers gepeild naar welke deze verwachtingen bij Vlaamse onderwijsexperts zijn. Daarbij maken de informanten een onderscheid tussen het komen tot actiepunten (en de kenmerken waaraan deze actiepunten moeten voldoen) enerzijds en concrete resultaten anderzijds.

a. KOMEN TOT ACTIEPUNTEN (EN DE KENMERKEN WAARAAN DEZE MOETEN VOLDOEN)

Zelfevaluatie moet volgens de informanten voldoende heldere en relevante informatie verschaffen om beslissingen te nemen voor de verdere ontwikkeling van de school. Een kwaliteitsvolle zelfevaluatie blijft volgens hen dus niet steken bij de vaststellingen maar resulteert in het ondernemen van acties (try-outs, bijsturingen, reorganisaties, …). Herhaaldelijk benadrukt men dat er iets moet gebeuren met de resultaten van de zelfevaluatie. Kwaliteitsvolle zelfevaluatie staat dus duidelijk in het teken van actieplannen en behoort volgens de informanten aanleiding te geven tot acties die de schoolontwikkeling bevorderen. Het belangrijkste doel van zelfevaluatie moet het realiseren zijn van aantoonbare en duurzame verbeteringen. (Vernieuwde) doelen moeten vertaald worden in concrete acties, prioriteiten moeten vastgelegd worden en volgens een afgesproken tijdspad uitgevoerd. Een zelfevaluatie is in de ogen van de informanten dus maar kwaliteitsvol als er een vertaling plaatsvindt van de resultaten naar beleidsaanbevelingen voor klas- en schoolpraktijk. Op die manier willen de informanten verzekeren dat men komt tot de voorbereiding en planning van verbeteractiviteiten. Of tot de borging van wat goed gaat. De mate waarin een zelfevaluatieproces in gepaste actiepunten resulteert, is volgens de informanten de ultieme parameter voor de kwaliteit van zelfevaluaties.

> *'Je zou criteria kunnen hanteren omtrent de actiepunten waarin een zelfevaluatie uitmondt. 95% van de actiepunten is veel te vaag, te algemeen en te abstract. Bijvoorbeeld: We moeten meer doen aan de teamwerking. We moeten meer 'zelfstandig leren' leren. Dit is dus een criterium waar je de kwaliteit van zelfevaluaties op kan beoordelen. Hoe concreet zijn de actiepunten? Hoe realistisch zijn ze bijvoorbeeld? Zijn er op korte termijn mogelijk zaken te realiseren of niet? Begint men met vrij eenvoudige zaken zodat dat mensen een succeservaringen opdoen, om de grotere groep mee te krijgen? Of blijft het gewoon bij het formuleren van een aantal dingen waar dat dan - naar implementatie toe - verder niets mee gebeurt.'*

Kwaliteitsvolle zelfevaluatie resulteert volgens de respondenten dus in passende actiepunten. Informanten verwijzen herhaaldelijk naar de criteria waaraan deze actiepunten behoren te voldoen. De criteria die men aanhaalt, komen in wezen neer op het SMART-principe. De actiepunten moeten zodus:

- *Specifiek zijn*. Dat betekent dat ondubbelzinnig omschreven wordt wat men precies gaat doen.
- *Meetbaar zijn*. Dat betekent dat actiepunten zodanig geformuleerd worden dat ze op hun beurt het voorwerp kunnen uitmaken van een evaluatie. Het is dan duidelijk waaraan kan afgetoetst worden of men het actiepunt al dan niet bereikt.

- *Acceptabel zijn*. Dat betekent dat de doelgroep het actiepunt aanvaardt en dat er een draagvlak voor is.
- *Realistisch zijn*. Dat betekent dat het actiepunt haalbaar moet zijn.
- *Tijdsgebonden zijn*. Dat betekent dat wordt aangegeven tegen wanneer het actiepunt gerealiseerd zou moeten zijn.

Actiepunten die aan deze kenmerken voldoen zijn volgens de informanten richtinggevend. Ze geven aan wat men wil bereiken en sturen gedrag van het schoolteam. Bovendien wordt aangegeven welke resultaten wanneer moeten worden bereikt. Door actiepunten op deze manier te formuleren is de kans volgens de informanten groter dat er in de praktijk iets van terecht komt. Vandaar dat ze als criterium voor het beoordelen van de kwaliteit van zelfevaluaties naar voren worden geschoven.

b. CONCRETE RESULTATEN

Het komen tot actiepunten staat in de bovenstaande beschrijving los van concrete inhouden. Er zijn echter ook informanten die concrete resultaten vooropstellen om de kwaliteit van zelfevaluaties te beoordelen. Zelfevaluatie is volgens hen dan ook niet alleen een middel om tot actiepunten te komen. Het uitvoeren van zelfevaluatie kan ook op zich tot concrete waardevolle resultaten leiden. De uitkomsten die herhaaldelijk als waardevol resultaat van zelfevaluatie worden gerapporteerd zijn:

- Stimuleren van gesprekken over de eigen opvattingen omtrent goed onderwijs.
- Bevorderen van een meer 'reflectieve houding' en van de wil om kritischer tegen de zaken aan te kijken.
- Onderkennen van zowel positieve als negatieve aspecten van het schoolfunctioneren.
- Resulteren in extra motivatie om kwaliteit te leveren.
- Versterken van de betrokkenheid van teamleden bij het schoolbeleid.
- Verduidelijken van verantwoordelijkheden.
- Bijdragen tot samenwerking en gezamenlijke doelgerichtheid in de school.
- Versterken van de samenhang in de school.

Naast de vraag of zelfevaluatie tot passende actiepunten leidt, is volgens de informanten dus ook de vraag naar de mate waarin concrete resultaten zoals de bovenstaande gerealiseerd worden indicatief voor de kwaliteit van zelfevaluaties.

3.2 De voorspellers van kwaliteitsvolle zelfevaluatie

Uit de omschrijving van het begrip 'zelfevaluatie' kan men volgens de informanten verscheidene algemeen geldende elementen afleiden die bepalend zijn voor de kwaliteit van een zelfevaluatie. Daarnaast schuiven zij een grote batterij aan losse indicatoren naar voren waarvan zij verwachten dat ze kwaliteitsbevorderend zijn. De indicatoren die in overweging zouden kunnen worden genomen bij meta-evaluaties zijn dan ook talrijk. In het volgende presenteren we deze indicatoren in een overzichtelijk kader. Een eenvoudige maar zeer bruikbare indeling daartoe bestaat uit het categoriseren van de verschillen-

de indicatoren naar het *waarom*, het *wat*, het *hoe,* het *wanneer* en het *wie* van een zelfevaluatie (Van Petegem, 1998). Figuur 2 visualiseert de kern van de verschillende voorspellers.

Figuur 2: De voorspellers van kwaliteitsvolle zelfevaluatie volgens Vlaamse experts

We kunnen volgens de informanten stellen dat een zelfevaluatie kwaliteitsvol zal zijn wanneer deze doelgericht verloopt (waarom?) en zich op relevante aspecten van het schoolfunctioneren richt (wat?). Daarnaast voldoen de voorbereiding en het verloop van de zelfevaluatie best aan een aantal eisen (hoe?); is het van belang dat de juiste mensen bij de zelfevaluatie betrokken worden (wie?) en dat de schoolcultuur voldoende rijp is voor het uitvoeren van zelfevaluaties (wanneer?). Deze verschillende vragen worden op basis van de inbreng van de informanten verder ingevuld. Het antwoord op de wanneer-vraag wordt daarbij geïntegreerd in het antwoord op de hoe-vraag.

3.2.1 Het waarom? - Voorspellers met betrekking tot de doelen van de zelfevaluatie

De vraag naar waarom een zelfevaluatie uitgevoerd wordt, speelt volgens de informanten een cruciale rol in het bewaken van de kwaliteit van zelfevaluaties. Kwaliteitsvolle zelfevaluaties zetten volgens hen duidelijke en gedeelde doelen voorop, zoeken aansluiting bij de bestaande schoolvisie en gaan uit van haalbare intenties. We lichten deze kwaliteitskenmerken in volgende alinea's verder toe.

a. DOELGERICHTHEID

De doelgerichtheid van zelfevaluaties wordt sterk in de verf gezet door de informanten. Bij de start van een zelfevaluatie moet duidelijkheid geschapen worden over de intenties die men bij de zelfevaluatie heeft. De doelen van de zelfevaluatie dienen voor alle betrokkenen duidelijk te zijn en er mag volgens de informanten geen sprake zijn van ver-

borgen agenda's. De school moet weten op welke doelen ze mikt en moet proberen deze doelen meetbaar te maken. De resultaten van de zelfevaluatie moeten met andere woorden aan de doelen kunnen afgetoetst worden. Een kwaliteitsvolle zelfevaluatie begint daarom volgens sommige informanten met een reflectie op hoe men de kwaliteit van de zelfevaluatie zal evalueren.

b. GEZAMENLIJKE DOELGERICHTHEID

Om tot kwaliteitsvolle zelfevaluatie te komen, moeten de doelen niet alleen duidelijk zijn, ze moeten volgens de informanten ook gedeeld worden door het schoolteam. Zowel de schoolleiding als het leerkrachtenkorps moeten overtuigd zijn van de meerwaarde van de zelfevaluatie. De informanten schuiven de mate waarin het schoolteam achter de zelfevaluatie staat en de manier waarop men in de wandelgangen over de zelfevaluatie spreekt herhaaldelijk als voorspellers van de kwaliteit van zelfevaluaties naar voren. De schoolleiding moet daartoe vanaf het begin duidelijk aangeven dat er rekening zal worden gehouden met de resultaten van de zelfevaluatie. Het team moet er – zo geven de informanten aan - van overtuigd zijn dat de resultaten van de zelfevaluatie indien nodig zullen leiden tot aanpassingen van het bestaande beleid. Een zekere gedragenheid door interne (personeelsleden, leerlingen en ouders) en externe partners (inspectie, werkgevers, vervolgonderwijs) is volgens de informanten dus een voorwaarde om tot kwaliteitsvolle zelfevaluaties te komen.

> *'In het beste geval is het doel voor iedereen duidelijk. Dat lijkt mij toch wel een groot winstpunt. Toch zit men altijd met de spanning tussen de persoon en de organisatie, en die is dat, stel dat je inderdaad als persoon niet akkoord gaat met dat doel. Bij mij, in elk geval, komen er dan spanningen op. En ik zal dan wel een aantal formele acties doen om dat doel mee te realiseren, maar toch voel ik mij dan een beetje, … voel ik dat dat toch slabakt of dat ik mij uiteindelijk toch onttrek aan een aantal dingen.'*

> *'Ja. Euhm… daar heb ik grote twijfels bij: in welke mate… men zegt wel, we bepalen allemaal samen mee de organisatie, maar ik zie toch heel dikwijls dat het bepaalden zijn die eigenlijk al lijnen uitzetten, en ikzelf heb daar eigenlijk last van.'*

Kortom: de inbreng van de informanten leert dat kwaliteitsvolle zelfevaluatie berust op een dragende visie die door iedereen gekend is en gedeeld wordt. De zelfevaluatie moet met andere woorden gerealiseerd worden vanuit een zo breed mogelijk draagvlak.

c. VERANKERING IN EEN RUIMER KADER

De doelen die men met de zelfevaluatie wil bereiken, dienen volgens de informanten verder ook inhoudelijk aan te sluiten bij de bestaande schoolvisie. Dat betekent dat een kwaliteitsvolle zelfevaluatie volgens hen een duidelijke visie impliceert over wat kwaliteitsonderwijs is. Kwaliteitsvolle zelfevaluaties vereisen volgens een respondent uitgeschreven visieteksten. In essentie gaat het daarbij om een duidelijke kijk op de kwaliteit van leren en onderwijzen enerzijds en op de schoolcultuur en het daarvan afgeleide personeels-, onderwijskundig- en logistiek beleid anderzijds. Verder wordt ook aangegeven

dat zelfevaluaties kwaliteitsvoller zullen zijn naarmate ze meer geïntegreerd worden in de bestaande denkkaders en in bestaande structuren.

> *'Wat toch zeker zou moeten terugkomen, is dat je bij evaluatie vertrekt vanuit duidelijke indicatoren. Dat je duidelijk op voorhand bepaalt rond welke punten dat je meet om iets in zicht te krijgen. Dus met andere woorden, je moet heel goed weten waar dat je met je school naar toe wil gaan, om te kunnen gaan bepalen op welke manier en welke dingen je gaat meten.'*

> *'In eerste instantie, die visie moet er zijn. Een school moet een beleidsplan hebben. (...) De essentie is dat iedereen die bij de zelfevaluatie betrokken is goed weet, wat dat de bedoeling is, waar men naar toe wil. Dat dat niet iets is, dat u door de directie opgedrongen wordt. Dat het duidelijk gemaakt wordt voor iedereen die daar bij hoort.'*

3.2.2 Het wat? - Voorspellers met betrekking tot de inhoud van de zelfevaluatie

Een aantal informanten verwijst naar de keuze van zelfevaluatiethema's om het onderscheid te maken tussen een kwaliteitsvolle en een niet of minder kwaliteitsvolle zelfevaluatie. Dat criterium leunt enigszins aan bij de doelgerichtheid van een zelfevaluatie. Zo wordt er bijvoorbeeld gesteld dat de mate waarin men het in de school eens is over de te evalueren thema's een voorspellende kracht heeft. Ook de link met de schoolvisie is hier aanwezig. De inhoudelijke keuze van zelfevaluatiethema's kan volgens een groep informanten best aansluiten bij de visie en het pedagogische project van de school. De selectie van te evalueren domeinen of thema's moet met andere woorden relevant zijn voor het gehele schoolgebeuren. Men geeft bovendien aan dat het moet gaan om elementen waarover de school zelf controle heeft.

> *'Omdat je anders veel dingen gaat vast stellen… ofwel is dat een sociologische realiteit ofwel is dat de overheid. Dat is nu eenmaal omdat ons schoolbestuur zo is georganiseerd of weet ik veel. Dat is dan vaak wel zeer triestig. (...) Uw zelfevaluatie is in principe gericht op schoolontwikkeling, dus dat je zelf er ook iets aan kan verbeteren hé.'*

Over welke thema's daartoe in aanmerking komen worden geen uitspraken gedaan. Dat kan volgens een aantal informanten ook niet eenduidig vastgelegd worden zonder weet te hebben van het specifieke karakter, inzake doelen en inhoud, van elke zelfevaluatie.

3.2.3 Het hoe? - Voorspellers met betrekking tot het verloop van de zelfevaluatie

De informanten doen verschillende suggesties omtrent de manier waarop zelfevaluatieprocessen best verlopen en omtrent de criteria die men kan gebruiken om beoordelende uitspraken te doen over die processen. Het gaat als het ware om de voorwaarden die men volgens de informanten zou moeten realiseren opdat een zelfevaluatie de vooropgestelde verwachtingen ook effectief zou realiseren.

> *'Ik kan me nauwelijks voorstellen, dat het uiteindelijk resultaat op iets kan trekken, als er geen enkele vorm van passend proces is aan vooraf gegaan. Dus met andere woorden, ik verdeel bijvoorbeeld IZES hier in de school, we nemen*

dat af en ik bekijk de resultaten. We bespreken dat met het directieteam en formuleren daar 3 à 4 conclusies uit. Dat is flauwekul. Als je niet met uw team vooraf besproken hebt, wat dat de vragenlijst inhoudt, wat je eigenlijk gaat doen en dan ook de conclusies van uw vragenlijst ook bespreekt met uw team. En uiteindelijk, in functie daarvan de stap zetten naar mogelijke acties. Als je dat niet doet, dan ben je nergens.'

De informanten stellen verschillende sleutelelementen voorop met betrekking tot het verloop van zelfevaluaties. In het volgende onderscheiden we communicatie, de bereidheid om samen te werken en om kritisch te reflecteren, veranderingsbereidheid, realiteitszin en systematiek. De beschrijving maakt duidelijk op welke manier de informanten verwachten dat deze proceskenmerken van invloed zijn op de kwaliteit van zelfevaluaties.

a. HET BELANG VAN COMMUNICATIE

Wanneer hen gevraagd wordt naar voorspellers van kwaliteitsvolle zelfevaluatie beginnen de meeste informanten met het benadrukken van het belang van een passende communicatie. Communicatie mag volgens hen in geen enkele fase van het zelfevaluatieproces over het hoofd worden gezien. Een doordachte aanpak van de communicatie blijft tot aan het eind van het proces van cruciaal belang.

'De essentie is communicatie. Dat is een aandachtspunt bij de start, maar eigenlijk continu. In de verschillende stappen moet men ook tijd nemen, om te zorgen dat iedereen mee is. Dat iedereen blijft weten, waarover het gaat. Waar men naar toe wil. Dat het bovendien ook niet beschouwd wordt als opnieuw een briefje invullen.'

'Je moet ook goed weten hoe je dingen presenteert naar de groep waarmee je iets wil bereiken. Ik denk dat hier heel wat problemen in scholen gerezen zijn: hoewel men de eerste stappen goed zet, dat men in de laatste fase heel slecht communiceert, en dat maakt alles ongedaan. Want als je door het slecht presenteren weerstand opwekt, dan gaat het effect van je beleidsinitiatieven, de ontwikkeling die je koppelt aan je zelfevaluatie, verloren.'

In de eerste plaats verwijzen informanten naar de noodzaak van openheid in het communiceren over het eigen functioneren. Om tot kwaliteitsvolle zelfevaluaties te komen, moeten teamleden bereid zijn om op een open manier te reflecteren op hun eigen functioneren. Openheid is – in de ogen van de informanten - in de eerste plaats een kenmerk van een schoolcultuur waarin van de leden van het schoolteam verwacht wordt dat ze openhartig en eerlijk communiceren over de eigen ervaringen, opinies en verwachtingen. De informanten benadrukken ook het belang van toegankelijke fora voor het bespreken van procedures, resultaten en beleidsaanbevelingen. Een goed verloop van de communicatie doorheen een zelfevaluatieproces vergt volgens hen daarom de aanwezigheid van een coördinerende persoon (een ankerfiguur of kwaliteitscoördinator) of werkgroep.

b. HET BELANG VAN DE BEREIDHEID OM SAMEN TE WERKEN

Zelfevaluaties hebben volgens de informanten meer kans op slagen in scholen waar samenwerking sterk gewaardeerd wordt. Dat houdt in dat de teamleden er teamgericht denken en van mening zijn dat ze door samen te werken méér kunnen bereiken. Vlaamse onderwijsexperts omschrijven dit nogal eens als een collegiale sfeer of als een veilig klimaat. Cruciaal in dat opzicht is het bestaan van vertrouwen tussen de teamleden enerzijds en tussen de schoolleiding en het team anderzijds. Zelfevaluaties kunnen volgens de informanten maar slagen als er tenminste minimale samenwerkingsverbanden zijn. Daarbij mag echter niet uit het oog verloren worden – zo geeft men aan - dat dit niet ten koste mag gaan van een voldoende grote autonomie voor de teamleden.

c. HET BELANG VAN DE BEREIDHEID TOT REFLECTIE

Met reflectie verwijzen informanten naar de bereidwilligheid om het bestaande in vraag te stellen en alternatieven in overweging te nemen. Dat vergt wat informanten een 'open houding' noemen. Teamleden met een open geest ervaren een wederzijdse kritische en constructieve ingesteldheid als een vanzelfsprekendheid. Met de zogenaamde reflectiecultuur verwijzen de informanten naar de verwachting dat teamleden zichzelf als persoon en als team in vraag durven stellen en bereid zijn te leren uit de eigen fouten. Bereidheid tot reflectie impliceert dat verhoudingen tussen teamleden normaal blijven en dat de leden spontaan en open blijven functioneren, ook al is er sprake van een vorm van evaluatie. *'Veel zelfevaluaties lopen fout omdat men het als een bedreiging gaat zien.'* Men moet over alle schoolaspecten mogen en durven praten zonder dat dit een bedreiging inhoudt. Spanningen en conflicten moeten beheersbaar en bespreekbaar zijn alvorens zelfevaluaties tot waardevolle resultaten kunnen leiden.

> *'Je kan niet beginnen met zelfevaluatie, als je niet een beetje een schoolcultuur hebt, die daar open voor staat. Als je ook geen open attitude hebt van uw personeel, ten aanzien van wat zelfevaluatie is en impliceert. Als je dat niet hebt! Dan heeft zelfevaluatie zelfs geen enkele zin.'*

De informanten geven aan dat een reflectiecultuur zich maar kan ontwikkelen als ook de schoolleiding (directie, middenkader én de inrichtende macht) op een constructieve manier met feedback kan omgaan. Een respondent geeft aan dat het daarom voor sommige scholen goed zou zijn om eerst aan die openheid te werken vooraleer een zelfevaluatie als dusdanig op het getouw te zetten. Ook de aanwezigheid van een externe kritische vriend is daarbij volgens verschillende respondenten aangewezen. Men geeft immers aan dat het bijzonder moeilijk is om als organisatie de eigen blinde vlekken te zien. *'Het is gezond dat er in een context van goed gezag en vertrouwen een externe bij betrokken wordt.'* Reflectiebereidheid houdt verder ook een responsieve houding in ten aanzien van prikkels uit de omgeving. Ook dat zou de zelfevaluatie ten goede komen.

> *'Een school die zich opsluit binnen eigen groep en eigen muren, die mist één en ander. Die kan evengoed een zelfevaluatie doen, maar het gevaar is groot dat een aantal blinde vlekken niet zichtbaar gaan worden. En een school die heel ontvankelijk is voor die prikkels vanuit de omgeving, denk ik dat veel meer kansen heeft om daarop in te grijpen. (...) Ze mag zich daar ook niet door laten*

leiden, maar dat ze daar ontvankelijk voor is, en dat ook aan een kritische vraagstelling onderwerpt, intern dan, dat vind ik deel van verantwoordelijkheid nemen van een school. En uiteindelijk is het dat hé. Zelfevaluatie is verantwoordelijkheid nemen door een school.'

d. HET BELANG VAN BEREIDHEID TOT VERANDERING

Veranderingsbereidheid leunt sterk aan bij de bereidheid tot reflectie. Sommigen informanten spreken van een verbetercultuur en een enkeling spreekt van durf. Daarbij denkt men bijvoorbeeld aan de mate waarin problemen en werkpunten tijdens de zelfevaluatie gezien worden als kansen of uitdagingen. Om een zelfevaluatie tot een goed einde te brengen moeten teamleden oplossingen ook durven en willen zoeken in veranderingen in hun eigen functioneren. Momenteel verwacht men nog vaak dat (eerst) de anderen moeten veranderen.

'Ik herinner mij een directeur die mij onlangs zei: 'als je de mening vraagt van de leerkrachten, dan is maar één die moet veranderen, dat ben ik.'

e. HET BELANG VAN REALITEITSZIN

De informanten benadrukken herhaaldelijk dat scholen niet te ambitieus mogen zijn bij de start van een zelfevaluatie. Er moet gewerkt worden aan een haalbaar traject, zowel in de betekenis van inhoudelijke complexiteit als in termen van de duur en omvang van de bevraging. Ook het uittrekken van voldoende tijd en middelen voor de verschillende stappen in het zelfevaluatieproces blijkt cruciaal. Kwaliteitsvolle zelfevaluatie start volgens de informanten gefaseerd, in kleine stappen en biedt voldoende garanties voor consolidatie. Scholen moeten voldoende prioriteiten bepalen, maar niet alles tegelijk willen aanpakken. De informanten adviseren (zeer) beperkt en bescheiden te beginnen om daarna langzaam uit te bouwen. Dat vergt een realistisch beeld van de eigen mogelijkheden en een houding waarbij beide voeten op de grond worden gehouden. Ook de hoeveelheid vernieuwingen van de afgelopen jaren en de manier waarop de zelfevaluatie daarbij aansluit is volgens een aantal informanten een aandachtspunt om tot haalbare doelen te komen. Tenslotte waarschuwt men voor het gevaar van enquêtemoeheid.

f. HET BELANG VAN SYSTEMATIEK

In de omschrijving van kwaliteitsvolle zelfevaluatie bleek reeds het belang dat informanten hechten aan systematiek. Daarbij doelen ze zowel op het ruimere kader waarin de zelfevaluatie uitgevoerd wordt als op aspecten van de zelfevaluatie zelf.

- *In het ruimere kwaliteitszorgsysteem*

Er wordt door de informanten gewezen op het belang van een overstijgend kwaliteitsstappenplan waarin men de zelfevaluatie kan situeren. In dat plan zouden de beleidslijnen, de procedures, de strategie, de planning en de eigenlijke zelfevaluatieactiviteiten vastgelegd moeten worden. Exemplarisch wordt verwezen naar de mogelijkheden van een cyclisch verloop volgens de PDCA-cirkel: plan (kwaliteitsbepaling, normstelling), voer uit (doelen en acties), controleer (meet, evalueer) en borg, stel bij of verbeter.

> *'Scholen zien zelfevaluatie vaak als: we gaan dat nu eens doen, en we hebben dit en dat al ingevuld, en dat is het resultaat. Dat ze dus stoppen bij dat resultaat. Dat ligt daar dan op papier en dat is het dan. Dus dan denken ze dat ze zelfevaluatie hebben gedaan. Vandaar het bedrog van de term dus denk ik. Het zou beter opvolging en voortgang of zoiets zijn, iets dat meer op de dynamiek wijst.'*

Het ruime referentiekader is verder ook van belang in het beoordelen van de zelfevaluatieresultaten. Scholen moeten volgens een aantal informanten zorgen voor een leidraad, een legitimatiebron om het eigen functioneren en de eigen verwezenlijkingen aan te toetsen.

- *In de planning en coördinatie van de zelfevaluatie*

Een kwaliteitsvolle zelfevaluatie wordt volgens verschillende informanten gekenmerkt door de aanwezigheid van een systematisch opvolgingssysteem dat het mogelijk maakt te achterhalen wat er tijdens het zelfevaluatieproces goed loopt, wat minder en waarom. Er moet volgens de informanten werk gemaakt worden van een globaal plan waarin de zelfevaluatie plaatsvindt (met o.a. aandacht voor voorbereiding, communicatie, doelstellingen, procedures, verloop, gegevensverwerking, valorisatie). In wezen geven ze aan dat alle werkzaamheden in het kader van een zelfevaluatie zorgvuldig gepland en gecoördineerd zouden moeten worden. De procedures en verantwoordelijkheden worden verder best (voldoende) formeel vastgelegd om continuïteit en stabiliteit te verzekeren. Een informant stelt dat adequate procedures noodzakelijk zijn als men wil dat afspraken gerespecteerd worden. Deelnemers aan een zelfevaluatie moeten hun eigen plaats en verantwoordelijkheden in het zelfevaluatiegebeuren goed kennen. Pas als de verschillende stappen in het zelfevaluatieproces vooraf worden vastgelegd, kunnen ze volgens de informanten transparant en hanteerbaar gemaakt worden.

Het verzekeren van opvolging tijdens het zelfevaluatieproces wordt voortdurend benadrukt als een noodzakelijke voorwaarde voor een kwaliteitsvol verloop van een zelfevaluatie. Er wordt daartoe gepleit voor de aanwezigheid van een sturend, dragend 'platform' op school. De zelfevaluatie wordt in dat geval kritisch door een stuurgroep opgevolgd en begeleid. Verschillende informanten benadrukken daartoe het belang van een goed functionerend middenkader. Ondanks de noodzaak aan formele afspraken en procedures wordt evenwel ook het belang benadrukt van het waarborgen van de toegankelijkheid van de verantwoordelijken voor de zelfevaluatie.

- *In de gegevensverzameling*

Een systematische methodologie is volgens de informanten ook tijdens de eigenlijke gegevensverzameling een voorwaarde om tot bruikbare resultaten te komen. Dat betekent volgens hen onder meer dat er betrouwbare instrumenten gebruikt worden, dat de gegevensverzameling zorgvuldig gebeurt, dat de respons voldoende hoog is, dat er een correcte verwerking van gegevens plaatsvindt, en dergelijke. Deze eisen moeten garanderen dat de gegevens die door middel van een zelfevaluatie-instrument verzameld worden valide zijn en vlot gegeneraliseerd kunnen worden.

> *'Je moet zorgen, dat er een minimum is aan degelijkheid. Maar je moet niet verwachten dat de scholen onderzoekers worden. Een schooldirecteur is geen onderzoeker. Maar er moet een minimale orde zijn rond die gegevensverzameling, er moeten volgens mij minimale kwaliteitseisen gerespecteerd worden. Ik denk dan ook aan primaire dingen hé. Een respons van 12 % en toch uitspraken beginnen te doen over de school is zus of zo… aan zo'n dingen aandacht besteden is wel terecht.'*

> *'Je moet daar toch een aantal conclusies uit kunnen trekken, die niet ter discussie staan. Anders kan je er niks mee doen. Uw zelfevaluatie moet legitiem zijn.'*

Systematiek in de gegevensverzameling vergt de inzet van goede instrumenten. De informanten vullen het begrip instrument ruim in. Het gaat in de gegevensverzameling duidelijk niet alleen om schriftelijke vragenlijsten. Ook (meer) kwalitatieve manieren om gegevens te verzamelen kunnen op een systematische manier ingevuld worden. De informanten spreken bijvoorbeeld van interviews, focusgesprekken, reflectietanks, documentanalyses en cijferanalyses.

- *In de keuze van het zelfevaluatie-instrument*

De antwoorden van de informanten leren dat de keuze van een gepast zelfevaluatie-instrument geen vanzelfsprekendheid is voor scholen. Toch is die keuze bepalend voor het verloop en de resultaten van de zelfevaluatie. Afhankelijk van de visie, de zelfevaluatiedoelen, de bedoelde inhouden, de aanpak en de beschikbare tijd moet een school komen tot een zelfevaluatie-instrument dat bij de eigen situatie aansluit. De kwaliteit van zelfevaluaties ligt volgens informanten hoger wanneer scholen een systematische afweging maken van verschillende mogelijkheden/instrumenten om de zelfevaluatie uit te voeren.

> *'Je kan zeggen, we kiezen voor IZES bijvoorbeeld, daarom en daarom en daarom,… en dat dan bespreken. Dan daar openlijk een uitspraak over doen, van dat is beter dan andere instrumenten daarom en daarom en daarom. Dan zeg ik, dan bent u goed bezig. Maar je kunt ook zeggen, de directeur beslist ik ga IZES gebruiken, want dat is nu eenmaal een instrument dat makkelijk toe te passen is. Daar is zelfs een Cd-rom van en andere scholen doen dat ook. Dat is een minder goede vorm van hoe je het proces doormaakt (...) Je moet die discussie mogelijk maken en bij die keuze een openheid toelaten, en ook zorgen, dat het instrument voldoende aansluit bij, waar dat je als school naar toe wil.'*

Van een zelfevaluatie-instrument verwachten de informanten dat het de te bestuderen thema's ontnuchtert. Dat betekent dat men loskomt van percepties, dat men focust op feiten en dat 'kritiek op anderen' omgebogen wordt naar zelfkritiek. Het instrument moet het mogelijk maken om het onderzochte thema uit te drukken in éénduidige en meetbare indicatoren. Om dat op een verantwoorde manier te doen moet het gaan om een gevalideerd instrument dat aansluit bij de doelen en de thema's van de zelfevaluatie. Scholen moeten volgens de informanten een instrument kiezen dat meet wat men wil meten (cf. validiteit). Bij de keuze van een instrument moet een school daarom achterhalen wat de visie is achter de instrumenten die men ter beschikking heeft. Een informant geeft aan dat het *'ideaal zou het zijn wanneer de achterliggende visie van het instrument aansluit*

bij de visie van de school'. De visie van de ontwikkelaars van het instrument moet voor de deelnemers aan de zelfevaluatie duidelijk zijn en moet zo goed mogelijk gerespecteerd worden om het instrument op een geoorloofde manier aan te wenden. Gepaste instrumenten houden volgens de bevraagde informanten bovendien rekening met de context waarin de school functioneert. Dat betekent volgens hen dat het ook wenselijk kan zijn om als school zelf een eigen instrument te ontwikkelen.

De dataverzameling die met het gebruik van een instrument gepaard gaat, biedt volgens de informanten best ook een passende vergelijkingsbasis. Als een school zich kan afmeten tegen duidelijke benchmarks of standaarden komt dat volgens een aantal informanten de zelfevaluatie ten goede. De informanten bepleiten ook de herhaalbaarheid van zelfevaluatie-instrumenten zodat door opeenvolgende metingen evoluties vastgesteld kunnen worden. Zo mogelijk wordt ook een combinatie van kwalitatieve en kwantitatieve bevindingen als nastrevenswaardig vooropgesteld. Dat alles mag echter niet in te sterke mate ten koste gaan van de gebruiksvriendelijkheid van het instrument.

3.2.4 Het wie? - Voorspellers met betrekking tot leiderschap en betrokkenheid

Verschillende informanten benadrukken dat de schoolleiding voldoende competent moet zijn om tot een kwaliteitsvolle zelfevaluatie te komen. De schoolleider moet in de eerste plaats een motor zijn, maar hoeft niet noodzakelijk zelf de uitvoerende verantwoordelijkheid te dragen over het zelfevaluatieproces. De informanten lijken het meeste heil te verwachten van een krachtig leiderschap in combinatie met een participatieve besluitvorming. Dat houdt volgens de informanten in dat betrokkenen kunnen meedenken, mee overleggen en mee beslissen over het verloop van de zelfevaluatie en over de ontwikkeling van de school als geheel. Er wordt bijvoorbeeld aangegeven dat een zelfevaluatie kwaliteitsvol zal zijn als een representatieve werkgroep (liefst met een kritische vriend) instaat voor de coördinatie en uitwerking van heel het proces. De meeste informanten gaan ervan uit dat zelfevaluaties kwaliteitsvoller zullen verlopen naarmate meer personen die betrokken partij zijn bij het onderzochte subject betrokken worden. Een participatieve benadering doorheen de zelfevaluatie is volgens hen met andere woorden een sterke voorspeller, niet in het minst omwille van het motiverende effect dat ervan zou uitgaan. Scholen waar de leerkrachten het gevoel hebben dat er naar hen geluisterd wordt, zullen volgens de informanten tot kwaliteitsvollere zelfevaluaties komen. Zelfevaluaties hebben in deze logica baat bij een minder hiërarchische en meer gelijkwaardige verhouding tussen teamleden.

Toch zijn er nuanceverschillen in de antwoorden omtrent de concrete invulling van de wie-vraag. Zo stelt men bijvoorbeeld dat een zelfevaluatie kwaliteitsvol is als (1) er zoveel mogelijk participanten betrokken worden, (2) als het personeel betrokken wordt bij alle stappen in het proces, of (3) als de betrokkenheid van de schoolleider en van een grote groep teamleden of een duidelijk gemandateerde kerngroep gegarandeerd is. De invulling van het antwoord op de wie-vraag kan dan ook niet a priori vastgelegd worden. Het komt er volgens de informanten in wezen op neer dat de zelfevaluatie gedragen moet worden door het schoolteam en niet alleen door enkele ingewijden. Een informant stelt:

'Ik vind, van wie het initiatief uitgaat is niet zo bepalend, in de zin van wie de eerste is die met het voorstel komt is niet bepalend. Maar wie het initiatief draagt is wel belangrijk.' De informanten geven aan dat de betrokkenheid en de verantwoordelijkheden van de teamleden - afhankelijk van het doel en de thema's van de zelfevaluatie - telkens anders moeten ingevuld worden.

Bij het beantwoorden van de wie-vraag wordt door verschillende informanten ook verder gekeken dan de betrokkenheid van zij die deel uitmaken van de school zelf. De kwaliteit van zelfevaluatie zou ook wel varen bij de betrokkenheid van een extern iemand die het zelfevaluatiegebeuren ondersteunt en vanuit die positie feedback kan geven op de vaststellingen.

> *'Ik denk dat wanneer een school in haar proces van zelfevaluatie externen meeneemt, of kritische vrienden meeneemt of… als je uit de onwilligen ook put bijvoorbeeld. Ik bedoel, dat zijn allemaal objectiveerbare elementen die de kwaliteit verhogen. Maar zeker ook de omgeving meenemen, externen. Als TSO-school de bedrijven waar je stages mee doet laten meespelen.'*

3.3 Stand van zaken inzake zelfevaluatie in Vlaamse scholen

In deze paragraaf stellen we ons de vraag naar de mogelijkheden en beperkingen van Vlaamse scholen om tot kwaliteitsvolle zelfevaluaties te komen. We gaan na of de voorwaarden om een kwaliteitsvolle zelfevaluatie uit te voeren volgens de informanten (voldoende) vervuld zijn. Die vraag is volgens de informanten evenwel niet eenduidig te beantwoorden. De realiteit bestaat volgens hen uit erg uiteenlopende situaties. Er zijn verschillende ritmes. Sommige scholen hebben al een hele weg afgelegd terwijl andere nog terugschrikken om hun werking in vraag te stellen. Hoewel algemeen geldende antwoorden dus niet mogelijk zijn, lijkt de algemene teneur evenwel eerder pessimistisch. Het aantal scholen met een degelijke zelfevaluatiepraktijk is volgens de informanten vooralsnog beperkt. Het hoeft dan ook niet te verbazen dat de lijst met troeven minder omvangrijk is dan het overzicht van de bestaande tekorten in Vlaamse scholen.

3.3.1 De troeven van Vlaamse scholen

De vaststelling dat zelfevaluatie voor Vlaamse scholen steeds meer een middel blijkt om tot schoolontwikkeling te komen, is volgens de informanten onder meer te verklaren door het stimulerende effect dat van de ruimere beleidscontext uitgaat. De informanten geven allereerst aan dat scholen kunnen werken aan zelfevaluatie zonder erop afgerekend te worden. Ondanks de sterke aanbeveling door de doorlichtende inspectie is het uitvoeren van zelfevaluaties op dit moment een vrijwillig proces. Bovendien zijn scholen volgens een informant vertrouwd met passende referentiekaders omdat ze daar via de schooldoorlichting mee geconfronteerd werden. Verder wordt aangehaald dat de kwaliteit van zelfevaluaties een impuls krijgt van de klemtoon op lokale autonomie en het groeiende besef bij scholen dat hiermee verantwoording ten opzichte van de maatschappij samen-

hangt. De voorbije jaren zou mede daardoor het besef dat scholen zelf verantwoordelijk zijn voor hun kwaliteit gegroeid zijn.

De informanten zien bovendien verscheidene ontwikkelingen die meer kwaliteitsvolle zelfevaluaties met zich mee kunnen brengen. Zo wordt er meer en meer gestreefd naar de participatie van relevante actoren in beleidsprocessen zodat het gevoel van betrokkenheid groeit. Scholen willen volgens de informanten het gevoel van verbondenheid en van onderlinge afhankelijkheid versterken. Er is ook steeds meer sprake van het doelbewust vooropstellen van doelstellingen op school- en klasniveau en van het werken met functiebeschrijvingen en functioneringsgesprekken. Scholen willen volgens de informanten hun professionele ontwikkeling meer in eigen handen nemen. De informanten rapporteren op basis van hun eigen ervaringen een groeiend engagement en reflecterend vermogen van schoolleiders en leerkrachten. Men geeft aan dat scholen - ondanks statutaire en reglementaire beperkingen - proberen om toch een personeelsbeleid te voeren waarbinnen taakdifferentiatie onder meer met het oog op het uitvoeren van zelfevaluaties mogelijk wordt.

Tenslotte zou ook de schaalgrootte van Vlaamse scholen de kwaliteit van zelfevaluaties ten goede komen. Toch wordt ook aangehaald dat een aantal scholen – voornamelijk basisscholen - door een te beperkte schoolgrootte en de daarmee samenhangende schoolorganisatie, personeelsbezetting en financiële situatie nauwelijks aan zelfevaluatie toekomt.

3.3.2 De knelpunten in Vlaamse scholen

De zwakke punten die door de informanten aangehaald worden, kunnen telkens in de vorm van een tekort beschreven worden. In het volgende onderscheiden we zes tekorten: een gebrek aan visie en doelgerichtheid, een gebrek aan openheid, een gebrek aan betrokkenheid, een gebrek aan deskundigheid en opvolging, een gebrek aan realisme en een gebrek aan middelen.

Een gebrek aan visie en doelgerichtheid. Volgens de informanten ontbreekt het in de meeste scholen aan visie omtrent de plaats van zelfevaluatie in het geheel van kwaliteitszorg. Scholen zien nog te weinig in dat zelfevaluatie niet los staat van de rest van de schoolwerking. Dat maakt dat werkpunten te weinig opgenomen worden in bestaande beleidsplannen.

> *'Een belangrijke vraag is: in hoeverre zijn die acties die men doet gelinkt aan het pedagogisch project van de school? En op welke manier maakt die zelfevaluatie daar eigenlijk ook een deel van uit. Ofwel doet men zo maar iets, ofwel zit er een zekere visie achter. Dat is toch wel een groot verschil. Bouwt men voort op iets waar dat men mee bezig is of zit men zomaar lukraak iets te doen hé. Hoe ad hoc is de actie?'*

Een gebrek aan openheid. Op dit moment is een gebrek aan openheid in veel scholen nog een belangrijk tekort. Dit uit zich op verschillende manieren. In nogal wat scholen heerst er volgens de informanten een eerder negatieve sfeer omdat mensen niet vrijuit over hun

bezorgheden en problemen kunnen praten. In veel scholen zou de afstand tussen bestuur en dagelijkse leiding, maar vooral tussen bestuur, personeel en leerlingen zeer groot zijn. Men geeft aan dat leerkrachten vaak nog een solistenrol spelen en dat er nog maar weinig sprake is van een cultuur van echt teamgericht werken. Daarnaast ervaren de informanten dat omgaan met feedback (kritiek) niet vanzelfsprekend is. Dat vertaalt zich onder meer in een laag welbevinden bij het personeel. Werkpunten worden nog teveel bekeken als problemen en niet als uitdagingen. Dit alles leidt tot een afwijzende opstelling ten aanzien van vernieuwingen waar ook de kwaliteit van zelfevaluaties onder kan lijden.

> *'Zelfs al is het in het kader van ontwikkeling, toch heeft iedere groep de neiging om er zo goed mogelijk te willen uitkomen. Bij onze vakwerkgroepen is dat ook. Je hebt doelstellingen, en je moet op het einde nakijken, zijn die nu al dan niet bereikt. Daar zit toch altijd een stukje in van fraaier de zaken voorstellen, want je kunt toch niet zeggen... Want dan moet je aanduiden waarom het fout gelopen is. Dan merk je toch dat iedereen een beetje de neiging heeft om dingen te verbloemen. Dat heeft met openheid en een houding van vertrouwen in de school te maken.'*

Een gebrek aan betrokkenheid. In veel scholen is het draagvlak volgens de informanten onvoldoende groot om tot kwaliteitsvolle zelfevaluaties te komen. Nogal wat personeelsleden zouden initiatieven inzake zelfevaluatie als bedreigend aanvoelen omdat er een te sterke sturing is door directie, omdat men vreest voor negatieve persoonsgebonden vaststellingen of omwille van de onduidelijkheid over het gebruik (misbruik) van de zelfevaluatieresultaten. Als de beslissing om een zelfevaluatie uit te voeren enkel genomen wordt door de schoolleiding kan dit bij de rest van het korps voor een gevoel zorgen van '*u vraagt en wij draaien*'. De informanten geven aan dat zulke problemen kunnen resulteren in onverschilligheid vanwege het korps. Of de teamleden zich betrokken voelen en overtuigd zijn van de meerwaarde hangt volgens hen vooral af van (1) de vraag of de directie hieromtrent een duidelijke visie heeft (2) en of de directie die visie 'verkocht' krijgt bij de teamleden. Dat is volgens de informanten nog te weinig het geval.

> *'Ik denk dat dergelijke dingen op een participatieve manier tot stand moeten komen. Dat ook iedereen heel goed moet weten, wat de bedoeling is van die zelfevaluatie. Anders riskeer je dat er achterdocht ontstaat. Als mensen niet weten waar iets naartoe gaat, dan ontstaan er weerstanden en dan ga je dus de grootste moeite hebben om die zelfevaluatie kwalitatief te krijgen en om met die resultaten iets te kunnen doen.'*

Betrokkenheid bij zelfevaluatie beperkt zich volgens de informanten nog te vaak tot de gegevensverzameling als dusdanig. Het schoolteam neemt dan enkel deel aan de bevraging maar niet aan de besluitvorming. Participatie bij de besluitvorming zou minder vanzelfsprekend zijn, zowel vanuit het standpunt van de schoolleiding (want leerkrachten hebben geen zicht op het geheel) als vanuit het standpunt van de leerkrachten (het kost te veel tijd en energie). Deelname aan en zeggingskracht tijdens het vormgeven aan de zelfevaluatie is volgens de informanten echter cruciaal om tot betrokkenheid te komen. Op die manier kan bewaakt worden of het gebeuren wel dicht genoeg bij de praktijk van

de leerkrachten aanleunt. Pas als zij zich herkennen in de resultaten zal er actiebereidheid ontstaan.

Tenslotte wordt ook vermeld dat schoolleiders en leerkrachten in veel scholen nog te weinig betrokken zijn op elkaars activiteiten. Het schoolbeleid is volgens de informanten nog onvoldoende gericht op de interactie tussen het onderwijskundig luik en het onderwijsorganisatorisch luik. De leerkrachten en schoolleiders zijn elk nog te zeer op hun eigen terreinen actief. Men denkt – zo geven de informanten aan - nog te veel in vakkenscheiding en is zich nog te weinig bewust van de mogelijkheden om problemen teamgericht op te lossen. Door die beperkte betrokkenheid worden volgens een informant belangrijke kansen tot interne professionalisering en netwerking gemist. Dat kan ook wegen op de kwaliteit van zelfevaluaties.

Een gebrek aan deskundigheid en opvolging. Verschillende informanten benadrukken dat directies en het lerarenkorps niet zijn opgeleid om aan zelfevaluatie te doen. Het ontbreekt scholen volgens hen aan deskundigheid om aan correcte, valide en betrouwbare gegevensverzameling en -verwerking te doen. Scholen zijn zich volgens een informant bijvoorbeeld nog te weinig bewust van de valkuil van het blijven ronddraaien in percepties. Er is sprake van een gebrek aan systematiek, bijvoorbeeld in het maken van een bewuste keuze tussen bestaande zelfevaluatie-instrumenten. Ook de deskundigheid om het verloop van zelfevaluatieprocessen en de resultaten van zelfevaluaties systematisch op te volgen, moet nog sterk groeien. Scholen zijn vaak wel bereid om problemen te lokaliseren en op te lijsten. Sommigen werken ook wel procedures uit om bij te sturen maar ondernemen onvoldoende actie om deze bijsturingen op te volgen en als uitgangspunt te gebruiken voor de volgende cyclische stap. Er zijn volgens de informanten nog te veel scholen die met de resultaten van een zelfevaluatie niets doen. Scholen zouden er moeilijk in slagen om veranderingsprocessen op gang te brengen of en te beheersen.

In teveel scholen zou er bovendien sprake zijn van een zwakke of een zeer restrictieve leiding.

> *'In sommige scholen moet je volgens mij niet aan zelfevaluatie beginnen. Daar moet je proberen een oplossing te zoeken, voor het probleem van de directeur. Er moet eerst een andere directeur komen, voor dat je daar iets kunt doen.'*

De informanten stellen nog steeds vast dat het niet de meest geschikte mensen zijn die voor een directiefunctie kandideren en/of gekozen worden. Daarbij worden deze mensen vaak niet opgeleid voor een leidinggevende functie. Dat maakt dat sterke ankerfiguren om het kwaliteitszorgbeleid in scholen te dragen volgens de informanten nog maar nauwelijks aanwezig zijn. Dat neemt niet weg dat sommige mensen zich ondanks een gebrek aan opleiding wel ontpoppen tot degelijke leiders, zij het met het nodige 'vallen en opstaan'.

Een gebrek aan realisme. Begeleiders van zelfevaluatieprocessen geven aan dat scholen hun behoefte aan vorming en begeleiding vaak onderschatten. *'Soms beseft men zelfs niet dat men ondersteuning nodig heeft omdat men onvoldoende kennis heeft van alles wat met zelfevaluatie te maken heeft'*. Het gebrek aan deskundigheid uit zich dus ook in

een tekort aan realisme. Scholen schatten hun eigen situatie en de impact van de zelfevaluatie volgens de informanten vaak niet goed in. Vlaamse scholen pakken zelfevaluaties dan ook dikwijls te ambitieus aan. Dat geldt zowel voor de omvang als de aard van de inhouden die bevraagd worden. Er worden nogal eens te veel en/of te gevoelige zelfevaluatiethema's vooropgesteld. In die zin hebben veel scholen volgens de informanten geen realistisch beeld van wat hun mogelijkheden zijn. Men geeft aan dat zelfevaluaties in sommige scholen moeilijk verlopen omdat de schoolleiding of het kernteam de schoolcultuur met de bijhorende weerstanden ten opzichte van (zelf)evaluatie verkeerd inschatten.

Een gebrek aan middelen. De problematiek van het gebrek aan middelen wordt vanuit verscheidene invalshoeken aangekaart. Elk van de volgende invalshoeken kan volgens de informanten mee helpen verklaren waarom scholen met goede intenties er toch niet in slagen om tot kwaliteitsvolle zelfevaluaties te komen.

Leerkrachten zijn in de eerste plaats begaan met de eigen kerntaken en zouden minder geïnteresseerd zijn in problemen op schoolniveau. Een informant schrijft dit (onder meer) toe aan het feit dat een lerarenopdracht nog steeds bestaat uit een lesopdracht waarop de leerkracht wordt afgerekend. Toch wordt er ook steeds meer een beroep gedaan op ruimere vaardigheden van de leraar. Het probleem is dat aandacht voor deze nieuwe uitdagingen ten koste kan gaan van de kwaliteit van de oorspronkelijke kerntaak.

Herhaaldelijk wordt ook het belang van coördinatie en opvolging in de verf gezet om tot kwaliteitsvolle zelfevaluaties te komen. De vraag voor veel scholen is echter hoe de tijdsbesteding door personeelsleden die een zelfevaluatie coördineren, verrekend kan worden. Omdat dergelijke taken vaak bovenop een bestaand takenpakket terecht komen, houden veel teamleden de boot af. Uit angst voor de planlast die met zelfevaluaties gepaard gaat, is men volgens de informanten zeer terughoudend om zich te engageren. Als van scholen bijkomende activiteiten verwacht worden, zou men hen logischerwijze met de middelen daartoe moeten uitrusten.

> *'Alle scholen die niet over een kwaliteitscoördinator beschikken, kunnen te weinig tijd maken voor zelfevaluatie. En er zijn heel weinig scholen die een kwaliteitscoördinator hebben. Of iemand die tijd kan investeren, om in kwaliteitszorg te steken. Terwijl iedereen weet, dat het een behoorlijke opgave is om gegevens te verzamelen, te analyseren, procedures duidelijk uit te tekenen. Dat zijn witte raven…'*

Het uitvoeren van een zelfevaluatie houdt volgens een informant ook in dat men verder kijkt dan het dagelijkse. Een aantal scholen kampt echter met een snel wijzigende personeelsbezetting of met het aantrekken van een problematisch leerlingenpubliek. Deze scholen focussen in de eerste plaats op de dagelijkse leerlingen- en daarmee samenhangende personeelsproblemen. In dergelijke moeilijke onderwijssituaties kan men volgens een aantal informanten niet van scholen verwachten dat ze zonder extra noemenswaardige middelen tot kwaliteitsvolle zelfevaluaties komen. Zonder deze middelen is er volgens hen (zeker in kleine scholen) te weinig tijd beschikbaar en zijn er te veel ernstige pro-

blemen en administratieve of logistieke beslommeringen om de voordelen van zelfevaluaties aan den lijve te kunnen ondervinden.

3.4 (De doeltreffendheid van de) beschikbare ondersteuning

Komen tot kwaliteitsvolle zelfevaluaties vergt specifieke competenties. Informanten verwijzen naar uiteenlopende elementen, gaande van het belang van voldoende technische knowhow voor het organiseren van een gegevensverzameling tot kennis en kunde inzake veranderingsmanagement. Scholen moeten volgens de informanten bewaken dat deze competenties voldoende in huis zijn. Een voorwaarde voor kwaliteitsvolle zelfevaluaties is volgens de informanten dan ook dat de noodzakelijke expertise in de school aanwezig is en/of door een beroep te doen op externe ondersteuning in de school wordt binnengehaald. Pas als dat aanbod voldoende breed en toegankelijk is, kan een inventarisatie van de eigen sterktes en zwaktes in het uitvoeren van zelfevaluaties een gepast antwoord vinden. Wanneer scholen naar begeleiding op zoek gaan en van een passende ondersteuningsvorm kunnen genieten, dan lukt het de meeste scholen volgens de informanten wel om een zelfevaluatie succesvol uit te voeren.

Verschillende informanten geven aan dat het huidige belang van een goede ondersteuning van zelfevaluatieprocessen door velen onderschat wordt. Dat geldt zowel voor beleidsmakers als voor de scholen zelf. Bijna alle scholen zouden immers ondersteuning nodig hebben.

> *'Als je werkelijk in je onderwijs, je schoolbeleidsplan zelfevaluatie een prominente plaats wil geven, denk ik dat het belangrijk is dat je in je school iemand zoekt die zich daar professioneel mee wil bezighouden, die zich daar ook voor schoolt, en dat niet overlaat aan de plaatselijke secretaris. Je moet toch wel wat meer bagage hebben.'*

Scholen hebben volgens de informanten momenteel nog onvoldoende knowhow in huis om de klus alleen te klaren. En zelfs wanneer een school al wel heel wat ervaring bezit, is het volgens de informanten wenselijk om een 'kritische vriend' bij het zelfevaluatieproces te betrekken. De informanten beschrijven concrete indicaties voor de heilzame impact van ondersteuning. Zo zou er in de meeste scholen bijvoorbeeld pas onder impuls van nascholing en/of begeleiding aandacht besteed worden aan het informeren en sensibiliseren van de betrokkenen. Bovendien wordt gerapporteerd dat sommige zaken gemakkelijker bespreekbaar worden wanneer een externe de resultaten in de school komt toelichten. Er kunnen dan prikkels worden gegeven die men van een collega of een directeur niet zou nemen, maar die de zelfreflectie of een veranderingsproces wel kunnen stimuleren. Het zou bovendien wenselijk zijn dat de persoon die de resultaten komt bespreken ook de verdere begeleiding op zich kan nemen, omdat het implementatieproces dan minder vlug stilvalt.

Uit het bovenstaande blijkt dat de informanten het belang van ondersteuning sterk in de verf zetten. Daarover bestaat eensgezindheid. Minder overeenstemming is er bij het beantwoorden van de vraag of scholen momenteel voldoende ondersteuning kunnen ge-

nieten bij het uitvoeren van zelfevaluaties? Sommigen benadrukken dat de ondersteuning nu al voldoende uitgebreid is terwijl anderen belangrijke tekorten rapporteren.

De ondersteuning voor het uitvoeren van zelfevaluatieprocessen is volgens een aantal informanten voldoende divers. Universiteiten bieden op basis van wetenschappelijk onderzoek concrete zelfevaluatiemodellen aan en zorgen voor begeleiding bij het gebruik van deze zelfevaluatie-instrumenten. De pedagogische begeleidingsdiensten van de koepels fungeren als stimulator en procesbegeleiders. Zij kunnen hulp bieden bij de introductie en opvolging van zelfevaluaties, kunnen scholen samen brengen en goede voorbeelden verspreiden. Indien nodig - en dat geldt in het bijzonder voor de meer beroepsgerichte opleidingen - kan men in bedrijven een kritische vriend vinden. Bovendien komt het thema 'zelfevaluatie' ook voor in nascholingsprogramma's. In deze logica is er dus wel voldoende ondersteuning beschikbaar.

Anderen delen de bovenstaande mening niet. Een aantal informanten geeft aan dat de ondersteuning van scholen in veel opzichten tekort schiet. *'Versnipperd, ondoorzichtig, soms van een heel lage kwaliteit.'* Volgens hen zijn de netwerken van scholen die zelfevaluatie-instrumenten gebruiken te beperkt, zijn er te weinig systeembegeleiders, zijn universiteiten nog te weinig bezig met zelfevaluatie en is de begeleiding door commerciële instanties te duur. Scholen zouden bovendien te weinig zicht hebben op de geschikte instrumenten. Momenteel is de ondersteuning volgens deze informanten te zeer beperkt tot incidentele initiatieven en tot de pedagogische begeleidingsdiensten. Het luidt dat er nog meer nood is aan professionele ondersteuning van scholen.

Waar de informanten het wel in grote lijnen over eens zijn is dat de ondersteuning op het einde van de rit niet doeltreffend genoeg is. Volgens de informanten die vinden dat er in principe voldoende ondersteuningsmogelijkheden zijn, is dit grotendeels toe te schrijven aan het feit dat scholen het beschikbare aanbod onvoldoende of onverstandig gebruiken. Men geeft aan dat dit onder meer komt doordat de meeste scholen niet bewust bezig zijn met het verwerven van inzicht in de eigen deskundigheid. Bijgevolg is het eerder zeldzaam dat de school(leiding) een duidelijk geformuleerde vraag naar ondersteuning stelt. Scholen komen er - volgens de inbreng van de informanten - niet toe om verschillende deelproblemen aan elkaar te linken tot één geheel en om daarrond de juiste vragen te formuleren om te gaan shoppen op de vraaggestuurde nascholingsmarkt.

> *'Waar ontstaat het probleem? Er is altijd één knelpunt dat er altijd blijft boven hangen: blijkbaar missen scholen het mechanisme, de vaardigheid of de handigheid om de goede vragen de stellen voor zichzelf, naar de buitenwereld. Want het is een vraaggestuurd systeem van ondersteuning. En dan weer niet die goede scholen, maar die scholen die ook met zelfevaluatie problemen hebben, die met alles wat te maken heeft met verandering miserie hebben, dat zijn ook de scholen die niet in staat zijn om de goede vragen te stellen.'*

Het maakt volgens de informanten niet zoveel uit welke instantie de ondersteuning in het uitvoeren van zelfevaluaties verzorgt, zolang deze instanties maar aan de volgende voorwaarden voldoen: ervaringsgericht en interactief kunnen werken; bereid zijn om ook in te staan voor ondersteuning tijdens de implementatie; kunnen werken vanuit een inbedding

in de lokale expertise en cultuur en op geen enkele manier gerelateerd zijn aan externe evaluatoren (cf. het verantwoordingsperspectief). Deze voorwaarden zouden evenwel onvoldoende gerealiseerd zijn. Er wordt bijvoorbeeld opgemerkt dat ook ondersteuningsdiensten tot op heden nog onvoldoende knowhow in huis hebben om tot doeltreffende ondersteuning van zelfevaluaties te komen. Men verklaart dit door het feit dat zij vanuit hun verleden vaak meer gericht zijn op onderwijsleerprocessen dan op schoolontwikkelingsprocessen. De meeste begeleiders hebben volgens een aantal informanten dan weer weinig tot geen ervaring met het uitoefenen van leidinggevende functies in scholen. Er wordt daarom ook aangegeven dat schoolondersteuners op hun beurt gevormd moeten worden om scholen te begeleiden bij het uitvoeren van zelfevaluaties.

4 ALGEMENE CONCLUSIES

Hoewel de informanten aangeven dat zelfevaluatie in wezen geen nieuw fenomeen is, juichen ze de hedendaagse aandacht ervoor in scholen sterk toe.

> *'Het is natuurlijk sinds een paar jaar de grote mode, het is ook zo gelanceerd, terecht. Want het gaat er over dat mensen gewoon natrekken wat ze doen. Ik denk dat dat al zo oud is als de straat. Het verschil is, denk ik, het belang van systematiek, van instrumenten en van totaliteit van de zaak, het in een groot perspectief zetten.'*

De verschillen met voorheen komen tot uiting in de invulling die de informanten geven aan het begrip 'zelfevaluatie'. Op basis van de inbreng van de informanten omschreven we zelfevaluatie als een '*een proces waarbij een school, op eigen initiatief en vanuit een overkoepelend kwaliteitszorgconcept, zelf op een systematische manier aspecten van het eigen functioneren beschrijft en beoordeelt met als doel (indien nodig) te komen tot concrete verbeterprocessen*'. Dit is een omschrijving die zeer sterk aanleunt bij andere omschrijvingen van het concept (Van Petegem, 1998; Devos & Verhoeven, 2000). In een andere formulering en deels andere woordkeuzes komen immers dezelfde elementen aan bod. De informanten geven wel aan dat de feitelijke invullingen van het concept zeer verscheiden kunnen zijn. Afhankelijk van de mate waarin de bestanddelen gerealiseerd zijn en van de concrete invulling ervan, komt men tot verschillende activiteiten die allen onder de noemer 'zelfevaluatie' vallen.

Het vervolg van deze paragraaf schetst de conclusies bij het voorgaande en integreert daarin een aantal lessen voor het vervolgonderzoek. Het geheel is gestructureerd op basis van de oriënterende onderzoeksvragen die bij de start van deze voorbereidende studie vooropgesteld werden.

4.1 Indicatoren voor kwaliteitsvolle zelfevaluatie

We stelden bij de start dat een (voor Vlaanderen) algemeen aanvaarde set van criteria om de kwaliteit van een zelfevaluatie te beschrijven niet voorhanden was. Het achterha-

len van indicatoren om uitspraken te doen over de kwaliteit van zelfevaluaties was dan ook een belangrijk doel van deze voorbereidende studie.

De informanten geven eerst aan dat scholen zeer uiteenlopende verwachtingen kunnen koesteren met het aangaan van een zelfevaluatieproces. Het beoordelen van de kwaliteit van dat proces zou dus op basis van schooleigen indicatoren moeten plaatsvinden. Toch geven de informanten ook aan dat er eveneens schooloverstijgende indicatoren vooropgesteld kunnen worden om de kwaliteit van zelfevaluatie te boordelen. Deze hebben dan betrekking op de resultaten van de zelfevaluatie. De mate waarin op school strategieën geïmplementeerd worden om verbeterbare elementen te verbeteren of om te borgen wat goed gaat, blijkt een algemeen aanvaarde indicator voor de kwaliteit van zelfevaluaties. Om als kwaliteitsvol te kunnen worden bestempeld, behoort zelfevaluatie te resulteren in concrete acties. Dé verwachting is dat zelfevaluatie aanleiding geeft tot acties die de schoolontwikkeling bevorderen. Het eindresultaat van een zelfevaluatie moet volgens de informanten dan ook bestaan uit actiepunten die voldoende specifiek, meetbaar, acceptabel, realistisch en tijdsgebonden zijn.

Naast het belang van de actiepunten werden door de informanten ook een aantal concrete resultaten naar voren geschoven. Het gaat om resultaten waarover bij de informanten een grote mate van consensus bestaat dat het om waardevolle effecten gaat. In het vervolgonderzoek zullen deze concrete resultaten (zoals ‘het versterken van de samenhang in de school’) mee opgenomen worden als mogelijke indicatoren voor het kwaliteitsvolle karakter van zelfevaluaties.

4.2 Hoe kwaliteitsvol zijn zelfevaluatie in Vlaamse scholen?

Als we vertrekken van de veronderstelling dat kwaliteitsvolle zelfevaluatie resulteert in goede actiepunten, dan is de Vlaamse stand van zaken inzake zelfevaluaties volgens de informanten (eerder) negatief. Het is evenwel moeilijk om veralgemenende uitspraken te doen. Er zijn volgens de informanten grote verschillen in de kwaliteit van zelfevaluaties. De vraag naar de kwaliteit van zelfevaluatie in Vlaanderen behoeft dus een genuanceerd antwoord.

Vlaanderen heeft volgens de informanten in de eerste plaats een aantal troeven. De onderwijskwaliteit is er volgens de bevraagde deskundigen in het algemeen goed en er is een grote bezorgdheid om kwaliteitsvolle groei. Het belang dat aan zelfevaluatie gehecht wordt is groeiend. Hoewel de intenties om aan zelfevaluatie te doen steeds vaker oprecht aanwezig zijn, laat de feitelijke uitwerking echter nogal eens te wensen over. De manier waarop het proces vormgegeven wordt zou het realiseren van deze intenties hypothekeren. Er zijn met andere woorden nog steeds verscheidene zwakke punten in Vlaamse scholen die de kwaliteit van zelfevaluaties volgens de informanten in de weg staan. Deze zijn het gebrek aan visie en doelgerichtheid, aan openheid, aan betrokkenheid, aan deskundigheid en opvolging, aan realisme en het gebrek aan middelen. Scholen hebben bovendien – volgens de informanten vaak zonder het zelf te beseffen - een grote behoefte aan ondersteuning. Die is volgens de bevoorrechte getuigen nog ontoereikend omdat

scholen hun weg nog niet goed vinden in het bestaande aanbod maar ook omdat het bestaande aanbod verdere professionalisering behoeft en meer dient te focussen op de begeleiding van scholen op systeemniveau.

Al de elementen die in de volgende paragraaf als voorspeller van kwaliteitsvolle zelfevaluatie naar voor worden geschoven, kunnen beschouwd worden als een ideaalbeeld waarnaar de culturele en organisatorische kenmerken van scholen moeten evolueren. Het is volgens de informanten duidelijk dat de weg voor scholen, hun begeleiders en hun overheid nog lang is.

4.3 Hoe kunnen de verschillen in kwaliteit verklaard worden?

De eensgezindheid tussen informanten omtrent de relevantie van verschillende voorspellers van kwaliteitsvolle zelfevaluaties blijkt groot. Hoewel de verschillende informanten andere klemtonen leggen, lijken ze het onderling in grote mate eens over welke kenmerken van het zelfevaluatieproces bepalend zijn om tot kwaliteitsvolle zelfevaluaties te komen. Die kenmerken hebben betrekking op een aantal randvoorwaarden zoals de beschikbaarheid van voldoende middelen en ondersteuning. Daarnaast speelt ook de schoolcultuur volgens de informanten een bepalende rol in de manier waarop zelfevaluatieprocessen verlopen. Er worden verschillende waarden en normen naar voren geschoven die kwaliteitsverschillen in zelfevaluaties kunnen verklaren. Daarnaast is er een sterke nadruk op aspecten die men samenvat onder de noemer van ‘beleidsvoerend vermogen’ (Sleegers, 1991).

Uit de batterij elementen die volgens de informanten bijdragen tot de kwaliteit van zelfevaluatie blijkt doelgerichtheid cruciaal. Het antwoord op de vraag waarom een zelfevaluatie uitgevoerd wordt, dient door de betrokkenen gedragen te zijn en verankerd in een ruimer kader. Daarnaast moeten de voorbereiding en het verloop van de zelfevaluatie doordacht gebeuren, moeten de gepaste mensen bij de zelfevaluatie betrokken worden en behoort de schoolcontext rijp te zijn voor een zelfevaluatie. Binnen het verloop van de zelfevaluatie wordt voornamelijk het belang van communicatie, samenwerking, reflectie, veranderingsbereidheid, realiteitszin en systematiek benadrukt. De eerste vijf criteria brengen ons ertoe zelfevaluatie te beschouwen als een ‘beleidsdaad’; systematiek brengt ons ertoe de zelfevaluatie te bekijken als een ‘onderzoeksdaad’. We lichten beide conclusies kort toe.

Zelfevaluatie als *beleidsdaad*. De antwoorden van de informanten lijken er op te wijzen dat bij het uitvoeren van een zelfevaluatie dezelfde kenmerken van het schoolfunctioneren een rol spelen als bij het voeren van andere vormen van schoolbeleid zoals personeelsbeleid of het opzetten van een systeem voor leerlingenbegeleiding. Zelfevaluaties kunnen in die zin beschouwd worden als het uitvoeren van een beleidsdaad en vergen naast een aantal specifieke vaardigheden in de eerste plaats een voldoende groot beleidsvoerend vermogen van scholen. De informanten verwijzen herhaaldelijk naar het belang van effectieve beleidsvoering om tot kwaliteitsvolle zelfevaluaties te komen. De volgende dragers van beleidsvoerend vermogen die door Van Petegem et al. (2005b) naar

voren worden geschoven blijken sterk aan te leunen bij de inbreng van de informanten: doeltreffende communicatie, ondersteunende professionele en persoonlijke relaties, gedeeld leiderschap, gezamenlijke doelgerichtheid, responsief vermogen, innovatief vermogen, geïntegreerd beleid en tenslotte reflectief vermogen. Naarmate scholen meer aan deze dragers tegemoet komen, is de effectiviteit van het beleid dat scholen voeren groter. De inbreng van de informanten brengt ons bij de hypothese dat hetzelfde geldt voor het uitvoeren van zelfevaluaties.

Zelfevaluatie kan daarnaast ook als *onderzoeksdaad* benaderd worden. De informanten benadrukken dat in een zelfevaluatie het niveau van toevallig opgedane indrukken moet overstegen worden. Dat vergt een systematische aanpak van het gebeuren. Enkel voldoende objectieve en zorgvuldig verzamelde gegevens zijn bruikbaar. Scholen moeten daartoe in voldoende mate inzicht hebben in het uitvoeren van sociaalwetenschappelijk onderzoek.

5 AFRONDEND

Met deze Delphi-studie wilden we de verschillende argumenten, ervaringen en opinies van bevoorrechte getuigen in kaart te brengen. Met de beschrijving van de bovenstaande resultaten werd daartoe een exploratieve aanzet gegeven. Dit is een vertrekpunt voor verdere studies. Daarbij dient de stap van percepties naar empirische evidentie te worden gezet. De stap van de genoemde voorspellers van kwaliteitsvolle zelfevaluatie naar mogelijke criteria voor een meta-evaluatie lijkt overigens logisch. In vervolgonderzoek zal de link tussen de (f)actoren die er volgens de informanten toe doen enerzijds en de feitelijke en gepercipieerde kwaliteit van zelfevaluaties anderzijds verkend worden. Op dit moment is het immers onduidelijk in welke mate scholen beschikken over voldoende beleidsvoerend vermogen en onderzoeksvaardigheden om strategieën voor kwaliteitsvolle zelfevaluatie uit te werken en toe te passen. Ook de vraag naar de verschillen die er op dit moment bestaan tussen scholen is tot op heden onbeantwoord. Het is veelbelovend om het hoger beschreven kader met daarin de criteria die de informanten aanleverden, te leggen op de manier waarop scholen zelfevaluaties uitvoeren. Op die manier kan onderzocht worden in welke mate de opinies van Vlaamse deskundigen stroken met de realiteit in Vlaamse scholen.

INVALSHOEK 2: LITERATUURSTUDIE

De tweede invalshoek om het theoretisch kader vorm te geven was logischerwijze het uitvoeren van een literatuurstudie. Ook de literatuurstudie moest leiden tot een meer diepgaand inzicht in de bestudeerde materie. We wilden in kaart te brengen wat er over (het beoordelen van) de kwaliteit van zelfevaluatie al bekend en nog niet bekend is.

Hoewel tijdsintensief zullen we over deze literatuurstudie bondig rapporteren. We lichten in deze paragraaf voornamelijk toe op welke manier we bij de literatuurstudie tewerk

zijn gegaan en op welke manier de systematische consultatie van bronnen nagestreefd werd. De focus ligt daartoe eerst op het proces van de literatuurstudie. Daarna wordt ingegaan op de resultaten van dit proces. De antwoorden die het literatuuronderzoek opleverde, worden echter niet als dusdanig in dit deel beschreven. Daarvoor verwijzen we naar deel 4 waarin het theoretisch kader van deze studie gepresenteerd wordt.

1 HET PROCES: DE ZOEKTOCHT NAAR EN SELECTIE VAN RELEVANTE BRONNEN

De centrale richtvragen die aan de basis liggen van de literatuurstudie waren de onderzoeksvragen zoals deze na de probleemstelling geformuleerd werden. Zij vormden als het ware het structurerend kader voor het zoekplan. De invalshoek was dus in de eerste plaats 'het uitvoeren van zelfevaluaties door scholen'. Daarnaast worden in de onderzoeksvragen ook concepten als 'kwaliteit', 'het functioneren van scholen' en 'ondersteuning' centraal gesteld. Om een antwoord te kunnen geven op de onderzoeksvragen moest eerst duidelijk zijn wat er onder deze concepten verstaan wordt, hoe ze geoperationaliseerd kunnen worden en wat hun onderlinge verhoudingen zijn. Deze vragen waren het uitgangspunt van de literatuurstudie.

In de trefwoorden die gebruikt werden om tot relevante literatuur te komen, kan een onderscheid gemaakt worden tussen kerntrefwoorden en verruimende trefwoorden. Tot de eerste rekenen we: zelfevaluatie, meta-evaluatie, interne evaluatie, zelfanalyse, doorlichtingen, schooleffectiviteit, schoolontwikkeling, beleidsvoerend vermogen, lerende organisatie, (onderwijs)kwaliteit, reflectie, ondersteuning. Deze kernwoorden – en hun Engelse tegenhangers - vormden het startpunt van het literatuuronderzoek. De relevantie van verruimende trefwoorden bleek als het ware pas gaandeweg en ze werden dan ook pas later in het zoekproces aangewend. Tot de belangrijkste voorbeelden van verruimende trefwoorden rekenen we bijvoorbeeld: onderwijsvernieuwing, veranderingscapaciteit, schoolorganisatie, (school)cultuur, professionele leergemeenschap, leiderschap, onderwijsorganisatie, organisatie-effectiviteit, onderwijsbegeleiding, (collectief) leren, (sociaal)constructivisme, 'enkele lus'- en 'dubbele lus'-leren, en kenniscreatie.

Aan de hand van deze trefwoorden werden de belangrijkste onderwijsgerichte databanken zowel op titel als op onderwerp doorzocht. Startpunten waren onder meer ADIOV (Automatisch Documentatie- en Informatiesysteem voor het Onderwijs in Vlaanderen), ERIC (Educational Resource Information Center), Educati*on-line*, Academic Search Elite maar ook internetzoekmachines. Eenmaal relevante literatuur op het spoor werd gekomen, zijn de databanken op naam van auteurs geraadpleegd. Daarnaast is literatuur gezocht naar aanleiding van de literatuurverwijzingen in reeds gevonden tijdschriftartikelen en boeken.

2 HET RESULTAAT: EEN VARIATIE AAN GERAADPLEEGDE BRONNEN

De combinatie van verschillende databanken heeft tot een grote variatie in geraadpleegde werken geleid: abstracte en praktijkgerichte (hand)boeken; conceptuele, wetenschappelijke en op ervaring gebaseerde tijdschriftartikelen, 'grijze literatuur' met onderzoeksrapporten, congresbijdragen, toespraken en beleidsintenties.

De raadpleging van deze literatuur heeft voor deze studie tot een aantal belangrijke resultaten geleid. De meest belangrijke bevinding was wellicht dat over onderzoek naar de kwaliteit van zelfevaluaties (zoals door ons gedefinieerd) tot op heden slechts in beperkte mate gerapporteerd wordt. Blok, Sleegers & Karsten (2005, p.3) stellen niet voor niets dat het om een nog betrekkelijk onbekend fenomeen gaat. De raakpunten die er toch zijn, worden in het volgende deel in het theoretisch kader geïntegreerd.

3 AFRONDEND: VERFIJNDE ONDERZOEKSVRAGEN

Eén van de doelen van de Delphi-studie en het literatuuronderzoek was te komen tot beargumenteerde keuzes voor het uitwerken van een theoretisch kader voor het uitvoeren van de volgende onderzoeksactiviteiten. De bovenstaande bevindingen hebben dan ook een duidelijke impact op het theoretisch kader dat in deel 4 uitgewerkt zal worden. Dat uit zich bijvoorbeeld in de meer precieze omschrijving van de voorliggende onderzoeksvragen. Vanuit het hiervoor behandelde denkkader komen we immers tot vier onderzoeksvragen voor het vervolgonderzoek. Deze onderzoeksvragen zijn een uitwerking van onze oriënterende onderzoeksvragen. Opnieuw zijn ze van verschillende aard: we onderscheiden ontwerp-, beschrijvings- en verklaringproblemen (Swanborn, 2000).

1. Aan de hand van welke indicatoren kan men de kwaliteit van een zelfevaluatie operationaliseren?

Het woord ‘kwaliteit’ wordt voortdurend in de mond genomen, ook in de context van het uitvoeren van zelfevaluaties. Het begrip is evenwel uiterst moeilijk in te vullen. Het is tot op heden bijvoorbeeld onduidelijk aan welke verwachtingen een zelfevaluatie dient te voldoen om als kwaliteitsvol te worden bestempeld. Met de aandacht voor deze eerste onderzoeksvraag willen we verdere stappen zetten om het begrip ‘kwaliteit van een zelfevaluatie’ te operationaliseren. De nadruk ligt hier dus op ontwerp.

2. Hoe kwaliteitsvol verlopen zelfevaluaties in Vlaamse scholen en wat zijn de verschillen tussen scholen terzake?

Over de wijze waarop zelfevaluatieprocessen in Vlaamse scholen verlopen, is momenteel weinig geweten. Door deze onderzoeksvraag naar voren te schuiven willen we in kaart brengen hoe zelfevaluaties momenteel plaatsvinden. De nadruk ligt op het beschrijven van de variatie in de bestaande zelfevaluatieprocessen en in de kwaliteit ervan.

3. In welke mate wordt het verloop van een zelfevaluatie bepaald door (a) kenmerken van het functioneren van de school en (b) de ondersteuning die scholen genieten?

Het antwoord op deze onderzoeksvraag moet verduidelijken of het uitvoeren van een zelfevaluatie al dan niet losstaat van de rest van het functioneren van scholen. Daarmee bedoelen we of het verloop van zelfevaluaties samenhangt met ruimere kenmerken van de school die in wezen los staan van de eigenlijke zelfevaluatie. Daartoe proberen we de verschillen in het verloop van zelfevaluatie te verklaren op basis van kenmerken van het ruimere functioneren van scholen. Daarnaast onderzoeken we de impact van ondersteuning.

4. In welke mate worden de resultaten van een zelfevaluatie bepaald door (a) het verloop van de zelfevaluatie en (b) de ondersteuning die scholen genieten?

Ook voor de variatie in de kwaliteit van zelfevaluaties willen we verklaringen vinden. Zoals uit de onderzoeksvraag blijkt, richten we ons voor mogelijke verklaringsgronden op elementen die betrekking hebben op het verloop van de zelfevaluatie enerzijds en de ondersteuning die scholen tijdens het zelfevaluatieproces genieten anderzijds.

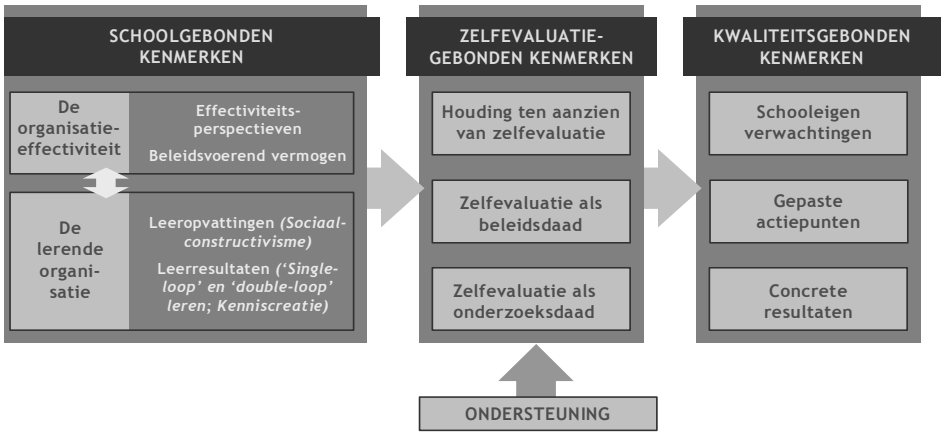
Deze onderzoeksvragen staan centraal in het uitwerken van de verdere onderzoeksactiviteiten. Deel 4 schetst een passend theoretisch kader om ze te onderzoeken. Dat theoretisch kader geeft een overzicht van literatuur (cf. invalshoek 2) en opinies (cf. invalshoek 1) die tot een goed begrip van de onderzochte verschijnselen bijdragen. Het brengt de voornaamste argumenten, onderzoeksresultaten en theorieën samen.

DEEL 4

ONTWERP EN OPERATIONALISATIE VAN HET THEORETISCH KADER

In dit deel bundelen we de inzichten uit de literatuurstudie en de Delphi-studie tot een kader dat het uitgangspunt zal vormen voor verdere gegevensverzamelingen en analyses. Op basis van de onderzoeksactiviteiten waarover hierboven gerapporteerd werd, inventariseren we de kenmerken waarvan verwacht wordt dat ze van invloed zijn op de kwaliteit van zelfevaluaties. Voor elk kenmerk bestaat in zekere mate (empirische) evidentie. Toch is het in deze fase van het onderzoek nog niet geweten welke de specifieke bijdrage is van elk kenmerk. Het in kaart brengen daarvan zal dan ook centraal staan in het verdere verloop van deze studie. Daartoe schetsen we eerst een kader dat het theoretische uitgangspunt zal zijn voor het beantwoorden van de vooropgestelde onderzoeksvragen. Dat kader wordt visueel samengevat in Figuur 3 en toegelicht in de volgende paragrafen.

Figuur 3: Theoretisch kader



Of een (geïnstrumentaliseerde) zelfevaluatie kwaliteitsvol[3] is in termen van haar resultaten, is volgens dit theoretisch kader zowel afhankelijk van kenmerken van het functioneren van de school in het algemeen als van kenmerken van het zelfevaluatieproces als dusdanig. Onze hypothese is dat het functioneren van de school in sterke mate bepalend is voor het verloop van de zelfevaluatieprocessen en de kwaliteit ervan. Daarnaast is er in het model ook een belangrijke rol weggelegd voor de ondersteuning die scholen krijgen. Er worden dus vier kernconcepten onderscheiden.

1. Om de relevante *schoolgebonden kenmerken* in kaart te brengen, doen we een beroep op twee invalshoeken: de organisatie-effectiviteit van de school en de mate waarin de school als een 'lerende organisatie' geboekstaafd kan worden. De organisatie-effectiviteit wordt in kaart gebracht aan de hand van de effectiviteitsperspectieven van Quinn en Rohrbaugh (1983) en de mate waarin scholen tegemoet komen aan de dragers van beleidsvoerend vermogen (Van Petegem e.a., 2005b). Of een school al dan niet als 'lerende organisatie' bestempeld kan worden, wordt in het model door middel van drie theoretische benaderingen geconcretiseerd: de sociaalconstructivistische visie op leren, 'single loop' en 'double loop' leren (Argyris & Schön, 1978) en het model van kenniscreatie (Nonaka & Takeuchi, 1995).

2. Ook de *zelfevaluatiegebonden kenmerken* worden door middel van complementaire benaderingen onderzocht. Een eerste variabele is de houding ten aanzien van zelfevaluatie. Verder introduceren we het concept 'zelfevaluatie als beleidsdaad'. Daarmee willen we aantonen dat de dragers van beleidsvoerend vermogen ook waardevol zijn om het verloop van zelfevaluaties te beschrijven. Ten derde worden onder de noemer 'zelfevaluatie als onderzoeksdaad' de eerder technische en onderzoeksgerichte kenmerken van het zelfevaluatieproces bestudeerd.

3. De derde factor in het model is de *ondersteuning* die scholen genieten. Zoals in het model kan afgelezen worden, verwachten we van de ondersteuning die scholen genieten een rechtstreeks effect op de kenmerken van de zelfevaluatie en daardoor op de kwaliteit ervan.

4. Tenslotte is er *'de kwaliteit van de zelfevaluatie'*. Finaal is dit de afhankelijke variabele in onze onderzoeksopzet. Het gaat hem hier om de mate waarin de zelfevaluatie in de vooropgestelde resultaten uitmondt. Ook dit concept behoeft verdere omschrijving en operationalisatie.

---

[3] In het door ons gehanteerde begrippenkader is een zelfevaluatie kwaliteitsvol als deze aan de vooropgestelde resultaatsverwachtingen voldoet. Tenzij expliciet vermeld verwijst kwaliteit dus altijd naar de resultaten van de zelfevaluatie. Een 'kwaliteitsvolle zelfevaluatie' is bijvoorbeeld iets anders als een 'kwaliteitsvol verloop van de zelfevaluatie'. In het laatste geval is er een verwijzing naar het verloop (het proces) van de zelfevaluatie. Het gaat dan om een zelfevaluatie die aan de procesverwachtingen voldoet. Zonder specificatie is kwaliteit systematisch resultaatgebonden.

In de volgende paragrafen werken we de vier kernconcepten in het kader voor meta-evaluatie verder uit. We willen daarbij benadrukken dat dit model niet het enige mogelijke kader is om naar (de kwaliteit van) zelfevaluaties te kijken. Het stelt specifieke aspecten en karakteristieken centraal. We willen met het model structuur aanbrengen in de manier waarop we naar de te bestuderen verschijnselen kijken, zonder te ambiëren een volledig beeld van de werkelijkheid weer te geven. Het model stelt ons in staat om een aantal voor deze studie richtinggevende onderzoekshypothesen voorop te stellen (zie pagina 114) en zal zowel een cruciale rol spelen in de operationalisering als bij het interpreteren van de onderzoeksresultaten. We integreren in dit deel overigens de beschrijving én de operationalisering van het theoretisch kader[4]. We geven daartoe in grijze kaders telkens aan hoe de theoretische concepten gemeten zullen worden in het survey-onderzoek.

1 META-EVALUATIE GECONTEXTUALISEERD: FOCUS OP SCHOOLGEBONDEN KENMERKEN

De manier waarop scholen vormgeven aan een zelfevaluatie en de kwaliteit van de zelfevaluatie die hiermee gepaard gaat, moeten gesitueerd en begrepen worden in de context waarin deze ontstaan. Die context is in wezen zeer ruim en dient daarom te worden afgebakend. Wij focussen op door de school beïnvloedbare aspecten van het schoolfunctioneren. De macro-beleidscontext wordt in het volgende dus niet opgenomen. Als invalshoeken hanteren we de concepten 'organisatie-effectiviteit' (zie 1.1) en 'de school als lerende organisatie' (zie 1.2).

1.1 De organisatie-effectiviteit van scholen

Organisatie-effectiviteit wordt aan de hand van twee verschillende modellen geoperationaliseerd. In het volgende verwijzen we eerst naar de effectiviteitsperspectieven van Quinn en Rohrbaugh (1983) en daarna naar het beleidsvoerend vermogen van scholen (Sleegers, 1991; Van Petegem e.a., 2005b).

1.1.1 De effectiviteitsperspectieven van Quinn en Rohrbaugh

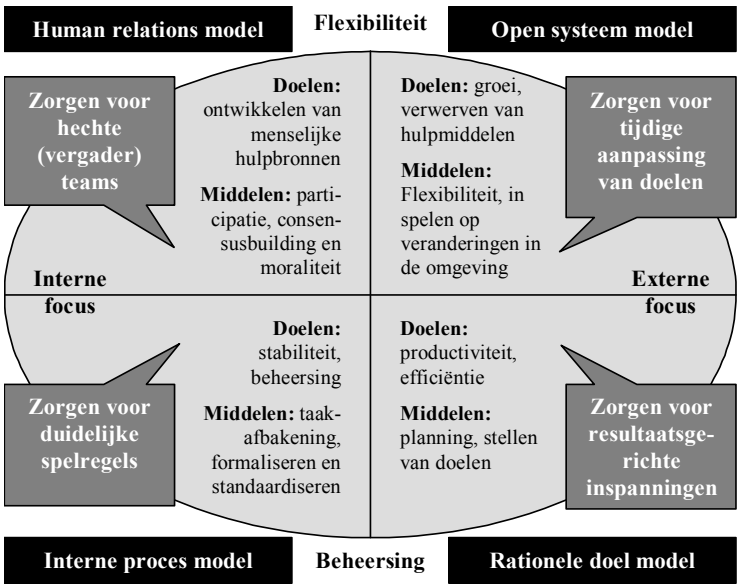
Het feit wat wij uitspraken willen doen over de organisatie-effectiviteit van scholen impliceert dat er sprake is van doelgerichtheid. Het effectiviteitsbegrip is immers gekoppeld aan de assumptie dat organisaties doelgericht zijn (Scheerens, 1997). Daarbij stelt zich onmiddellijk de vraag om welke doelen of welke soort doelen van de organisatie het gaat. De doelen die in scholen vooropgesteld worden, kunnen echter sterk uiteen lopen en ook de basis waarop de effectiviteit van organisaties beoordeeld dient te worden is dus niet uniform geldig. Om de variatie in doelgerichtheid en effectiviteitsbenaderingen in scho-

[4] Voor een overzicht van de gebruikte schalen verwijzen we naar bijlage 1.

len in kaart te brengen, doen we een beroep op de effectiviteitsperspectieven die door Quinn en Rohrbaugh (1983) worden onderscheiden. Hun kader is gebaseerd op multidimensionele schaaltechnieken waardoor in een lijst van indicatoren van organisatie-effectiviteit (Campbell, Bownas, Peterson & Dunnette, 1974) dimensies onderscheiden kunnen worden. Op die manier is het mogelijk een ordening aan te brengen in de criteria die doorgaans aangewend worden om de organisatie-effectiviteit (van scholen) te beschrijven en te beoordelen.

Quinn en Rohrbaugh (1983, p. 369) operationaliseren de effectiviteit van een kennissysteem aan de hand van vier effectiviteitsperspectieven die bepaald worden door twee onafhankelijke dimensies. De eerste dimensie richt zich op de organisatorische focus van de (school)organisatie en loopt van een interne focus op de mensen in de organisatie tot een externe focus op de organisatie zelf. De tweede dimensie representeert het contrast tussen enerzijds stabiliteit en controle en anderzijds flexibiliteit en verandering. De twee dimensies leiden tot vier kwadranten (zie Figuur 4). De 4 perspectieven zijn het 'Human relations model', het 'Open systeem model', het 'Interne proces model' en het 'Rationele doel model'.

Figuur 4: De effectiviteitsperspectieven van Quinn en Rohrbaugh



Het leerrijke van deze kwadranten is het uitgangspunt dat ze alle vier in minimale mate van toepassing moeten zijn om van een effectieve organisatie te kunnen spreken (Quinn & Rohrbaugh, 1983; Koster & Bouman, 1997; Scheerens, 1997; Quinn, Faerman, Thompson & McGrath, 2003). Om vlot te kunnen functioneren, moet een school meer bepaald *in voldoende mate* aan de vier hieronder opgesomde vereisten voldoen. De omschrijving 'in voldoende mate' is hierbij cruciaal vermits het verband met organisatie-effectiviteit niet lineair is. Quinn (1988) spreekt in dat opzicht van een negatieve zone en een positieve

zone. De negatieve zone omvat zowel het tekort als een teveel binnen een bepaald kwadrant. Een school heeft bijvoorbeeld duidelijke regels nodig, maar een teveel aan regels leidt tot bureaucratie en hypothekeert de goede werking van de organisatie (Stijnen, 1993; Vanhoof & Van Petegem, 2006b).

a. HET INTERNE PROCES MODEL: ZORGEN VOOR DUIDELIJKE SPELREGELS (INTERNE FOCUS – BEHEERSING)

Het beheersen van het functioneren van (school)organisaties vergt duidelijke spelregels. Om tot vlot werkende systemen te komen, moet de organisatie zodanig vorm worden gegeven dat er stabiliteit ontstaat. Het doel in dit kwadrant is dan ook beheersing en stabiliteit. Dat vergt strategieën als het duidelijk omschrijven en afbakenen van verantwoordelijkheden, zorgvuldige planning en opvolging van activiteiten door middel van het formaliseren en standaardiseren van overleg- en besluitvormingsprocessen. Ook aandacht voor de communicatie van informatie en doelen wordt gehanteerd om het doel te bereiken. De schoolleiding heeft in dit model een belangrijke rol inzake coördinatie en controle (Quinn et al., 2003).

> *De operationalisering van de effectiviteitsperspectieven is telkens gebaseerd op de 'Schoolcultuur Inventaris' (Maslowski, 2001). De vraag luidt in welke mate respondenten van mening zijn dat anderen in hun school een aantal waarden belangrijk vinden. Voor het interne proces model gaat het om de volgende waarden: stabiliteit, continuïteit, formele aanpak, rechtlijnigheid, controle, reglementering en zekerheid.*

b. HET HUMAN RELATIONS MODEL: ZORGEN VOOR HECHTE (VERGADER)TEAMS (INTERNE FOCUS – FLEXIBILITEIT)

Scholen bestaan bij uitstek uit menselijke bronnen en dat maakt het 'human relations model' bijzonder relevant. De uitdaging bestaat erin groepen mensen bij elkaar te brengen en te komen tot een harmonieus geheel. Het doel in dit kwadrant is de ontwikkeling van de menselijke hulpbronnen. Daartoe wil men openheid, collaboratie, loyaliteit, betrokkenheid en motivatie bij de teamleden bevorderen. Strategieën daartoe zijn teambuilding, participatie, consensusbuilding en aandacht voor moraliteit. De activiteiten van de schoolleiding sluiten in dit kwadrant aan bij de rol van stimulator, teambuilder, mentor en coach (Quinn et al., 2003).

> *Het 'human relations model' wordt gemeten aan de hand van de volgende waarden: engagement, behulpzaamheid, wederzijds vertrouwen, loyaliteit, wederzijds begrip, aangenaam klimaat en samenwerking.*

c. HET OPEN SYSTEEM MODEL: ZORGEN VOOR TIJDIGE AANPASSING VAN DOELEN (EXTERNE FOCUS – FLEXIBILITEIT)

Organisaties staan niet los van de omgeving waarin ze gelegen zijn. Scholen hebben bijvoorbeeld een maatschappelijke functie en worden publiek gesubsidieerd. Ze moeten hun werking vormgeven op een manier die door de maatschappelijke stakeholders geaccepteerd wordt. Om te kunnen groeien en om bijkomende hulpmiddelen te verwerven, is een gerichtheid op de omgeving dan ook een belangrijk doel. De middelen daartoe zijn flexibiliteit en de bereidheid om in te willen spelen op veranderingen in de omgeving. Respon-

siviteit en adaptatie zijn sleutelwoorden. Er is eerder sprake van inspiratie dan van controle. De bekommernis is als het ware als organisatie de goede dingen doen. De schoolleiding speelt hier de rol van onderhandelaar en bemiddelaar tussen de school en de omgeving (Quinn et al., 2003).

> *Het 'open systeem model' wordt gemeten aan de hand van de volgende waarden: Vernieuwingsgezindheid, flexibiliteit, diversiteit, op de omstandigheden inspelen, openheid naar nieuwe ontwikkelingen, veelzijdigheid en aanpassingsvermogen.*

d. HET RATIONELE DOEL MODEL: ZORGEN VOOR RESULTAATSGERICHTE INSPANNINGEN (EXTERNE FOCUS – BEHEERSING)

(School)organisaties moeten zo worden vormgegeven dat met een minimum aan middelen een zo goed mogelijk resultaat kan neergezet worden. In dit kwadrant zijn de doelen dan ook productiviteit en efficiëntie. Om het prestatie- en resultaatsgericht gedrag van leden van de school te bevorderen worden strategieën als het specificeren en kwantificeren van doelstellingen en resultaten vooropgesteld. Er is sprake van goed omschreven projecten met tussentijdse mijlpalen en een rationele opvolging. Eerder dan de goede dingen doen is de bekommernis hier de dingen goed doen. De schoolleiding speelt hier als het ware de rol van producent (Quinn et al., 2003).

> *Het 'rationele doel model' wordt gemeten aan de hand van de volgende waarden: Resultaatgerichtheid, effectiviteit, prestatiegerichtheid, efficiëntie, doelgerichtheid, succes nastreven en bekwaamheid.*

Naar het bovenstaande model wordt ook verwezen als het 'concurrerende waarden model' (cf. 'Competing Values Approach' (Quinn & Rohrbaugh, 1983, p. 369)) omdat de dimensies in het model op het eerste zicht tegenstrijdige boodschappen uitdragen. We willen bijvoorbeeld dat scholen controleerbare en stabiele organisaties zijn, maar verwachten tevens adaptiviteit en flexibiliteit. De verschillende kwadranten worden dan ook beter niet als op zich staand benaderd. Beter is ze te benaderen als een onderdeel van een ruimer construct, namelijk organisatie-effectiviteit. Om het handelen van mensen in scholen te begrijpen zijn de vier kwadranten - en de waarden die er achter schuilgaan - elk op zich noodzakelijk. Bovendien: de verschillende kenmerken van het model mogen elkaar dan wel tegenspreken, ze sluiten elkaar niet uit (Quinn & Rohrbaugh, 1983, p. 374). Organisatie-effectieve scholen zijn zowel flexibel als stabiel en hebben zowel een interne als externe focus. Er moet dus met andere woorden sprake zijn van een minimumprestatie op elke dimensie. Daarnaast komt het er voor scholen op aan een gepaste balans te vinden tussen de dimensies. Die balans kan niet eenduidig vast worden gelegd (Maslowski, 2001). In de analyse zullen wij daarom ook werken met clusters van scholen (schoolprofielen) om de link te leggen met het verloop en de kwaliteit van zelfevaluaties.

1.1.2 Het beleidsvoerend vermogen van scholen

Om het functioneren van scholen in kaart te brengen doen we ook een beroep op het 'beleidsvoerend vermogen' van scholen. Onze beschrijving van dat concept is gebaseerd op (Sleegers, 1991; Lagerweij & Voogt, 1997; Van Petegem e.a., 2005b, p. 32) en luidt:

Een school met een sterk beleidsvoerend vermogen wendt haar beschikbare beleidsruimte succesvol aan om te komen tot een voortdurend proces van behouden of veranderen van haar (beheersmatige en onderwijskundige) functioneren met als doel het verbeteren van haar onderwijskwaliteit en het bereiken van de haar opgelegde en eigen doelen. Het laatste deel van de omschrijving verwijst expliciet naar het aspect van doelbereiking. Daarbij wordt zowel verwezen naar het algemene doel 'het verbeteren van de onderwijskwaliteit' als naar meer specifieke doelen die aan scholen worden opgelegd (bijvoorbeeld het bereiken van de eindtermen of het vergroten van gelijke onderwijskansen) en die scholen voor zichzelf kunnen vooropstellen. Scholen met een sterk beleidsvoerend vermogen gebruiken hun beschikbare beleidsruimte maximaal om die doelen waar te maken. Zij richten hun functioneren zodanig in dat het bijdraagt tot doelbereiking (Van Petegem, Devos, Mahieu, Dang Kim & Warmoes, 2006).

a. HERKOMST VAN DE DRAGERS VAN BELEIDSVOEREND VERMOGEN

De effectiviteit van het beleid dat scholen voeren, wordt in sterke mate bepaald door de mate waarin scholen tegemoet komen aan een aantal dragers van beleidsvoerend vermogen. Op basis van onderzoek is één en ander bekend over de sleutelkenmerken en dragers van het beleidsvoerend vermogen van scholen. Tabel 3 inventariseert een aantal belangrijke bevindingen van Vlaams en internationaal onderzoek naar aspecten van de effectiviteit van *school*beleid. We hebben ernaar gestreefd om zowel inhoudelijk als methodologisch een gevarieerd beeld te schetsen. Hoewel in de verschillende onderzoeken ook andere relevante kenmerken (op micro-, meso- en macroniveau) gerapporteerd worden, blijken de volgende acht in de context van schoolbeleid het meest pertinent: doeltreffende communicatie(COM), ondersteuning en samenwerking (OND & SAM), gedeeld leiderschap (GED LS), gezamenlijke doelgerichtheid (GEZ DOEL), responsief vermogen (RESP), innovatief vermogen (INNO), geïntegreerd beleid (INT) en reflectief vermogen (REFLEC).

Tabel 3: Empirische evidentie voor de dragers van beleidsvoerend vermogen

Onderzoek	Focus van het onderzoek en onderzoeksmethodologie	Dragers
(Mortimore, Sammons, Stoll, Lewis & Ecob, 1988)	Survey-onderzoek in 49 lagere scholen om te achterhalen of er verschillen zijn in de effectiviteit van scholen nadat met de kenmerken van de instroom rekening werd gehouden.	GED LS, OND & SAM, RESP
(Levine & Lezotte, 1990)	Review van onderwijsonderzoek naar de effectiviteit van scholen. De focus ligt op het bestuderen van scholen die in onderzoek buitengewoon effectief bleken (cf. outlier-onderzoek).	COM, GEZ DOEL, GED LS, OND & SAM, RESP, REFLEC
(Staessens, 1991)	Kwalitatieve studie naar de professionele cultuur van Vlaamse basisscholen. 9 cases werden op basis interviews met de directeur en minstens de helft van alle leerkrachten onderzocht.	OND & SAM, GED LS, GEZ DOEL
(Sleegers, 1991)	Studie naar de benutting van beleidsruimte door Nederlandse secundaire scholen. Survey bij schoolleiding, schoolbestuur en leerkrachten.	OND & SAM, GED LS, RESP, INT

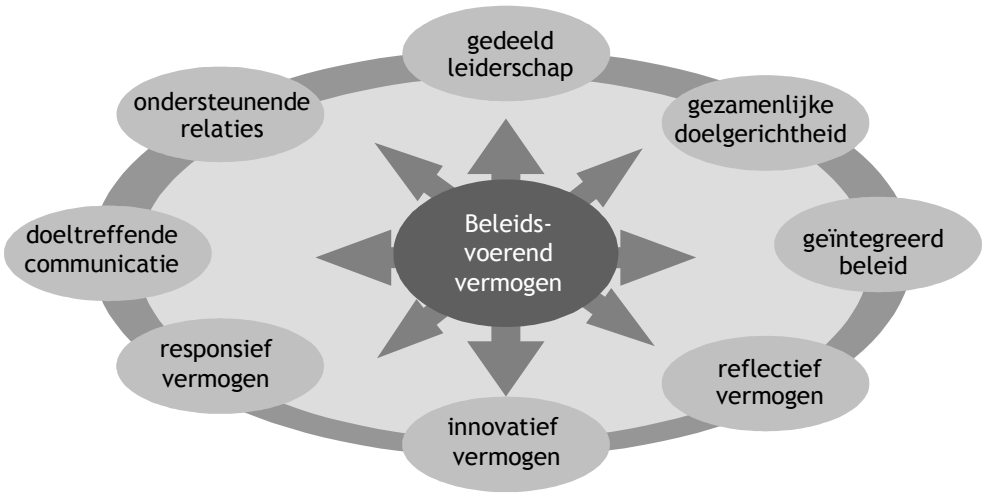
Onderzoek	Focus van het onderzoek en onderzoeksmethodologie	Dragers
(Van den Berg, Sleegers, Bakx & van der Eerden, 1994)	Literatuurstudie en kwalitatief onderzoek naar de kenmerken van innovatieve organisaties.	COM, OND & SAM, GED LS,
(Sammons, Hillman & Mortimore, 1995)	Literatuurreview naar de sleutelelementen in de effectiviteit van scholen (BaO en SO). Gebaseerd op de studie van 186 studies die werden uitgevoerd in de Verenigde Staten, Canada en het Verenigd Koninkrijk.	OND & SAM, GED LS, GEZ DOEL, RESP, INNO, REFLEC
(Hooge, 1998)	Studie naar de invloed van de interne organisatie en externe omgeving op de beleidsvoering in basisscholen. Gebaseerd op vergelijkende gevalstudies en surveys bij schoolleiders.	GED LS, RESP, INT
(McBeath, 1999)	Een opinieonderzoek naar de kenmerken van een goede school. Gegevensverzameling bij leerlingen, leerkrachten, ouders, schoolleiding, ondersteunend personeel en beleidsmakers.	COM, OND & SAM, GED LS, RESP
(Wikeley, Stoll & Lodge, 2002)	Gevalstudies in basis- en secundaire scholen naar succesfactoren bij het implementeren van vernieuwingen. Gegevensverzameling door middel van documentanalyse en schoolbezoeken bestaande uit observaties en interviews.	GEZ DOEL , GED LS ,INT
(Potter, Reynolds & Chapman, 2002)	Een Literatuurreview naar 'wat werkt?' in de ontwikkeling van scholen die in moeilijke omstandigheden functioneren of functioneerden.	GEZ DOEL, GED LS ,INT, OND & SAM, RESP, INNO
(Griffith, 2003)	Een onderzoek naar het effect van beleidsperspectieven van Quinn en Rohrbaugh (1983) op de prestaties van leerlingen basisonderwijs. Survey bij leerkrachten en leerlingen.	OND & SAM, GED LS, GEZ DOEL, RESP
(De Maeyer & Rymenans, 2004)	Effectiviteitsonderzoek in 95 Vlaamse scholen voor technisch en beroepssecundair onderwijs. Gegevensverzameling bij leerlingen, leerkrachten en schoolleiding. De hier gerapporteerde dragers hebben betrekking op resultaten voor wiskunde en begrijpend lezen.	OND & SAM, GED LS, REFLEC

De verschillende onderzoeken naar de sleutelelementen van de effectiviteit van scholen komen weliswaar vaak bij dezelfde factoren terecht (Devos e.a., 1999; Maslowski, 2001) maar de matige consistentie over de onderzoeken moet onderzoekers echter tot bescheidenheid aanmanen. Illustratief in dat opzicht is de meta-analyse van Bosker & Witziers (1996) waaruit geconcludeerd wordt dat over de hele lijn niet kan gesteld worden dat er empirische evidentie is voor het effect van leiderschap op de effectiviteit van (Europese) scholen. De hierboven aangehaalde evidentie is dus zeker niet absoluut. Het gaat eerder om indicaties dat de aangehaalde elementen de moeite waard zijn om ook in deze studie te worden opgenomen.

Figuur 5 brengt de dragers die wij uit het literatuuronderzoek meenemen visueel samen. De genoemde dragers van beleidsvoerend vermogen kunnen niet los van elkaar gezien worden. Ze zijn in hoge mate recursief in die zin dat ze ons inziens tegelijkertijd zowel 'doel op zich' zijn als 'middel om een doel te bereiken'. Er zijn dus duidelijke samenhangen tussen de dragers (Van Petegem e.a., 2006). Barnett, McCormick & Conners (2001) geven bijvoorbeeld aan dat de samenwerking tussen de leden van de school een belangrijke voorwaarde is om te komen tot een gezamenlijke doelgerichtheid. Dat geldt ook

voor het implementeren van vernieuwingen (Little, 1990a). Zulke voorbeelden zijn talrijk. Dit kan verklaren waarom de duidelijk theoretisch te onderscheiden dragers in bijvoorbeeld factoranalyses moeilijk als unieke factoren gerepresenteerd worden (Gemmeke & van der Vegt, 1997). Empirische evidentie wijst echter op het bestaan van een concept 'beleidsvoerend vermogen van scholen' dat in sterke mate de variantie in de verschillende dragers verklaart (Van Petegem e.a., 2006; Van Petegem, Vanhoof & De Maeyer, 2007).

Figuur 5: De dragers van de beleidsvoerend vermogen van scholen



b. OMSCHRIJVING VAN DE DRAGERS VAN BELEIDSVOEREND VERMOGEN

In het volgende omschrijven we per drager van beleidsvoerend vermogen hoe deze zich in scholen manifesteert. De veronderstelling is dat scholen die in hoge mate tegemoet komen aan de kenmerken van de verschillende dragers beleidseffectiever zijn dan scholen die er in mindere mate aan voldoen. We ronden elke beschrijving af met het oplijsten van een aantal mogelijke operationalisaties van de drager in kwestie.

- *Doeltreffende communicatie*

Om het beleid in scholen te begrijpen kan men niet voorbij aan de rol van communicatie. De Fleur et al. (1993; Van Petegem e.a., 2007) omschrijven communicatie in organisaties als een verzending van boodschappen via formele en informele netwerken die resulteert in de constructie van kennis en beïnvloeding van individuen en groepen. De manier waarop dergelijke communicatie vorm krijgt in scholen wordt in sterke mate bepaald door de waarden en normen die de leden van de schoolgemeenschap daaromtrent delen. Volgens Devos (2000) uit zich dit onder meer in openheid nastreven, tijd nemen om standpunten te begrijpen, luisteren en feedback geven. Te verwachten is dat het beleidsvoerend vermogen van scholen groter is naarmate mensen meer open communiceren over hun motie-

ven, ideeën, verzuchtingen en onzekerheden, en naarmate men meer bereid is tot actief luisteren en tot het geven van feedback (Levine & Lezotte, 1990). Naast de houding van de schoolleden is ook de structurele omkadering - denk aan structurele informatiekanalen, overlegmomenten en ontmoetingsmomenten - binnen de school van belang. Die moet de horizontale én verticale doorstroming van communicatie mogelijk maken. Communicatie verloopt echter niet uitsluitend via formele kanalen. Informele communicatie binnen de school heeft minstens evenveel impact, zowel ten goede als ten kwade. Roddel, achterklap, rookgordijnen en spookverhalen kunnen de gemoederen bijvoorbeeld lang bezig houden en kunnen het te voeren beleid in scholen lange tijd verlammen.

Indicatief voor de doeltreffendheid van communicatie in scholen is onder meer de mate waarin men er in slaagt om informatie bestemd voor leerkrachten vlot beschikbaar te maken; om zo volledig als gewenst te communiceren; om de communicatie tussen directie en leerkrachten in beide richtingen te doen verlopen; en om overleg- en besluitvormingsprocedures overzichtelijk te houden.

- *Ondersteunende relaties en samenwerking*

Uit onderzoek van Kwakman (1999) is gebleken dat samenwerking door docenten zelf genoemd wordt als een belangrijke bron om van te leren. Om de draagkracht voor beleidsvoering te ontwikkelen en in stand houden moet er in scholen sprake zijn van samenwerking en voldoende toegankelijke en beschikbare ondersteuning (Levine & Lezotte, 1990; Senge, 1997; Onderwijsraad, 2001; Jeris, 2003). Deze ondersteuning kan materieel van aard zijn (didactisch materiaal e.d.), maar evenzeer immaterieel in de vorm van ondersteunende relaties en vertrouwen (Swieringa & Wierdsma, 1990; Tschannen-Moran, 2001). We spreken van een ondersteunende professionele relatie wanneer men het gevoel heeft niet alleen te staan in de uitoefening van zijn taak, maar actief gevolgd en begeleid wordt door collega's (directie en leerkrachten) en door instanties zoals pedagogische begeleidingsdiensten en Centra voor Leerlingenbegeleiding. Wanneer er daarnaast ook (positieve) aandacht is voor de persoonlijke ontwikkeling en persoonlijke ervaringen is er sprake van ondersteunende persoonlijke relaties (Heck & Marcoulides, 1996; Geijsel & Sleegers, 2001). Een interessant perspectief op de invulling van ondersteunende relaties, is 'collegialiteit'. Elementaire waarden die in beschrijvingen van collegialiteit voorkomen (Little, 1990b; Hargreaves & Fullan, 1992; von Krogh, Ichijo & Nonaka, 2000; Marlow, Kyed & Connors, 2005) zijn: samenwerking, steun, vertrouwen, openheid en betrokkenheid. Ook begrip, consensus, solidariteit, dialoog en gedeelde verantwoordelijkheid zijn uitingen van collegialiteit. Vertrouwen en openheid zorgen ervoor dat leerkrachten hun eigen klaspraktijk, eventuele problemen of vragen, maar ook positieve resultaten aan anderen durven voorleggen en dat ze zich meer bij het schoolbeleid betrokken voelen (Smylie, 1992; Verbiest & Vandenberghe, 2002).

Indicatief voor de ondersteunende relaties en samenwerking in scholen is onder meer de mate waarin men er in slaagt om leerkrachten kansen te geven om activiteiten samen voor te bereiden; om d.m.v. samenwerking tussen leerkrachten zoveel mogelijk dubbel werk te vermijden; om er voor te zorgen dat leerkrachten zicht hebben op elkaars leerstof; en om er voor te zorgen dat leerkrachten bij elkaar steun vinden.

- *Gedeeld leiderschap*

Wij schuiven hier één van de vele conceptuele invullingen van leiderschap in scholen naar voren, namelijk de mate van gedeeld leiderschap (Leithwood & Duke, 1998; Spillane, Halverson & Diamond, 2001; Harris, 2003). Leiderschap beweegt zich op een continuüm met aan de ene pool een concentratie van leiderschap in handen van één individu (autocratisch leiderschap) en aan de andere pool de situatie waarin leiderschap gedeeld wordt onder alle teamleden (Rice & Schneider, 1994; Smith & Piele, 1997). Er zijn bepaalde beleidsdomeinen waarbij een leerkrachtenteam aanvaardt dat de directeur individueel beslist. Maar uiteraard zijn er heel wat domeinen en situaties waarbij (een deel van) het team niet aanvaardt dat de directeur autonoom optreedt. Teamleden willen dan mee beslissen. Gedeeld leiderschap verbetert in dat geval de kwaliteit van beslissing (Levine & Lezotte, 1990). Beslissingen die samen genomen worden zijn vaak makkelijker uit te voeren en hebben meer slagkracht omdat de betrokkenheid van het team op het beleid groeit; leerkrachten gaan zich doelmatiger voelen en het team krijgt meer begrip voor de rol en moeilijkheden van de directeur (Kessler, 1992; White, 1992; Tjepkema, 1993; Jongmans, Beijaard & Biemans, 1996; Simons, 1997; Harris, 2003). Sleegers (1991) heeft aangetoond dat scholen waar leerkrachten in hoge mate participeren aan de beleidsvoering beter in staat zijn om een zelfstandig beleid te voeren. Toch zijn er ook nadelen aan verbonden. Daarbij denken we aan verhoogde werklast en tijdsinvestering, en de mogelijkheden van kritiek en conflict waaraan deze vorm van leiderschap zich bloot stelt (McBeath, Moos & Riley, 1998). Daarenboven werkt opgelegde (gedwongen) participatie volgens Smith & Piele (1997) meestal contraproductief. Het is dus moeilijk om te spreken van één algemeen geldende leiderschapsstijl die te allen tijde de meest effectieve is (Harris, 2002; Vanhoof & Van Petegem, 2006b).

> *Indicatief voor het leiderschap in scholen is onder meer de mate waarin men er in slaagt om de personeelsleden bij beleidsvoorstellen te betrekken; om met het volledige personeelskorps te overleggen voordat er belangrijke beslissingen genomen worden; om zichzelf als leider te laten adviseren; om te komen tot een duidelijke visie over het eigen leidinggeven; om taken succesvol te delegeren; en om de personeelsleden te stimuleren tot het nemen van initiatieven.*

- *Gezamenlijke doelgerichtheid*

De kans dat scholen succesvol zijn in het beleid dat ze voeren, is mede afhankelijk van de mate waarin ze erin slagen om als gemeenschap te komen tot een gezamenlijke doelgerichtheid (Peters & Waterman, 1982; Levine & Lezotte, 1990; Sammons e.a., 1995; Hooge, 1998; Majoor, 2000). Het ontwikkelen van een gezamenlijke doelgerichtheid bestaat erin de verschillende wensen en verschillende ideeën in overeenstemming te brengen met wat ze gemeenschappelijk hebben, namelijk 'school maken' (Senge, 2001). Net door het gezamenlijke in overeenstemming te brengen, ontstaat er een vorm van betrokkenheid op een toekomstig doel van waar men met de school naar toe wil (Simons, 1997; Maslowski, 2001; Verbiest & Vandenberghe, 2002). Doelen die worden ervaren als duidelijk, uitdagend en realiseerbaar verhogen de motivatie van teamleden om zich in te zetten (Geijsel & Sleegers, 2001). Er wordt van de aanwezigheid van duidelijke en specifieke doelen een sterke impact verwacht op de keuze en uitvoering van activiteiten, en daar-

mee ook op het succes waarmee de vooropgestelde doelen bereikt worden (Vandenberghe & Van der Vegt, 1992; Vandenberghe, 1995). Zowel de doelgerichtheid als het gezamenlijk karakter ervan zijn dus van belang.

In het operationaliseren van gezamenlijk doelgerichtheid gebruiken wij twee invalshoeken, de 'sterkte van de visie' in de school en de mate van 'gezamenlijke doelgerichtheid'. Indicaties voor de 'sterkte van de visie' zijn onder meer duidelijke prioriteiten voorop stellen; een lange termijnvisie voorop stellen, de visie vertalen in beleidskeuzes en duidelijke doelen, en de visie bespreekbaar maken. Indicatief voor de 'gezamenlijke doelgerichtheid' is de mate waarin de school erin slaagt om alle leerkrachten achter de huidige aanpak te scharen, om overeenstemming te brengen in de aanpak van problemen, om de aanpak van problemen tot ieders verantwoordelijkheid te maken en om leerkrachten de visie van de school te laten vertalen in hun concrete aanpak.

- *Responsief vermogen*

Om het beleidsvoerend vermogen van scholen ten volle in kaart te brengen, beschouwen we de schoolorganisatie als een open systeem (Senge, 1997). Scholen moeten in relatie treden met hun omgeving en inspelen op de impulsen die vanuit de omgeving op hen afkomen (Simons, 1997). De omgeving heeft immers talrijke vragen en verwachtingen ten aanzien van de school. Uiteindelijk is de school er voor de leerlingen, hun ouders en de samenleving als geheel. Responsieve scholen treden in relatie met die omgeving en spelen succesvol in op impulsen vanuit die omgeving (Onderwijsraad, 2001). Voor een school is het niet zozeer de kunst om te kunnen inspelen op al de vragen en verwachtingen, dan wel zo adequaat mogelijk te reageren vanuit de bestaande schoolvisie (Bolhuis & Simons, 1999; Wilkins, 2002). Om responsief te kunnen reageren op omgevingselementen moet er in de school in de eerste plaats een open en toegankelijke cultuur heersen die maakt dat directie en leerkrachten beschikbaar zijn voor de verschillende actoren uit de schoolomgeving. Leden van het schoolteam moeten op de hoogte zijn van recente ontwikkelingen in de omgeving en bereid zijn zich daaromtrent te informeren en eventueel bij te scholen. De relatie met de omgeving kan zich ontplooien in formele relaties, informele relaties, partnerschappen/allianties of netwerken (Hargreaves & Fullan, 1998).

We maken in het operationaliseren van responsief vermogen een onderscheid tussen ouders enerzijds en de ruimere omgeving anderzijds. Indicatief voor het 'responsief vermogen van scholen ten aanzien van ouders' is onder meer de mate waarin men er in slaagt om de contacten met ouders zo in te richten dat leerkrachten ze als ondersteunend ervaren en de mate waarin men om open staat voor impulsen van ouders. Responsiviteit ten aanzien van de ruimere omgeving blijkt uit de mate waarin de school er in slaagt om voorstellen van externen in het schoolbeleid op te nemen, om bij problemen bij externen ten rade te gaan en om in te spelen op evoluties in de samenleving.

- *Innovatief vermogen*

Omgaan met vernieuwingen is een cruciale drager van het beleidsvoerend vermogen van scholen (Visscher, 1993). Het innovatief vermogen van scholen weerspiegelt de mate waarin zij in staat zijn vernieuwingen op een succesvolle manier te implementeren (Van den Berg e.a., 1994; Geijsel, 2001; Potter e.a., 2002) maar ook hoe zij omgaan met veranderingen vanuit de overheid en vanuit de school zelf (Geijsel, van den Berg & Sleegers, 1998; Kwakman, 1999). Innovatieve scholen stellen zich open voor nieuwe ideeën, werk-

wijzen, … (Bakx & Van der Eerden, 1993). Vernieuwingen die niet onmiddellijk met succes geïmplementeerd worden, worden niet als een achteruitgang beschouwd, maar eerder als een mislukte poging waaruit men kan leren. In een innovatieve school is de drempel om over problemen te praten laag (Bakx & Van der Eerden, 1993; Van den Berg e.a., 1994). De directie spoort er leerkrachten bijvoorbeeld aan om nieuwe dingen uit te proberen (Van den Berg e.a., 1994).

Niet iedere innovatie is voor scholen evenwel relevant en bruikbaar. Scholen moeten nagaan wat de consequenties van vernieuwingen zijn (Vandenberghe & Van der Vegt, 1992). Elke vernieuwing impliceert immers bepaalde normen en waarden die (tot op zekere hoogte) verenigbaar moeten zijn met de waarden en normen die de leidraad van het schoolgebeuren uitmaken (Staessens, 1993, p. 127). Denk daarbij bijvoorbeeld aan de mate waarin taakomschrijvingen aandacht schenken aan innoverende domeinen (Van den Berg e.a., 1994; Schollaert, 2000). We willen tenslotte benadrukken dat vernieuwingen weinig kans maken zonder de steun van de andere dragers van beleidseffectiviteit (Geijsel, 2001; Mintzberg, 2001). Fullan (1991, p. 1) schrijft bijvoorbeeld: 'If we know one thing about innovation and reform it is that it cannot be done successfully to others'. Een school met een sterk innovatief vermogen is bijvoorbeeld in staat de tegenstanders van de vernieuwing de zinvolheid ervan te laten inzien zonder hen de vernieuwing op een dwingende manier op te leggen. Ze weet weerstanden en onzekerheid te vermijden en streeft naar een participatieve benadering (Simons, 1997; Geijsel, Sleegers, Van den Berg & Kelchtermans, 2001).

> *Indicatief voor het innovatief vermogen van scholen is onder meer de mate waarin men er in slaagt om leerkrachten een positieve houding te doen aannemen tegenover vernieuwingen; om leerkrachten ertoe aan te zetten om te durven experimenteren en om een flexibele houding van alle leerkrachten te bekomen.*

- *Geïntegreerd beleid*

In dit onderzoek gebruiken we de term 'geïntegreerd beleid' om te verwijzen naar beleid dat zich niet beperkt tot één of twee beleidsdomeinen, maar zich afspeelt op meerdere (alle) domeinen. Het gaat hier dan om een 'globaal' of 'alomvattend' beleid, als tegenhanger van 'partieel' of 'gedeeltelijk' beleid (Verhoeven, 1986). Geïntegreerd houdt daarenboven in dat er een zekere coherentie tussen de verschillende beleidsdomeinen terug te vinden is. Dit sluit bijvoorbeeld aan bij het belang dat onder meer Hanson (1979) en Peters (1997) hechten aan de mate waarin de beheersmatige en onderwijskundige zones op elkaar afgestemd worden en de impact die daarvan uitgaat op het beleidsvoerend vermogen van scholen. Dergelijke afstemming vergt een houding die teamleden stimuleert om zich te informeren over de activiteiten, doelen en wensen van anderen en die hen uitnodigt om met deze informatie rekening te houden in hun eigen handelen (Giesbers & Sleegers, 1994). Als dit het geval is, spreken we van een 'geïntegreerd beleid' (of van afstemming); indien dit niet zo is, van een 'gesegmenteerd beleid' (Verhoeven, 1986). De beroepsopvatting van leerkrachten en schoolleiders is hier verhelderend. Van Wessum (1996) maakt bij leerkrachten bijvoorbeeld een onderscheid tussen een geïntegreerde en een atomistische beroepsopvatting. Leerkrachten die de waarden

en normen van de atomistische opvatting aanhangen, gedragen zich zeer autonoom, plegen weinig overleg en werken zelden samen met collega's. Leerkrachten daarentegen met een geïntegreerde beroepsopvatting, staan veel positiever t.o.v. samenwerking en menen ook dat onderlinge afstemming een goede zaak is.

Indicatief voor het geïntegreerde karakter van het beleid in scholen is onder meer de mate waarin men er in slaagt om duidelijke afspraken te maken m.b.t. de werking van verschillende raden/werkgroepen, om leerkrachten wegwijs te maken in het geheel van raden/werkgroepen, om begeleidingsinitiatieven op elkaar af te stemmen, en om verschillende werkgroepen op een complementaire manier te doen samenwerken.

- *Reflectief vermogen*

Reflectie is essentieel om tot leren te komen, dat geldt voor individuen maar ook voor scholen (Verbiest & Vandenberghe, 2002). De dominante cultuur in veel scholen is er echter één van doen en handelen, met weinig of geen aandacht voor reflectie en leren (York-Barr, Sommers, Ghere & Montie, 2001). Tjepkema (2002) stelt zelfs dat de te sterke actiegerichtheid een belemmering is voor het leren in scholen. Om het complexe handelen in scholen beheersbaar te maken, worden vaak gewoonten en routines ontwikkeld die ten koste gaan van een reflectief bewustzijn dat uitnodigt om opmerkzaam na te denken over wat men doet. *'Scholen zijn zelf nog niet gewend om systematisch na te gaan of ze hun doelen (als die al geformuleerd zijn) bereikt hebben'* (Reezigt, 2001, p. 101). Een school waarin dit reflectief bewustzijn wèl sterk aanwezig is, bestempelen we als een school met een sterk reflectief vermogen.

Onder het reflectief vermogen van scholen verstaan we de persoonlijke bereidwilligheid van de teamleden om hun eigen benaderingen en handelen in vraag te stellen en te verbeteren (Simons, 1997; York-Barr e.a., 2001). Het gaat om een algemeen kritische, vragende oriëntatie t.o.v. het eigen handelen en een bereidwilligheid om in dat eigen handelen zowel positieve als negatieve elementen te ontdekken en te analyseren (Verbiest, 2003). Daartoe wordt op weloverwogen tijdstippen doelbewust afstand genomen van het dagdagelijkse handelen om met een open blik (voor andere werkwijzen) stil te staan en na te denken (Nonaka & Takeuchi, 1995; Devos, 2000). Het doel is te komen tot een beter begrip van de situatie zodat de acties die ondernomen worden aan effectiviteit winnen (Bright, 1996; Greenwood, 1998).

Indicatief voor het reflectief karakter van het beleid in scholen is onder meer de mate waarin men er in slaagt om leerkrachten systematisch te evalueren, om weerstanden van leerkrachten bij het vooruitzicht te worden geëvalueerd te overwinnen, om leerkrachten hun eigen aanpak systematisch in vraag te doen stellen, en om op schoolniveau aan zelfevaluatie te doen.

Met de beschrijving van het reflectief vermogen van scholen ronden we de toelichting bij de acht dragers van het beleidsvoerend vermogen van scholen af. Onze verwachting is dat scholen die sterk verschillen wat deze dragers betreft ook zullen verschillen in de manier waarop ze vorm geven aan zelfevaluaties. Een gelijkaardig verband verwachten we bij de andere invalshoek om schoolgebonden kenmerken in kaart te brengen, namelijk 'de lerende organisatie'. Deze invalshoek wordt in de volgende paragraaf uitgewerkt.

1.2 De school als lerende organisatie

Het voeren van een effectief schoolbeleid wordt steeds minder beschouwd als het incorporeren van elders ontwikkelde kennis in de eigen aanpak. Nonaka en Takeuchi (1995) stellen bijvoorbeeld dat innovatie en beleidsvoering in organisaties afhankelijk zijn van de capaciteiten om kennis binnen de organisatie te creëren. Het vermogen van een organisatie om op alle niveaus kennis te ontwikkelen en toe te passen wordt steeds meer een kernbekwaamheid van organisaties (Verbiest & Vandenberghe, 2002). Waar het gaat om professionalisering (in scholen) is er met andere woorden een accentverschuiving van opleiden (kennis verwerken) naar leren (kennis creëren) (Wenger, 1998; Donnenberg, 1999; Patriotta, 2004). Uit onderzoek blijkt dat organisaties in problemen doorgaans niet of nauwelijks in staat zijn om te leren. De meest succesvolle organisaties - profit maar ook non profit - blijken in staat om de persoonlijke ontwikkeling van individuen in de organisatie te koppelen aan een betere effectiviteit op organisatieniveau (Bomers, 1990; Senge, 2001).

Ook in de context van onderzoek naar zelfevaluaties verwacht men veel van de beginselen van de lerende organisatie (Devos e.a., 1999, p. 38). In dit onderzoek worden scholen daarom ook benaderd als 'lerende organisaties'. Het begrip heeft de voorbije jaren steeds meer aandacht en aanhang gekregen waardoor er een veelheid aan omschrijvingen van het concept in omloop zijn. Verschillende onderzoeksbenaderingen (zoals organisatieleer, psychologie, economie, management- en innovatiestudies, schoolontwikkeling) hebben het concept immers vanuit zeer verscheidene invalshoeken bestudeerd (Dodgson, 1993). Verbiest (2002, p. 14) stelt nochtans dat de conceptuele en empirische basis van de 'lerende organisatie' zwak ontwikkeld is. De omschrijvingen van 'de lerende organisatie' komen dan ook vaak neer op een aantal relevante kenmerken of aspecten (Güthe, 1997) en kennen weinig overeenstemming (Bolhuis & Simons, 1999, p. 40). Hoedanook, de ideeën over de lerende organisatie hebben ook het denken en onderzoek over de organisatie van scholen beïnvloed.

1.2.1 Van de invulling van begrippen naar een onderzoeksfocus

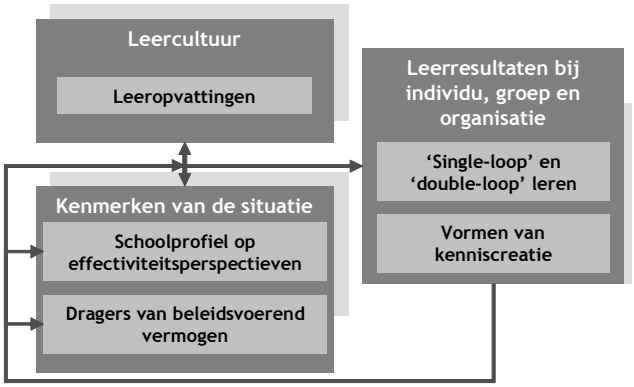
Wij gebruiken de volgende beschrijving van het concept 'lerende organisatie': 'Een lerende organisatie is een organisatie die *als bewust beleid* het leren van al haar leden faciliteert en die zichzelf voortdurend verandert *ter optimalisering van haar effectiviteit*' (dit is de beschrijving van Pedler et al. (in Dodgson, 1993, p. 377) uitgebreid met deze van Bomers (1990, p. 21) (*in italic*)). Deze beschrijving grijpt sterk terug naar algemene omschrijvingen. Dat hoeft echter geen probleem te zijn. Het gaat immers eerder om een manier om naar scholen te kijken - een perspectief - dan om een begrip waarvan de betekenis in concrete kenmerken kan bepaald worden (Güthe, 1997). Begin jaren '90 schreven Swieringa & Wierdsma (1990, p.49) dat 'hoe een dergelijke lerende organisatie er uit zal zien een vraag is waar we nog weinig kunnen over zeggen om de eenvoudige reden dat dit soort organisaties vrijwel nog niet bestaat'. Ruim een decennium later spreekt Verbiest (2002, p.14) in dezelfde lijn van een 'niet al te concreet ideaalbeeld

waarvan geen of weinig voorbeelden beschikbaar zijn'. Het ideaalbeeld is breed gedragen maar wordt (ook in het eigen handelen) slechts moeizaam en weinig toegepast (Bolhuis & Simons, 2001, p. 38). De lerende organisatie is nog steeds geen duidelijk einddoel dat definitief bereikt kan worden; het uitgangspunt is in eerste instantie een gemeenschappelijk streven (Bolhuis & Simons, 1999; Geijsel e.a., 2001).

In een onderwijscontext wordt ook vaak verwezen naar 'professionele leergemeenschappen'. Hord (1997, p. 6) definieert een professionele leergemeenschap als *'het permanent zoeken en delen van leren van leerkrachten en schoolleiders met als doel het verbeteren van het onderwijs aan de leerlingen. Het gaat om een gemeenschap van voortdurend onderzoek en verbetering.'* Deze omschrijving deelt met andere omschrijvingen van het concept een nadruk op het samen (of collectief) leren over de onderwijspraktijk, met de expliciete bedoeling te verbeteren (Verbiest & Vandenberghe, 2002, p. 9). Het onderscheid tussen een 'lerende organisatie' en een 'professionele leergemeenschap' berust op een verschillende invulling van het concept gemeenschap (Mitchell & Sackney, 2000). In een lerende organisatie zouden de belangrijkste doelen betrekking hebben op groei, productiviteit en effectiviteit van de organisatie. Deze doelen worden bovendien door een 'elite' van de organisatie bepaald. In een professionele leergemeenschap worden de doelen door de leden van de gemeenschap bepaald. Ze hebben er eerder betrekking op de groei en ontwikkeling van de mensen zelf.

In de volgende paragrafen wordt het concept 'leren in organisaties' verder uitgediept. Om één en ander te ordenen, inspireren we ons op het 'heuristisch model voor collectief leren in schoolorganisaties' (zie Figuur 6) (Bolhuis & Simons, 1999; Verbiest, 2002). Er is sprake van collectief leren wanneer mensen samen leerprocessen doormaken (De Laat, Poell, Simons & van der Krogt, 2001). Op die manier kan de kennis uitgewisseld en onderhandeld worden met anderen in de organisatie (Verbiest, 2002, p. 15). Het model toont de condities van collectief leren en de leerresultaten ervan in samenhang.

Figuur 6: Het heuristisch model voor collectief leren (naar Verbiest, 2002)



De *leerresultaten* bij individuen, groepen of bij de school als geheel worden bepaald door twee elkaar beïnvloedende groepen van kenmerken: de *leercultuur* en de *kenmerken van de situatie* (Bolhuis & Simons, 1999; Verbiest, 2002; Arts, Kok, Sleegers, Verbiest & de Wit, 2003). De kenmerken van de situatie werden hoger beschreven bij het focussen op organisatie-effectiviteit van scholen (cf. het profiel van scholen op de effectiviteitsperspectieven van Quinn en Rohrbaugh (1983) en de mate waarin de school tegemoet komt aan de dragers van beleidsvoerend vermogen). In de twee volgende paragrafen concentreren we ons op de leercultuur en de leerresultaten in scholen.

1.2.2 De leercultuur in scholen: Opvattingen van teamleden over het eigen leren

Individueel leervermogen kan worden omschreven als de mate waarin iemand in staat en bereid is zijn eigen leren te sturen en vorm te geven (Simons, 1991). Gedurende hun leven ontwikkelen mensen een eigen leerstijl; een redelijk consistente aanpak van het leren (Kolb, 1984; Vermunt, 1992). Net als individuen hebben organisaties een soort leerstijl die door Verbiest (2002) een collectieve leercultuur genoemd wordt. De impact van de leercultuur op het leren in organisaties blijft evenwel vaak impliciet. Door in deze eerste invalshoek te focussen op de opvattingen van individuen over leren willen we de leercultuur in scholen expliciteren. Ook voor scholen geldt immers dat de houding van individuen ten aanzien van hun eigen leren een belangrijk onderdeel is van de organisatie (Verschaffel & de Corte, 1998).

Om de opvattingen ten aanzien van het eigen leren en om leervaardigheden in kaart brengen doen we een beroep op het (sociaal)constructivistische gedachtegoed inzake leren. We staan eerst – bij wijze van verantwoording - stil bij de opkomst van het (sociaal)constructivistische gedachtegoed om vervolgens het gedachtegoed als dusdanig toe te lichten.

a. DE OPKOMST VAN HET (SOCIAAL)CONSTRUCTIVISTISCHE GEDACHTEGOED INZAKE LEREN

In het huidige informatietijdperk worden informatie en kennis niet meer aanzien als statisch en eindig, maar als dynamisch en infinitief (Bomers, 1990; Birenbaum, 1995). Er is teveel kennis om ze allemaal over te dragen en paraat te houden bij de leden van een (school)organisatie (Verbiest & Vandenberghe, 2002). Het traditionele onderwijs wordt door deze ontwikkelingen steeds meer als ontoereikend ervaren en termen als levenslang leren, professionele ontwikkeling en leren op de werkplek hebben hun ingang gedaan (bijvoorbeeld Verbiest, 2002; Simons, 1991). Een heroriëntatie op het leren van individuen was/is dan ook noodzakelijk. In de laatste helft van de 20ste eeuw brak geleidelijk aan een nieuw psychologisch leermodel door: leren is een activiteit van het individu zelf dat handelend en in gesprek met anderen zijn cognitieve (kennis)structuren opbouwt (Bolhuis, 2000; Van der Haar, 2001). Op basis van de ontwikkeling van nieuwe leertheorieën is er veel meer inzicht ontstaan in de complexiteit van het leerproces. Leerprocessen kennen geen standaard; iedereen leert op zijn of haar eigen wijze. Dat maakt het aanbieden van 'leerervaringen' er niet eenvoudiger op. 'Leren van nu' moet daarom gericht zijn op voortdurende verandering, op zelfontplooiing en actief bezig zijn. De

evolutie van een *acquisition metaphor* naar een *participation mathaphor* is tekenend in dit opzicht (Sfard, 1998). Kennisoverdracht vervaagt en leren wordt steeds meer het aanreiken en ontdekken van vaardigheden om zelf kennis te construeren (Bolhuis & Simons, 2001). Er ligt minder nadruk op het overbrengen van vaststaande feiten en overgeerfde kennis, maar meer op interpretatieschema's voor de complexiteit die ons omringt (Herman, Aschbacher & Winters, 1992; Arter & Bond, 1996; Vermunt, 1998). Dat heeft grote implicaties voor het leren van individuen in (school)organisaties. De manier waarop mensen denken over leren kan immers een invloed uitoefenen op de manier waarop zij zelf denken en handelen in een lerende context (Donche, 2005).

b. SOCIAALCONSTRUCTIVISME ALS COGNITIEVE LEERTHEORIE

In tegenstelling tot de traditionele onderwijsbenadering, die gebaseerd is op behavioristische leertheorieën, zijn de hedendaagse onderwijsbenaderingen gebaseerd op de cognitieve leertheorie en het sociaalconstructivistische gedachtegoed (De Klerk & Simons, 1988; Van der Veen & Van der Wal, 1989; Herman e.a., 1992; Birenbaum, 1995). De cognitieve psychologie ziet leren als een informatieverwerkingsproces; ze gaat na hoe individuen informatie verwerken en hoe die informatie in het geheugen wordt opgeslagen (De Klerk & Simons, 1988; Boekaerts & Simons, 1995). Een recente stroming binnen deze cognitieve psychologie is het sociaalconstructivisme. De kernidee achter deze theorie is dat leren een actief en constructief proces is, zelfs in die mate dat leren alleen maar goed mogelijk is als de lerenden zelf actief construerend met de leerstof bezig zijn (Cobb, 1996). Belangrijke lessen uit deze theorieën zijn dat men leren kan leren en dat er geen 'beste' manier bestaat om te leren. Het leren vindt plaats in de lerende die nieuwe kennis en vaardigheden opbouwt, gebruik makend van reeds aanwezige kennisschema's. Leren is dus een voortdurend proces waarin individuen continu informatie ontvangen, interpreteren en in verband brengen met wat ze reeds weten en ervaren hebben (voorkennis). Op basis van nieuwe informatie worden de eigen concepten van de wereld (mentale schema's) aangepast. De kunst van het leren is dat nieuwe informatie gekoppeld wordt aan bestaande voorkennis. Pas als de eigen concepten zijn veranderd, is er wat bijgeleerd. Mensen construeren dus kennis en die constructie verschilt van lerende tot lerende (Clemens et al., 1999).

Inzicht in de belangrijkste elementen van de sociaalconstructivistische leertheorieën is noodzakelijk om zicht te krijgen op het leren van individuen in (school)organisaties (Verschaffel & de Corte, 1998). Vandaar dat we kort stilstaan bij een aantal sleutelkenmerken van deze leertheorieën, namelijk: leren is een actief en constructief, cumulatief, zelfregulatief, doelgericht, contextgebonden en sociaal proces (Herman e.a., 1992; Arter & Bond, 1996; Vermunt, 1998; Verschaffel & de Corte, 1998; Van Petegem, Imbrecht & Brandt, 2005a).

- *Leren is actief en constructief*. Individuen nemen niet zozeer feitelijke informatie op, ze creëren hun eigen begrippen en structuren van de werkelijkheid.
- *Leren is cumulatief maar niet lineair*. Het cumulatieve karakter van leren zit in wezen al vervat in het constructieve: individuen construeren nieuwe kennis op

basis van wat ze al weten en kunnen. Voorkennis is een belangrijke factor bij het leren. Toch mag men hieruit niet concluderen dat leren een mooi lineair proces is. Er is geenszins sprake van een vastliggend hiërarchisch parcours.

- *Leren is zelfregulerend.* Hier gaat het over de activiteiten in verband met het beheren en het bewaken van het leerproces door de lerende zelf, de zogenaamde metacognitieve vaardigheden. Individuen in organisaties moeten leren om zelf de controle over hun eigen leren in handen te nemen.
- *Leren is doelgericht.* Leren zal slechts succesvol zijn als de lerende zich bewust is van het (zelfbepaalde) doel waar hij of zij naar toe werkt. Individuen in organisaties worden evenwel vaak geconfronteerd worden met leerdoelen van buiten af. Het is dan erg belangrijk dat zij zich deze doelen eigen maken, dat ze zich ermee kunnen identificeren.
- *Leren is context- en situatiegebonden.* Denk- en leerprocessen verlopen het best wanneer voortdurend interactie plaatsvindt met anderen en met de eigen omgeving. Leren vindt immers niet plaats in een vacuüm. Dat betekent bijvoorbeeld dat leren dus best een duidelijke band heeft met de realiteit.
- *Leren is collaboratief.* De kwaliteit van leren stijgt wanneer lerenden een 'lerende gemeenschap' vormen. Dit houdt bijvoorbeeld in dat lerenden naar elkaar luisteren, samen reflecteren, op elkaars ideeën verder bouwen. Een cruciaal aspect van het hedendaagse leren is dus de sociale interactie met anderen tijdens de leeractiviteiten.

Als het leren van individuen gekenmerkt wordt door de bovenstaande sociaalconstructivistische kenmerken, dan kunnen op basis daarvan een aantal kenmerken vooropgesteld worden waaraan een (school)organisatie dient te voldoen om tot leervermogen te komen (Snoek, 2004). We formuleren ze aan de hand van een aantal concrete vragen: Wordt de lerende benaderd als een competente persoon met eigen belangen en doelen? Onderschrijven de leden van de organisatie de gerichtheid op de vooropgestelde doelen (leren is doelgericht)? Starten leerprocessen met het activeren van bestaande kennis of vaardigheden en worden daaraan nieuwe kennis en vaardigheden gekoppeld? Wordt er voortgebouwd op en gebruik gemaakt van kennis en vaardigheden waarover de personen al beschikken (leren is cumulatief)? Zijn de lerenden actief betrokken bij het leren en sturen ze hun eigen leerproces? Wordt ook het leerproces als dusdanig bewaakt (cf. metacognitie) (Leren is zelfregulerend)? Wordt er geleerd in interactie met de omgeving (Leren is collaboratief)? Vindt het leren plaats in authentieke situaties? Voltrekken leerprocessen zich in voortdurende interactie met de sociale en culturele context waarin het geleerde verworven en gebruikt wordt (leren is contextgebonden)? De verwachting is dat naarmate het antwoord op de bovenstaande vragen in sterkere mate positief is, het leervermogen van de (school)organisatie groter zal zijn (Bomers, 1990, p. 31).

c. HET METEN VAN DE LEERCULTUUR

Om de percepties van de respondenten ten aanzien van hun eigen leren in kaart te brengen, formuleren we de verschillende kenmerken van het sociaalconstructivisme in contrasterende zinsneden. De respondenten spreken zich dan telkens voor één van beide contrasten of nemen een middenpositie in. De zinsneden zijn deels gebaseerd op bestaande vragenlijsten naar de opvattingen van volwassenen over leren (Bolhuis, 2000; Bolhuis & Simons, 2001) en deels een eigen operationalisering van het theoretisch kader.

Leren is passief informatie ontvangen.	⇔	Leren is kritisch denken en verbanden leggen.
Een goede nascholer vertelt mij precies hoe het zit en wat ik moet doen.	⇔	Een goede nascholer laat me zelf zaken uitvinden en ervaren hoe het werkt.
Leren staat los van wat je al weet.	⇔	Leren is voortbouwen op wat je al weet.
Leren is iets dat door anderen gestuurd wordt.	⇔	Leren is iets dat je zelf stuurt.
Leren staat los van een doel.	⇔	Leren is doelen nastreven.
Leren doe je best door afstand te nemen van de realiteit.	⇔	Je leert het best als er een duidelijke band is met de realiteit.
Ik geloof niet dat je van discussie met anderen veel opsteekt.	⇔	Je leert er veel van om je mening met die van anderen te vergelijken.
Leren is een individueel proces.	⇔	Leren is een sociaal proces.

1.2.3 Leerresultaten: Over de aard van leerprocessen die in scholen plaatsvinden

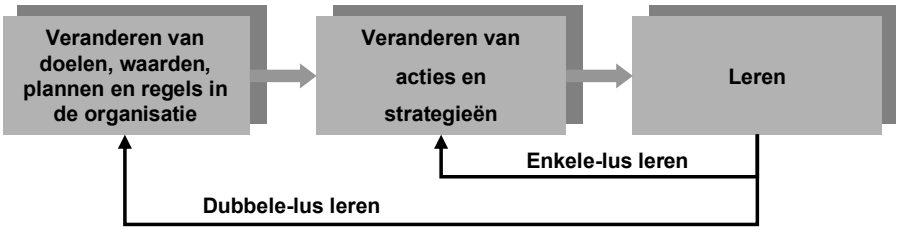
Mensen leren altijd en overal, binnen en buiten hun werk, binnen en buiten het onderwijs (Bolhuis & Simons, 2001). In het volgende focussen we weliswaar op het leren als lid van een schoolteam, maar ook de leerprocessen in organisaties zijn zeer verscheiden. Leren in (school)organisatie kent talrijke verschijningsvormen, gaande van roddel over observatie en dialoog tot het ontwikkelen of toepassen van theorieën (Paslowsky, Forslin & Reinhardt, 2001). In deze studie onderzoeken we deze verscheidenheid aan de hand van twee invalshoeken. Het gaat om het onderscheid tussen 'enkele-lus' en 'dubbele-lus' leren (Argyris & Schön, 1978) enerzijds en de theorie van kenniscreatie (Nonaka & Takeuchi, 1995) anderzijds.

a. 'ENKELE-LUS' EN 'DUBBELE-LUS' LEREN

Voor Argyris en Schön (1978, p. 2) omvat leren het detecteren en corrigeren van error. De mogelijke vormen van error in scholen zijn zeer verscheiden: teruglopende leerlingenaantallen, dalende jobsatisfactie bij leerkrachten, conflict, het niet afgestemd zijn op de huidige noden van de samenleving, een behoefte aan professionele ontwikkeling. De manier waarop (school)organisaties reageren op dergelijke vormen van error is indicatief voor het lerende karakter van de organisatie. Het leren in lerende organisaties is ruimer dan het leren in niet lerende organisaties. Leren in een lerende organisatie vormt een dynamisch proces waarbij niet alleen nieuwe kennis wordt toegevoegd aan de kennisbasis van de organisatie (*enkele-lus leren*), maar ook heersende opvattingen normen, beleidsrichtingen en doelen worden bijgesteld (*dubbele-lus leren*) (Argyris & Schön, 1978; De Laat e.a., 2001). Dubbele-lus leren is cruciaal om tot kwaliteitsvolle zelfevaluaties te

komen (Scheerens, 2004). De verschillende vormen van leren worden kort toegelicht. Het onderscheid kan op een eenvoudige manier schematisch worden voorgesteld (zie Figuur 7).

Figuur 7: Het onderscheid tussen enkele-lus en dubbele-lus leren



- *Enkele-lus leren*

Een organisatie kan proberen error te corrigeren door bestaande werkwijzen en regels binnen de geldende normen van de organisatie aan te passen. Er wordt - wanneer er iets mis loopt - gezocht naar een andere strategie die het mogelijk maakt tot een oplossing te komen. Dat gebeurt binnen het bestaande kader van doelen, waarden, plannen en regels. Het kader zelf wordt niet in vraag gesteld. Argyris en Schön (1978, p. 20-21) omschrijven deze vorm van leren als volgt: '*Single-loop learning is sufficient where error correction can proceed by changing organizational strategies within a constant framework of norms for performance. It is concerned primarily with effectiveness - that is, with how best to achieve existing goals and objectives and how best to keep organizational performance within the range specified by existing norms.*' Volgens Dodgson (1993, p. 383) gaat het bij enkele-lus leren om activiteiten die bijdragen tot de kennisbasis en schooleigen competenties of routines zonder de fundamentele aard van deze elementen of de activiteiten van de organisatie te veranderen. De context van enkele-lus leren laat zich kenmerken door vaagheid, ambiguïteit, oncontroleerbaarheid, onbespreekbaarheid, het achterhouden van informatie, versnipperde informatie, onzekerheid en incompatibiliteit van praktijktheorieën (Argyris & Schön, 1996).

- *Dubbele-lus leren*

In sommige gevallen moeten organisaties zich om een bestaande error te corrigeren engageren in een leercyclus waarin ook het bestaande kader kritisch in vraag wordt gesteld (Argyris & Schön, 1978, p. 21). Met dit laatste wijzen Argyris en Schön op het belang van dubbele-lus leren. Zulk een vorm van leren kan resulteren in een verandering van de doelen, waarden, plannen en regels van de organisatie. In tegenstelling tot bij enkele-lus leren worden aannames in twijfel getrokken en vergelijkt men (praktijk)theorieën om met een nieuwe theorie voor de dag te komen. De ontwikkeling van scholen naar professionele leergemeenschappen is een voorbeeld van dubbele-lus leren. Het is een verandering die de cultuur en structuur van een school fundamenteel raakt (Verbiest & Vanden-

berghe, 2002, p. 79). De reactie op error is bij dubbele-lus leren eerder gericht op specificeren, verduidelijken, toetsen, tot openheid en discussie trachten te komen, onderzoeken en het afstemmen van praktijktheorieën (Argyris & Schön, 1996).

➢ *Het meten van enkele-lus en dubbele-lus leren*

De reactie van een school op een concreet probleem (error) laat zich maar begrijpen door een diep inzicht in de concrete context waarin deze error zich voordoet. In een schriftelijke vragenlijst kan deze diepgang zeer moeilijk bereikt worden. We trekken de vraag over het leren in de school daarom op naar een meer abstract niveau en brengen in kaart hoe de school in algemene termen reageert op problemen die zich voordoen. Daarbij wordt niet verwezen naar concrete problemen. Van de respondenten wordt gevraagd dat ze aangeven hoe vaak het gebeurt dat men in de school op de volgende manieren reageert bij problemen.

- *Niets ondernemen.*
- *Het probleem niet erkennen of minimaliseren.*
- *De (echte) kern van de zaak onbesproken laten.*
- *Informatie achterhouden.*
- *Discussies over het probleem trachten te vermijden.*
- *Het probleem analyseren maar geen acties ondernemen.*
- *Aanpassingen doen zonder de achterliggende redenen aan te pakken.*
- *Oplossingen kiezen zonder deze te toetsen op hun kwaliteit.*
- De verschillende meningen van betrokkenen verzamelen.
- Onderlinge spanningen en meningsverschillen bespreekbaar maken.
- Objectieve gegevens verzamelen om het probleem in kaart te brengen.
- De gevoelens die het probleem te weeg brengt bespreken.
- De eigen waarden en normen in vraag stellen.
- De uitgangspunten van de school in vraag stellen.

Het bovenstaande lijstje valt uiteen in twee types reacties bij problemen. Schuingedrukt zijn eerst de doorgaans niet constructief geachte reacties weergegeven (eerder single loop). Daarna volgen de doorgaans wel constructieve geachte reacties (eerder double loop). De scores op de schuingedrukte items worden eerst gespiegeld en vervolgens bij de andere scores opgeteld. Op die manier heeft elke respondent voor zijn school een score op het continuüm 'enkele-lus' versus 'dubbele-lus' leren.

b. DE THEORIE VAN KENNISCREATIE (NONAKA & TAKEUCHI, 1995)

Nonaka & Takeuchi (1995) stellen in hun boek 'De kenniscreërende organisatie' dat het succes van Japanse bedrijven te verklaren is door hun vaardigheden en expertise op het gebied van organisatorische kenniscreatie. Een belangrijk verschil met het Westen is volgens hen dat Japanse organisaties anders denken over kennis. Kennis is er meer dan wat doorgaans in woorden en getallen uitgedrukt wordt; men ziet er kennis in eerste instantie als stilzwijgend, onzichtbaar en moeilijk expliciet te maken. Leren is dus meer dan bewust plannen en organiseren (Simons, 2003). Een ruimere opvatting opent de ogen voor leren dat plaatsvindt in de praktijk, tijdens het dagelijks functioneren (Kwakman, 1999). Kennis is immers niet altijd met woorden te omschrijven en bestaat in verschillen-

de vormen: impliciet en persoonsgeboden enerzijds maar ook expliciet en gecodificeerd anderzijds. Beide vormen staan niet volkomen los van elkaar en vullen elkaar aan (Nonaka & Takeuchi, 1995; Jorna, 2004).

- *Impliciete en expliciete kennis*

Impliciete (stilzwijgende of persoonsgebonden impliciete) kennis is sterk individueel gekleurd en moeilijk te formaliseren. Het gaat om de - grotendeels onbewuste - kennis die mensen (leerkrachten en schoolleiders) tijdens het werk ontwikkelen (Nonaka & Takeuchi, 1995; von Krogh e.a., 2000; Bolhuis & Simons, 2001). Daarbij denken we aan subjectieve inzichten, eigen intuïties en plotse invallen. Deze kennis is sterk ingebed in de praktijk en de expertise van individuen (of van een collectief). Ze is nauw verbonden met idealen, waarden en emoties van personen. Denk bijvoorbeeld aan de vraag naar het 'waarom en hoe' van klasmanagement. Hoewel ervaren leerkrachten deze technieken zeer goed beheersen, ervaren ze bijvoorbeeld problemen om hun aanpak te verwoorden wanneer beginnende leerkrachten om advies vragen (Ballet, Kelchtermans, Berens & Janssens, 2000). Omdat impliciete kennis zich moeilijk leent tot formalisering is het moeilijk ze te documenteren en met anderen te delen; tenzij voor mensen die een gezamenlijke werkplek delen. Voor hen is het makkelijker om impliciete kennis te delen. Denk bijvoorbeeld aan de vaststelling dat leerkrachten hetgeen ze elkaar vertellen over hun eigen praktijk, doorgaans meer inspirerend vinden dan een lezing van een externe expert (Hargreaves, 1992; Leenheer, 2002). Impliciete kennis wordt eerder geleerd door oefening, vallen en opstaan, imitatie en samen ervaringen opdoen.

Bij expliciete of (gecodificeerde) kennis gaat het om kennis die via een formele, systematische taal kan worden overgebracht op anderen. Deze vorm van kennis kennen we als geschreven teksten in handboeken, formules, schema's, beelden, richtlijnen, procedures, etc. Het grote voordeel ervan is dat het kan worden opgeslagen en makkelijk kan worden overgedragen tussen mensen. Expliciete kennis is (in tegenstelling tot impliciete kennis) niet persoons- en contextgebonden.

- *Vier vormen van kennisconversie*

Nonaka & Takeuchi (1995) zijn van mening dat in het Westen de nadruk nog veel te veel ligt op het expliciete leren in de vorm van bijvoorbeeld onderwijs en cursussen. Het impliciete leren waarbij het gebruik van ervaring en intuïties een belangrijke rol speelt, zou meer naar waarde moeten worden geschat. Zij vestigen de aandacht op vier types leerprocessen. Deze vormen de 'motor' van het proces van kenniscreatie. Er is telkens sprake van een sociaal proces van conversie van kennis. Het gaat dus niet om een proces binnen individuen maar om een proces tussen individuen. Het model stelt vier soorten van kennisconversie voorop: externaliseren, combineren, internaliseren en socialiseren (zie Figuur 8).

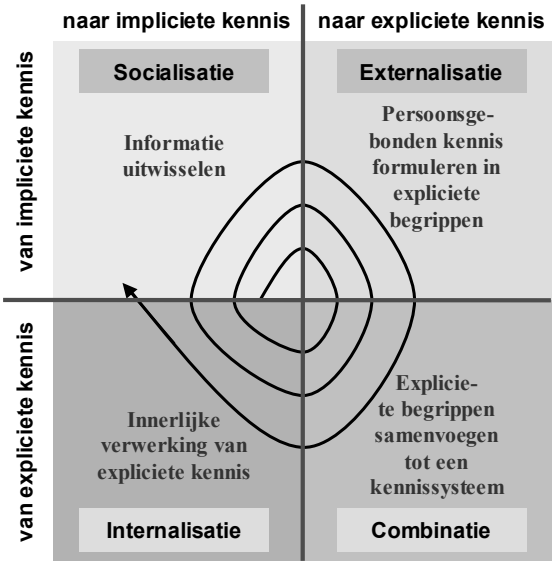
- *Externaliseren, van impliciete kennis naar expliciete kennis*

Externalisatie is het proces waarin impliciete kennis in expliciete begrippen wordt uitgedrukt. Persoonsgebonden kennis wordt in de vorm van metaforen, analogieën, concepten, hypothesen en modellen expliciet gemaakt. Simons (2003) spreekt in dit verband over uitwerken. Een belangrijke rol is hier weggelegd voor dialoog en gezamenlijke reflectie. Van de vier vormen van kennisconversie neemt externalisatie een sleutelpositie in. Het is immers hier dat nieuwe expliciete kennis geboren wordt.

- *Combineren, van expliciete kennis naar expliciete kennis*

Combinatie is een proces waarin begrippen tot een kennissysteem worden gesynthetiseerd. Verschillende structuren van expliciete kennis worden met elkaar gecombineerd. Mensen wisselen kennis uit en combineren die door middel van documenten, vergaderingen, telefoon- en internetverkeer. Nieuwe kennis ontstaat door het sorteren, aanvullen, combineren en categoriseren van bestaande informatie. Dit is een vorm van leren die we in het klassieke onderwijsdenken terugvinden.

Figuur 8: Vier manieren van kennisconversie (Naar Nonaka & Tacheuki, 1995, p. 85)



- *Internaliseren, van expliciete kennis naar impliciete kennis*

Wanneer expliciete kennis deel gaat uitmaken van de persoongebonden kennis is er sprake van internalisatie. Deze vorm van kennisconversie heeft veel te maken met 'al doende leren'. Het individu maakt zich nieuwe kennis eigen in de vorm van vernieuwde mentale modellen en uitgebreide technische knowhow. Wanneer de meeste leden van de organisatie hetzelfde mentale model delen, gaat de betreffende persoonsgebonden informatie deel uitmaken van de organisatiecultuur.

- *Socialiseren, van impliciete kennis naar impliciete kennis*

> Socialisatie is een proces waarbij ervaringen worden gedeeld en impliciete kennis wordt gecreëerd. Daarbij denken Nonaka & Takeuchi (1995) aan mentale modellen of technische vaardigheden. Technische vaardigheden kunnen bijvoorbeeld worden aangeleerd zonder taal (explicitatie) wanneer de lerende in de praktijk leert door waar te nemen en te imiteren. De sleutel tot deze vorm van kennisconversie is dan ook het delen van ervaringen. Simons (2003) benadrukt dat impliciet leren deels ook impliciet moet blijven. Men mag het niet 'verpesten' door het impliciete te snel te willen expliciteren.

Om organisatorische kenniscreatie te doen plaatsvinden, zal de opeenstapeling van persoongebonden kennis op individueel niveau door socialisatie moeten worden overgebracht op andere leden van de organisatie, waardoor een spiraal van kenniscreatie op gang komt (Nonaka & Takeuchi, 1995, p.82). Als dit niet gebeurt wordt de bestaande kennisbasis van de organisatie niet uitgebreid (Simons & Ruijters, 2001). Pas wanneer persoongebonden en expliciete kennis op elkaar inwerken, ontstaan er volgens Nonaka & Takeuchi (1995, p. 84) kansen voor innovatie. Ze noemen organisatorische kenniscreatie dan ook een 'voortdurende en permanente interactie tussen persoongebonden en expliciete kennis' en spreken van een bijhorende kennisspiraal.

Die kennisspiraal bestaat in wezen uit een doorlopen van de verschillende kwadranten, te beginnen bij socialisatie. Een eerste voorwaarde is dan ook een voldoende mate van socialisatie die het uitwisselen van ervaringen en mentale modellen mogelijk maakt. Vervolgens wordt een externalisatie gestart door een zinvolle discussie of collectieve overdenking. Door middel van metaforen en analogieën wordt persoonsgebonden kennis geëxpliciteerd. Daarna wordt de pas gecreëerde en reeds bestaande kennis gecombineerd tot een nieuwe handelingswijze. 'Al doende' zal deze tenslotte door de leden van de organisatie geïnternaliseerd worden en kan het proces van socialisatie opnieuw beginnen. Merk hierbij wel op dat er geen standaardprocedure bestaat waarmee persoonsgebonden kennis overgedragen kan worden in de organisatie. De aard van de verschillende kennisconversies kan verscheidene vormen aannemen en ook de volgorde van de kennisspiraal ligt niet eenduidig vast (Bolhuis & Simons, 2001).

Er is een aantal voorwaarden waarin de leercontext dient te voorzien om kenniscreatie mogelijk te maken: doelgerichtheid, zelfstandigheid, fluctuatie en creatieve chaos, redundantie en verscheidenheid (Nonaka & Takeuchi, 1995; Friedman, Lipshitz & Overmeer, 2001). De organisatie moet in eerste instantie een duidelijke intentie hebben inzake het ontwikkelen van kennis. Individuen en (kleine) teams moeten autonoom kunnen functioneren. Fluctuatie en 'creatieve chaos' zijn noodzakelijk om binnen de organisatie tot een goede interactie te komen. Dit betekent echter niet dat er wanorde moet heersen. Het gaat eerder om een moeilijk te voorspellen orde en om een regelmatige afbraak van vaste denkgewoontes en –patronen. Met de nadruk op een zekere redundantie in de beschikbare informatie wordt beklemtoond dat leden van de organisatie over meer informatie moeten beschikken dan voor hun strikte dagelijkse activiteiten noodzakelijk is. Dat creëert mogelijkheden om met elkaar (persoongebonden) kennis te delen. Tenslotte pleiten Nonaka & Takeuchi (1995) ook voor een voldoende grote verscheidenheid in de

aanwezige informatie. Daartoe moet informatie verschillend, flexibel, en snel te combineren zijn en moeten de leden van de organisatie toegang hebben tot de bronnen van informatie.

We willen deze beschrijving van kenniscreatie afronden met een kritische noot die ook bij het concept 'de lerende organisatie' opdook. Hoe beloftevol de bovenstaande beschrijving ook mag wezen, vooralsnog is weinig geweten over hoe het proces van kenniscreatie zich manifesteert in organisaties en hoe het zich bijvoorbeeld kan inzetten in bureaucratische en weinig beleidseffectieve scholen (Poell & van der Krogt, 1999). Men staat er in (school)organisaties in principe positief tegenover, maar het organiseren ervan blijft een probleem (Eraut, 2000; De Laat e.a., 2001; Simons & Ruijters, 2001).

- *Het meten van de vier soorten van kennisconversie in scholen*

De theorie van de kenniscreatie wordt geoperationaliseerd door na te gaan in welke mate processen binnen de vier onderdelen van de kennisspiraal in scholen plaatsvinden. Voor de vertaling van het theoretisch model naar meetbare items werd deels vertrokken van eigen itemontwikkeling en deels van bestaande literatuur (von Krogh e.a., 2000; Bolhuis & Simons, 2001; Simons, 2003; Chou & Tsai, 2004; Verbiest, 2004).

Tabel 4: Operationalisering van de vormen van kennisconversie

	In welke mate is het een gewoonte in uw school om:
Externaliseren	… over eigen ervaringen te rapporteren? … eigen overtuigingen te expliciteren? … de eigen manier van werken te presenteren aan andere teamleden?
Combineren	… nascholingen bij te wonen? … afspraken te maken over de betekenis van begrippen? … zich door experten te laten coachen? … vakliteratuur te lezen?
Internaliseren	… de ideeën die men op nascholing opdoet in praktijk om te zetten? … de eigen opvattingen aan te passen aan kennis van externen? … al doende te leren?
Socialiseren	… eigen ervaringen te bespreken met andere leden van het team? … zich door collega's te laten coachen? … anderen te observeren tijdens hun werkzaamheden?

Het bovenstaande lijstje zal het derde instrument vormen om het lerend karakter van scholen in kaart te brengen (tezamen met het onderscheid tussen enkele lus en dubbele lus leren en de opvattingen van teamleden over het eigen leren).

Zowel de invalshoek van de school als lerende organisatie en als deze van de organisatie-effectiviteit van scholen werden toegelicht en geoperationaliseerd. We ronden hiermee de beschrijving van de schoolgebonden kenmerken af. De volgende paragraaf richt zich op de zelfevaluatiegebonden kenmerken.

2 META-EVALUATIE ALS DUSDANIG: FOCUS OP ZELFEVALUATIEGEBONDEN KENMERKEN

Ook bij het focussen op de zelfevaluatiegebonden kenmerken gebruiken we verschillende invalshoeken. We creëren eerst een kader om de kernelementen van zelfevaluaties te beschrijven (zie paragraaf 2.1). Daarna operationaliseren we het concept 'kwaliteit van de zelfevaluatie' (zie paragraaf 2.2). Paragraaf 2.3 tenslotte brengt de variabelen in kaart waarvan verwacht wordt dat ze verschillen in de kwaliteit van zelfevaluaties kunnen verklaren.

2.1 Een profiel om zelfevaluaties in scholen te beschrijven

Om het verloop van zelfevaluaties te beschrijven doen we een beroep op de verschillende onderdelen die in de definitie van Petegem (1998, p. 104) onderscheiden worden. Zelfevaluatie is volgens deze omschrijving *'het proces, in hoofdzaak geïnitieerd door de school zelf, waarbij welgekozen participanten op een systematische wijze het functioneren van de school beschrijven en beoordelen met het oog op het nemen van beslissingen c.q. initiatieven in het kader van (aspecten van) de algehele school(beleids)ontwikkeling'*. Deze omschrijving stelt een aantal elementen aan de orde dat het beschrijven van zelfevaluaties in scholen kan structureren. We koppelen deze elementen telkens aan vragen die de variatie in verschijningsvormen van zelfevaluaties kunnen verduidelijken.

- *Zelfevaluatie is een proces*. Wordt de evaluatie opgevat als een permanent en cyclisch proces of eerder als een eenmalig te doorlopen stappenplan? Wordt zelfevaluatie eerder als een middel tot dan als een doel op zich opgevat? Is de zelfevaluatie een onderdeel van een permanent proces van kwaliteitszorg? Ligt de nadruk eerder op wat er met het eindresultaat gebeurt dan wel op dat product zelf?
- *Zelfevaluatie wordt in hoofdzaak geïnitieerd door de school zelf*. Is de stimulans om een zelfevaluatie uit te voeren schoolintern, extern of beide? Voert de school de zelfevaluatie op eigen kracht uit? Heeft zij controle over het concrete verloop ervan?
- *Zelfevaluatie vindt plaats bij welgekozen participanten.* Komt de zelfevaluatie er enkel op initiatief van de schoolleiding? Welke personen en/of groeperingen worden bij de zelfevaluatie betrokken? Zijn er verschillen in betrokkenheid in de verschillende fasen van de zelfevaluatie?
- *Zelfevaluatie gebeurt op systematische wijze*. Op welke manier worden zelfevaluatiegegevens verzameld? Doet de school een beroep op bestaande instrumenten of worden instrumenten zelf ontwikkeld? Is de gegevensverzameling kwalitatief (focusgroepen, fotoreportages, reflectieverslagen) of kwantitatief (vragenlijsten, dataanalyse) van aard? Of worden beide gecombineerd? Zijn er duidelijke procedures? Wordt het niveau van de intuïtieve vaststellingen overstegen? Worden de te evalueren aspecten van het schoolfunctioneren enkel beschreven, enkel beoordeeld dan wel beschreven én beoordeeld? Worden beide fasen van elkaar onderscheiden?

- *Zelfevaluatie heeft betrekking op het functioneren van de school*. Welke aspecten van het schoolfunctioneren worden geëvalueerd? Is de aandacht gericht op variabelen op contextniveau, inputniveau, procesniveau en/of outputniveau (cf. het geïntegreerd CIPO-model van schooleffectiviteit (Scheerens, 1990))? Focust men op het schoolniveau en/of op het leerkracht- en klasniveau?
- *Zelfevaluatie gebeurt met het oog op het nemen van beslissingen en initiatieven in het kader van school(beleids)ontwikkeling*. Komt de zelfevaluatie het functioneren van de school ten goede? Om de mogelijke doelen van zelfevaluaties te ordenen, doen we een beroep op de volgende vier functies van een systeem van evaluatie (van Aanholt & Buis, 1990; Liket, 1992).
    - *Een spiegelfunctie*. Bij deze functie ligt de klemtoon op interne procesontwikkeling. Men wil aan de hand van feedback over de eigen situatie zicht krijgen op het eigen functioneren en dat van collega's.
    - *Een legitimatiefunctie*. Hier ligt de klemtoon op verantwoording naar buiten toe. Men wil naar externen toe (bvb. ten aanzien van de onderwijsinspectie) verantwoording afleggen voor de activiteiten die men onderneemt.
    - *Een transferfunctie*. Het uitgangspunt in deze functie is dat de zelfevaluatie leerzaam kan zijn voor anderen. Men gaat ervan uit dat de processen die zich voltrekken in een bepaalde school of afdeling waardevol kunnen zijn voor andere scholen of afdelingen.
    - *Een beleidsvoorbereidende functie*. Bij deze functie ligt de klemtoon op beleidsontwikkeling. De resultaten van de zelfevaluatie zijn bedoeld om bij te dragen in de totstandkoming van nieuw beleid en de verbetering van bestaand beleid.

    Hoewel deze functies in theorie mooi te onderscheiden zijn, komen ze in realiteit vaak in gemengde vorm voor. Afhankelijk van de invulling van de functies van de zelfevaluatie, zal men tot specifieke soorten van actiepunten komen.

De bovenstaande kenmerken worden in eerste instantie gebruikt om de variatie in zelfevaluaties te beschrijven. Tabel 5 combineert de verschillende kenmerken in een schema dat het mogelijk maakt de belangrijkste kenmerken van een zelfevaluatie in kaart te brengen. De tabel is ingevuld naar het levensechte voorbeeld van een Vlaamse middenschool die in het kader van haar gelijke onderwijskansenbeleid (GOK) een zelfevaluatie uitvoerde (Van Petegem e.a., 2005b). In het vervolg verwijzen we hiernaar als het profiel van een zelfevaluatie.

Tabel 5: De kernkenmerken van een zelfevaluatie

De zelfevaluatie…	nee	**Kadert in een ruimere cyclus**			ja	**Is een eenmalige activiteit**		
Geïnitieerd door…	ja	**De school zelf**			ja	**Een externe overheid**		
Evaluatiefunctie(s)	ja	**Spiegel**	ja	**Legitimatie**	nee	**Transfer**	ja	**Beleidsvoorbereiding**
Participanten aan de zelfevaluatie (betrokken tijdens de planning, de dataverzameling en/of de bespreking van de zelfevaluatieresultaten)			**planning**		**gegevensverzameling**		**bespreking**	
		De directeur	ja		ja		ja	
		Middenkader	ja		ja		ja	
		Leerkrachten	nee		ja		ja	
		Adm. personeel	nee		nee		nee	
		Leerlingen	nee		nee		nee	
		Ouders	nee		nee		nee	
		Schoolbest./IM	nee		nee		nee	
		CLB	nee		ja		ja	
		Ped. beg.	nee		ja		nee	
Methodologische aanpak	ja	Kwantitatief: Survey bij leerkrachten			Ja	Kwalitatief: In groep bespreken van losse items van een zelfevaluatie-instrument		
	ja	Bestaand(e) instrument(en): ICO-SCOOP; GOK-instrument van het ECEGO			nee	Eigen instrument(en):		
Geëvalueerde thema's	Preventie en remediëring van ontwikkelings- en leerachterstanden van leerlingen , socio-emotionele ontwikkeling en intercultureel onderwijs (op school en klasniveau).							

2.2 De 'kwaliteit' van een zelfevaluatie

We stelden hoger dat een zelfevaluatie doelgericht dient te zijn. In die zin zijn de resultaten niet vrijblijvend. De afhankelijke variabele in ons onderzoeksmodel is dan ook de kwaliteit van zelfevaluaties. Dat roept de vraag op wanneer zelfevaluaties nu uiteindelijk kwaliteitsvol zijn. Om het concept 'kwaliteit van de zelfevaluatie' in kaart te brengen, hanteren we een multicriteriabenadering: met een schooleigen en een schooloverstijgende invalshoek. In het uitwerken van beide benaderingen beroepen we ons in sterke mate op de resultaten van de Delphi-studie.

2.2.1 De schooleigen benadering van de kwaliteit van 'zelfevaluaties'

In de probleemstelling bij deze studie definieerden wij (onderwijs)kwaliteit als het op een gepaste manier realiseren van de door belanghebbenden gedeelde verwachtingen. Ook de vraag naar de kwaliteit van zelfevaluaties zal vanuit die invalshoek benaderd worden. Omdat de doelen (cf. verwachtingen) die verschillende scholen vooropstellen bij het opstarten van een zelfevaluatie danig kunnen verschillen impliceert dit dat niet alle kwaliteitsindicatoren a priori vastgelegd kunnen worden. Afhankelijk van de opvattingen van de school zelf en van de context waarin de school functioneert kan kwaliteitszorg (c.q. zelfevaluatie) immers uiteenlopende doelen aannemen (Blok e.a., 2005). Van Petegem (1996, p. 13) schrijft dat het de school zelf is die deze keuze maakt afhankelijk van waar zij op dat moment aan toe is. Meer nog, ook binnen scholen kan er onenigheid be-

staan over de doelen die men met de zelfevaluatie wil bereiken (Van Petegem e.a., 2005b).

Om de schooleigen verwachtingen ten aanzien van de zelfevaluatie in kaart te brengen, doen we een beroep op een inventarisatie van functies en opbrengsten van zelfevaluaties die in de literatuur naar voren worden geschoven (van Aanholt & Buis, 1990; Liket, 1992; Devos e.a., 1999; Davies & Rudd, 2001; Meuret & Morlaix, 2003; Blok e.a., 2005). Dit brengt ons bij de volgende selectie van 10 verwachtingen ten aanzien van zelfevaluatie:

1. Zicht krijgen op sterke en zwakke punten in het handelen van leerkrachten.
2. Zicht krijgen op sterke en zwakke punten in het handelen van de schoolleiding.
3. Komen tot klasgerelateerde verbeterpunten.
4. Komen tot schoolgerelateerde verbeterpunten.
5. Bijdragen tot een grotere betrokkenheid van leerkrachten op het schoolbeleid.
6. Bijdragen tot een prettiger werkklimaat in de school.
7. Voorbereid zijn op een controle door de inspectie.
8. Een concreet probleem oplossen.
9. Het opsporen en delen van goede voorbeelden.
10. In orde zijn met de voorschriften van de overheid.

De verwachte situationele invulling van de verwachtingen ten aanzien van zelfevaluatie hoeft niet te betekenen dat over deze kwaliteit geen uitspraken kunnen gedaan worden. We zullen de schooleigen kwaliteit van zelfevaluaties invullen als de mate waarin de zelfevaluatie aan de door de betrokkenen vooropgestelde verwachtingen voldoet. We brengen daarom eerst in kaart welke verwachtingen een rol speelden bij het initiëren van de zelfevaluatie. Deze verwachtingen worden door de betrokkenen als het ware als een criterium naar voren geschoven om de kwaliteit van de zelfevaluatie aan af te toetsen. We gaan na in welke mate de zelfevaluatie de vooropgestelde verwachtingen ook daadwerkelijk realiseerde. Dit is dan de schooleigen indicator voor de kwaliteit van de zelfevaluatie.

2.2.2 De schooloverstijgende benadering van de kwaliteit van 'zelfevaluaties'

In een kwaliteitsvolle zelfevaluatie worden op basis van de beschrijving en beoordeling van het huidige schoolfunctioneren keuzes gemaakt omtrent op te zetten veranderings- en vernieuwingsactiviteiten (Blok e.a., 2005). Een school verandert haar werking op een aantal punten en selecteert bijvoorbeeld een aantal prioriteiten voor toekomstige acties, voor eventuele nascholing en voor externe begeleiding. De schooloverstijgende benadering bekijkt de kwaliteit van 'zelfevaluaties' vanuit deze invalshoek. Het gaat om de vraag of de zelfevaluatie resulteerde in actiepunten die door de teamleden ter harte worden genomen enerzijds en in een aantal concrete resultaten anderzijds.

a. RESULTEREN IN ACTIEPUNTEN

Uit de Delphi-studie onthouden we dat kwaliteitsvolle zelfevaluaties resulteren in specifieke, meetbare, aanvaardbare, realistische en tijdsgebonden actiepunten. De eerste schooloverstijgende benadering van de kwaliteit van zelfevaluatie gaat daarom na in welke mate de zelfevaluatie resulteerde in:

1. ... duidelijk omschreven actiepunten.
2. ... actiepunten waarvan de realisatie achteraf na te gaan is.
3. ... actiepunten waarvoor een groot draagvlak bestond.
4. ... realistische actiepunten.
5. ... actiepunten met een duidelijke tijdsplanning.

b. CONCRETE RESULTATEN

Zelfevaluatie is echter niet alleen een middel om tot actiepunten te komen. Het uitvoeren van zelfevaluaties kan ook op zich tot concrete waardevolle resultaten leiden. De kwaliteit van zelfevaluaties wordt binnen de schooloverstijgende benadering daarom ook gemeten door de betrokkenen in de school te bevragen over de mate waarin de volgende stellingen volgens hen van toepassing zijn op de zelfevaluatie die heeft plaats gevonden. De vraag is of de zelfevaluatie:

1. ... een impact heeft gehad op het schoolbeleid.
2. ... stimuleerde tot het bespreken van opvattingen over goed onderwijs.
3. ... de wil om kritischer tegen de zaken aan te kijken bevorderde.
4. ... resulteerde in extra motivatie om kwaliteit te leveren.
5. ... de betrokkenheid van teamleden bij het schoolbeleid versterkte.
6. ... heeft bijgedragen tot meer samenwerking in onze school.
7. ... ertoe heeft bijgedragen dat men in de school in sterkere mate dezelfde doelen is gaan nastreven.
8. ... de samenhang in de school versterkte.

Zoals aangegeven kan de kwaliteit van een zelfevaluatie dus op drie manieren geoperationaliseerd worden. In de volgende paragrafen focussen we niet verder op de resultaten van de zelfevaluatie maar op proceskenmerken. We gaan met name op zoek naar de proceskenmerken die ons in staat stellen om verschillen in de kwaliteit van zelfevaluaties te verklaren en te voorspellen.

2.3 Voorspellers van de kwaliteit van zelfevaluaties

Uit de hoger gepresenteerde omschrijving van het begrip ‘zelfevaluatie’ kunnen reeds verscheidene algemeen geldende elementen worden afgeleid die mede bepalend kunnen zijn voor de kwaliteit van zelfevaluaties: is de zelfevaluatie systematisch aangepakt, zijn de participanten welgekozen, wordt het functioneren van de school zowel beschreven als beoordeeld, vloeien er beslissingen voort uit de zelfevaluatie, staan deze beslissingen in het perspectief van de schoolontwikkeling (Van Petegem, 2001). Het zijn precies deze aanzetten tot beoordelingscriteria die verder moeten uitgewerkt worden. Daartoe bewandelen we drie complementaire sporen. Het eerste spoor richt zich op het beschrijven van de houding van het schoolteam ten aanzien van zelfevaluatie, het tweede spoor aanziet het uitvoeren van een zelfevaluatie als een beleidsdaad en het derde spoor benadert zelfevaluatie als een onderzoeksdaad.

2.3.1 De houding van teamleden ten aanzien van zelfevaluatie

Het eerste element dat we als mogelijke voorspeller opnemen, is de houding van de teamleden ten aanzien van zelfevaluatie. Terwijl we hoger wilden komen tot een 'feitelijke' beschrijving van zelfevaluatieaspecten in scholen, staan we hier stil bij de opvattingen van de betrokkenen. Zelfevaluatie kan enkel werken als er daartoe een gepast klimaat bestaat (Macbeath, 1999). '*Men moet al bij al dus niet al te lichtvaardig beginnen aan zelfevaluatie*' (van Aanholt & Buis, 1990, p. 19). Een gunstige voorwaarde om tot een succesvolle zelfevaluatie te komen, is dan ook het realiseren van een bewustwording bij de hele groep leerkrachten van het zinvol aspect van zelfevaluatie.

De houding van schoolleiders ten aanzien van zelfevaluatie is doorgaans positief. Zo vinden Nederlandse schoolleiders het bijvoorbeeld een nuttige en leerrijke onderneming. Zij zijn van mening dat zelfevaluaties resulteren in een betrouwbaar beeld (Blok e.a., 2005). Ze geven aan dat het een haalbare activiteit is maar zijn van mening dat het veel tijd kost. Tot gelijkaardige conclusies komen McBeath et al. (1999) in hun evaluatie van een project waarbinnen een 100-tal Europese scholen een zelfevaluatie uitvoerden. Er zijn evenwel indicaties dat de houding van schoolleiders beduidend positiever is dan deze van leerkrachten. Bij Van Petegem et al. (2005b, p. 288) lezen we bijvoorbeeld dat '*de houding van leerkrachten ten opzichte van zelfevaluatie vaak vrij afwijzend is. Directeurs staan er doorgaans positiever tegenover en zijn meer overtuigd van het nut van zelfevaluatieactiviteiten*'.

Hoewel de houding ten aanzien van zelfevaluatie als dusdanig voorheen reeds enige aandacht kreeg van onderzoekers (Devos e.a., 1999; McBeath e.a., 1999; Blok e.a., 2005) werd de link met de kwaliteit van zelfevaluaties nog niet expliciet onderzocht. In deze studie willen we die link wel leggen. Om de mogelijkheid tot vergelijking te vrijwaren, focussen we op opvattingen met betrekking tot de tegenstellingen over schoolzelfevaluatie die ook in vroegere studies reeds gebruikt werden (Meuret & Morlaix, 2003; Blok e.a., 2005). Deze tegenstellingen zijn:

Zelfevaluatie...		
vertelt ons niets nieuws	⇔	leert ons heel veel
leidt niet tot beter lesgeven	⇔	leidt tot beter lesgeven
kost veel tijd	⇔	kost weinig extra tijd
raakt alleen enkelen	⇔	raakt iedereen
leidt niet tot beter management	⇔	leidt tot beter management
is moeilijk te interpreteren	⇔	is makkelijk te begrijpen
is bij de meeste teamleden niet geliefd	⇔	is geliefd bij de meeste teamleden
is moeilijk uit te voeren	⇔	is goed uit te voeren
is niet kosteneffectief	⇔	is kosteneffectief
is afhankelijk van het toeval	⇔	resulteert in een betrouwbaar beeld
is subjectief	⇔	is objectief
is slechts een momentopname	⇔	geeft onze ontwikkeling goed weer

2.3.2 'Zelfevaluatie als beleidsdaad'

We ontwikkelden hoger een model dat de verschillende dragers van het beleidsvoerend vermogen van scholen in kaart brengt. In dit deel zullen wij beargumenteren dat deze dragers van toepassing zijn op verschillende beleidsterreinen en op verschillende beleidsactiviteiten binnen scholen. Dat geldt bijvoorbeeld voor het realiseren van kwaliteitszorg in het algemeen maar ook voor het uitvoeren van zelfevaluaties. Ook de voorbereidende Delphi-studie leert dat het in voldoende mate tegemoet komen aan de kenmerken van beleidsvoerend vermogen kan vooropgesteld worden als een structurerend kader om het proces van zelfevaluaties in kaart te brengen. Op die manier kunnen de dragers van beleidsvoerend vermogen een vernieuwend licht werpen op het zelfevaluatieproces. Figuur 9 inventariseert de acht dragers die we onderscheiden in het concept 'zelfevaluatie als beleidsdaad'.

Figuur 9: De dragers van 'zelfevaluatie als beleidsdaad'

De verwachting is dat de kwaliteit van een zelfevaluatie in scholen die (doorheen het zelfevaluatieproces) in sterke mate over deze dragers beschikken beter zal zijn dan in scholen die er in mindere mate over beschikken. Het is echter op basis van de nog uit te voeren gegevensverzamelingen dat het belang van de verschillende dragers getoetst zal worden aan empirische evidentie.

In de volgende paragrafen worden de verschillende dragers van 'zelfevaluatie als beleidsdaad' toegelicht. De dragers worden hier dus concreet uitgewerkt in de context van het uitvoeren van een zelfevaluatie. Dat doen we op basis van uitgangspunten en richtlijnen die in raamwerken voor zelfevaluatie als kwaliteitbevorderend naar voren worden geschoven. Ook de ervaringen waarover de onderzoeksliteratuur rapporteert, worden hierin geïntegreerd.

a. DOELTREFFENDE COMMUNICATIE: KWALITEITSVOLLE ZELFEVALUATIE BERUST OP DE MOGELIJKHEDEN OM

DUIDELIJK TE COMMUNICEREN DOORHEEN DE VERSCHILLENDE FASEN VAN HET ZELFEVALUATIEPROCES

Om het verloop van een zelfevaluatie te begrijpen, kan men niet voorbij aan de rol van communicatie. Door communicatie koppelt men de bedoelingen van verschillende betrokkenen aan elkaar. Volgens van Aanholt & Buis (1990, p. 25) wordt de mate waarin niet-direct betrokkenen geïnformeerd zijn bij een zelfevaluatie echter vaak overschat. Zelfevaluaties zullen volgens hen tot betere resultaten leiden als zij - afhankelijk van het doel en de inhoud van de zelfevaluatie - de relevante betrokkenen consulteren en inlichten. Dat vereist het bestaan van gepaste posities en procedures om het communicatieproces in goede banen te leiden. Belangrijke aspecten van het communicatieproces worden met andere woorden doelbewust gestandaardiseerd en geformaliseerd. Duidelijkheid van de te volgen procedures is dan ook één van de randvoorwaarden die Hendriks et al. (2002, p. 135) naar voren schuiven om tot succesvolle zelfevaluatie te komen. Scholen kunnen daarbij bijvoorbeeld een beroep doen op informatieve meetings op studiedagen en personeelsvergaderingen en op het verspreiden van relevante documenten (verslagen van vergaderingen, beschrijvingen van strategieën). Dat komt de effectiviteit van de zelfevaluatie volgens de auteurs ten goede doordat de betrokkenen weten waar ze terecht kunnen met vragen of klachten en doordat men weet welke procedures gevolgd moeten worden. Organisatorische kenmerken van het zelfevaluatieproces moeten tevens garanderen dat de communicatie zowel in de verticale als in de horizontale richting kan verlopen. De rol van informele communicatie tijdens zelfevaluaties mag hierbij niet onderschat worden. In bijvoorbeeld het bespreekbaar maken van moeilijke thema's en in het motiveren van leerkrachten voor het uitvoeren van zelfevaluaties zijn vaak meer mogelijkheden weggelegd voor informele communicatie dan voor formele overlegmomenten. We verwachten dan ook betere resultaten van zelfevaluaties die gebruik maken van een gepaste combinatie van formele en informele informatiekanalen en overlegmethoden. Hetzelfde geldt voor zelfevaluatieprocessen waarin mensen meer open communiceren over hun motieven, ideeën, verzuchtingen en onzekerheden omtrent de zelfevaluatie en de inhoud ervan. Tenslotte speelt ook het opstellen van een zelfevaluatierapport een belangrijke rol. Een goed zelfevaluatierapport is geen uiterst gedetailleerde weergave van alle bevindingen maar geeft op een toegankelijke manier een goede weerspiegeling van de resultaten.

Mogelijke indicatoren van doeltreffende communicatie tijdens een zelfevaluatie zijn de mate waarin scholen er in slagen om iedereen goed te informeren over de doelen van de zelfevaluatie; om duidelijke procedures voorop te stellen om over de zelfevaluatie te communiceren; om leerkrachten zover te krijgen dat ze hun bemerkingen bij de zelfevaluatie kenbaar maken en om met openheid te communiceren over de zelfevaluatieresultaten.

b. ONDERSTEUNENDE RELATIES EN SAMENWERKING: KWALITEITSVOLLE ZELFEVALUATIE BERUST OP EEN VOLDOENDE GROTE PROFESSIONELE EN PERSOONLIJKE DRAAGKRACHT VAN DE SCHOOLLEDEN

Als men van de directie en de leerkrachten van een school verwacht dat ze op een effectieve manier vormgeven aan een zelfevaluatie, dan moet men zich er van verzekeren dat zij daartoe over de nodige draagkracht beschikken (Scheerens e.a., 1999). Draagkracht kan zich maar ontwikkelen en in stand houden als er sprake is van voldoende en brede ondersteuning. We spreken van ondersteunende professionele relaties wanneer men niet het gevoel heeft dat men alleen staat in de uitoefening van zijn zelfevaluatietaak, maar

actief gevolgd en begeleid wordt door collega's (directie en leerkrachten) en/of door instanties zoals pedagogische begeleidingsdiensten of nascholingsinstituten. Wanneer er daarnaast ook (positieve) aandacht is voor de persoonlijke ontwikkeling en persoonlijke ervaringen tijdens de zelfevaluatie is er sprake van ondersteunende persoonlijke relaties.

Collegialiteit speelt een belangrijke rol in het uitvoeren van zelfevaluaties. In een collegiale schoolcultuur wordt samenwerking als een positief iets beschouwd, niet als een bedreiging voor de eigen autonomie van de leerkracht (Verbiest & Vandenberghe, 2002). Zelfevaluaties kunnen maar kwaliteitsvol zijn indien er sprake is van voldoende samenwerking, steun, vertrouwen, openheid en betrokkenheid (Hargreaves & Fullan, 1992; Watling & Arlow, 2002; Scheerens, 2004). Dat kunnen we als volgt begrijpen: vertrouwen en openheid zorgen ervoor dat leerkrachten hun eigen klaspraktijk, eventuele problemen of vragen, maar ook positieve resultaten durven voorleggen aan collega's, directie of andere ondersteunende instanties die bij de zelfevaluatie betrokken zijn (Devos e.a., 1999). Een klimaat van vertrouwen en openheid is dus noodzakelijk (van Aanholt & Buis, 1990; McBeath, 1999).

Om tijdens een zelfevaluatie te komen tot ondersteunende relaties moet aandacht besteed worden aan het bestaan van de nodige functies en procedures. Te verwachten is dat het bestaan van bijvoorbeeld vakwerkgroepen, lerarenraden, intervisiegroepen en vertrouwenspersonen resultaatsbevorderend werkt bij het uitvoeren van zelfevaluatieactiviteiten. Het bestaan van dergelijke structuren is dus zeker waardevol. In het bewerkstelligen van ondersteunende relaties moet evenwel een geschikte balans gevonden worden tussen flexibiliteit enerzijds en standaardisatie en formalisatie anderzijds. Om de ondersteuning voldoende toegankelijk en laagdrempelig te maken, moet deze bijvoorbeeld voldoende flexibel ingericht worden. Daar staat tegenover dat vastgelegde regels, procedures, bevoegdheden ook voor de nodige uniformiteit en continuïteit kunnen zorgen (Stijnen, 1993).

> *Mogelijke indicatoren van ondersteunende relaties tijdens een zelfevaluatieproces zijn de mate waarin scholen er in slagen om als een hecht team samen te werken; om door middel van samenwerking dubbel werk te vermijden; om tijdens de zelfevaluatie in te spelen op de specifieke ondersteuningsbehoeften van individuele leerkrachten; en om tijdens de zelfevaluatie een sfeer van wederzijds vertrouwen tussen de directie en leerkrachten te creëren.*

c. *GEDEELD LEIDERSCHAP: KWALITEITSVOLLE ZELFEVALUATIE BERUST OP DE PARTICIPATIE VAN DE GEPASTE PERSONEN BIJ DE VORMGEVING EN UITVOERING*

De directie speelt onmiskenbaar een centrale rol in het uitvoeren van zelfevaluaties op schoolniveau (Voogt, 1989; Devos & Verhoeven, 2003; Schildkamp, 2007). Ook in het uitvoeren van een zelfevaluatie is leiderschap echter geen zaak van de schoolleiding alleen. Effectieve leiders zijn zich bewust van de noodzaak om in anderen te investeren om zelfevaluatieactiviteiten tot een goed einde te brengen (Harris, 2003). Gezamenlijke besluitvorming creëert een breed draagvlak voor de zelfevaluatie (van Aanholt & Buis, 1990). Effectief leiderschap tijdens de zelfevaluatie bezit dan ook twee componenten; enerzijds het leiderschap van de schoolleider en anderzijds de participatie van leerkrachten in de besluitvorming tijdens de zelfevaluatie (Meuret & Morlaix, 2003). Davies en

Rudd (2001) concluderen uit hun ervaringen in Engelse scholen dat zelfevaluatie veelal een activiteit is van het management van de school, met een beperkte betrokkenheid van leerkrachten, leerlingen en ouders. Van Aanholt en Buis (1990, p. 21) pleiten voor 'power-sharing'; in die zin dat de schoolleiding om het zelfevaluatieproces te doen slagen bereid moet zijn een deel van haar taken te delegeren. Dat kan bijvoorbeeld door het uitvoeren van de voorbereidende fase toe te vertrouwen aan een stuurgroep. Deze voorbereidende werkgroep dient zeer doordacht te worden samengesteld (van Aanholt & Buis, 1990; Meuret & Morlaix, 2003). De school moet hierin trachten een voor haar gepaste balans te vinden tussen vrijwilligheid, representativiteit, charisma en eventueel (in)formele hiërarchie. Albrechts & Caals (2003) benadrukken dat de samenstelling een verticale en horizontale doorlichting van de organisatie moet mogelijk maken. Daarnaast geven ze aan dat men best niet alleen voortrekkers van de zelfevaluatie opneemt. Precies door ook twijfelaars te betrekken, wordt vaak een (kleine maar) belangrijke groep over de streep getrokken. Uit een analyse van zelfevaluatieprocessen blijkt dat het benoemen van een stuurgroep sterk samenhangt met het gepercipieerde succes van de zelfevaluatie (Davies & Rudd, 2001).

Gedeeld leiderschap betekent niet dat alle leerkrachten bij alle beslissingen en alle fasen van het zelfevaluatieproces betrokken moeten worden, maar wel dat ze waar nodig een rol spelen. Wanneer een zelfevaluatie zich richt op aspecten waar schoolleden niet onverschillig tegenover staan, moet hun betrokkenheid bij het zelfevaluatieproces overwogen worden (Mace, 2002). Om na te gaan in hoeverre verschillende teamleden betrokken zouden moeten worden tijdens de zelfevaluatie kunnen we een beroep doen op de kwadranten van gedeeld leiderschap van Hoy en Miskel (2003).

Tabel 6: Kwadranten van gedeeld leiderschap (Naar Hoy & Miskel, 2003)

		Is er sprake van een persoonlijk belang ?	
		Ja	Nee
Is er sprake van expertise?	Ja	I. Buiten de zone van onverschilligheid **Democratische of conflictsituatie** *(participatie)*	II. In de rand van de zone van onverschilligheid **Expertsituatie** *(occasionele participatie)*
	Nee	III. In de rand van de zone van onverschilligheid **Stakeholderssituatie** *(occasionele participatie)*	IV. Binnen de zone van onverschilligheid **Niet-collaboratieve situatie** *(geen participatie)*

Naargelang het kwadrant waarin de thema's van een zelfevaluatie zich bevinden, zal de invulling van betrokkenheid van schoolleden verschillen. In het eerste en het laatste kwadrant is de situatie doorgaans duidelijk. Wanneer er bijvoorbeeld sprake is van sterke persoonlijke belangen en sterke expertise wordt er best participatie in de zelfevaluatie nagestreefd. De situatie in de kwadranten 'in de rand van de zone van onverschilligheid' ligt enigszins moeilijker. Betrokkenheid van alle teamleden is er niet noodzakelijk het meest effectief. In de stakeholderssituatie bestaat immers het gevaar dat de personeelsleden inhoudelijk weinig kunnen inbrengen tijdens de zelfevaluatie. Participatie

dreigt door de betrokkenen in dat geval als een 'lege doos' te worden ervaren omdat de wezenlijke inbreng in de zelfevaluatie toch bij anderen, namelijk de experts, berust. In dit geval is participatie enkel aangewezen als de zelfevaluatieresultaten voldoende draagkracht moeten hebben en men een open communicatie wil verzekeren. Ook in de expertsituatie is participatie niet vanzelfsprekend. Participatie van experts mag dan de kansen op een kwaliteitsvolle beslissing wel vergroten, ook de kans op vervreemding en het uit het oog verliezen van de relevantie neemt toe. Uit het bovenstaande blijkt dat een grote participatie niet in alle zelfevaluatiecontexten aangewezen is en dat er in bepaalde omstandigheden beter gepleit wordt voor een occasionele en beperkte vorm van participatie. Het voorgaande impliceert echter niet dat men moet berusten in een situatie van beperkte betrokkenheid en beperkte expertise. Verbiest & Vandenberghe (2002) pleiten voor een cultuur van 'empowerment' zodat de kennis en belangstelling van alle personeelsleden kan toenemen. Dit om te vermijden dat enerzijds steeds dezelfde belanghebbenden of belangstellenden en deskundigen zich engageren, anderzijds om te vermijden dat niet-geïnteresseerden een bepaald beleid - al dan niet opzettelijk - gaan boycotten.

Mogelijke indicatoren van het participatieve karakter van de zelfevaluatie zijn de mate waarin scholen er in slagen om de verantwoordelijkheden voor het uitvoeren van de zelfevaluatie niet door één enkele persoon te laten dragen; om alle belanghebbenden bij de zelfevaluatie te betrekken; om de in de school beschikbare expertise aan te wenden om de zelfevaluatie tot een goed einde te brengen; en om bij het verdelen van zelfevaluatietaken rekening te houden met datgene waar de verschillende verantwoordelijken goed in zijn.

d. GEZAMENLIJKE DOELGERICHTHEID: KWALITEITSVOLLE ZELFEVALUATIE BERUST OP DUIDELIJK VOOROPGESTELDE DOELEN DIE DOOR DE BETROKKENEN GEDEELD WORDEN

Een primordiale conditie voor zelfevaluatie is het bestaan van een door het hele schoolteam gedragen behoefte om het huidige functioneren van de school te onderzoeken (Hendriks et al., 2002). Zelfevaluaties moeten dan ook plaats vinden met een duidelijk en eerlijk doel voor ogen (McBeath e.a., 1999; Schildkamp, 2007). Het is belangrijk dat duidelijk wordt gemaakt waarom er een zelfevaluatie plaatsvindt en waarvoor de resultaten zullen dienen (Nevo, 1995). Daarbij kunnen zowel onderwijskundige, pragmatische, politieke als juridische elementen naar voren worden geschoven. Om het belang van deze elementen correct te kunnen inschatten, is het van belang stil te staan bij wat de verschillende betrokkenen aan de zelfevaluatie willen en zullen overhouden. De behoeften van de verschillende stakeholders zijn immers niet noodzakelijk in overeenstemming. De prioriteiten van een leerkracht zijn niet altijd gelijklopend met deze van het schoolbestuur, ook al hoeven ze elkaar niet uit te sluiten. Bij de start van een zelfevaluatie komt het er op aan om de verschillende stakeholders tot overeenstemming te brengen omtrent de functie en doelen van de zelfevaluatie (Scheerens, 2004). Ook de vraag of een zelfevaluatie wel het geëigende middel is om deze doelen te bereiken is hier aan de orde (van Aanholt & Buis, 1990).

Een kwaliteitsvol zelfevaluatieproces start dus met een reflectie omtrent de vragen 'Wie is de opdrachtgever?' en 'Waarom doen wij een zelfevaluatie?'. Dergelijke reflectie resulteert in een heldere doelgerichtheid die impliceert dat men precies weet waar men met

de zelfevaluatie naartoe wil. Het antwoord op beide vragen is van belang voor de vormgeving aan het verdere verloop van de zelfevaluatie. Macbeath (1999, p. 106) stelt: 'first, start with the end in mind'.

Kenmerkend voor doelgerichte zelfevaluaties is verder dat men voortdurend de relatie bewaakt van de zelfevaluatieactiviteiten met de doelen die men met de zelfevaluatie wenst te bereiken. Die zelfevaluatiedoelen moeten niet alleen duidelijk geformuleerd worden. Ze dienen ook aan alle betrokkenen bekend gemaakt te worden (van Aanholt & Buis, 1990). Meer nog: het is van belang dat de betrokkenen de doelen onderschrijven en dat er dus sprake is van een gezamenlijke doelgerichtheid (Hendriks & Bosker, 2003). De kans dat scholen succesvol vorm geven aan een zelfevaluatie, is mede afhankelijk van de mate waarin ze erin slagen om als gemeenschap te komen tot een gezamenlijke doelgerichtheid voor die zelfevaluatie (Kyriakides & Campbell, 2004). De mate waarin iemand activiteiten als waardevol en belangrijk ervaart is een goede voorspeller voor de mate waarin die persoon gemotiveerd zal zijn om aan die activiteiten mee te werken (Devos e.a., 1999; McBeath, 1999). Bijvoorbeeld: voeren leerkrachten de zelfevaluatie enkel uit om administratief in orde te zijn of zien ze er ook voor zichzelf ontwikkelingsmogelijkheden in? Om tot een kwaliteitsvolle zelfevaluatie te komen moeten leerkrachten het nut van de zelfevaluatie inzien, zowel voor de school als voor zichzelf. Want, van zodra zelfevaluatie wordt beschouwd als een last, of als door externen opgelegd, zal het aan impact inboeten (Macbeath, 1999, p.108). Zelfevaluatie is dus niet sowieso goed, louter omdat het moet en iedereen het doet (Bunt & Van Hees, 1991).

Mogelijke indicatoren van de gezamenlijke doelgerichtheid van de zelfevaluatie zijn de mate waarin scholen er in slagen om voor de zelfevaluatie duidelijke doelen voorop te stellen; om alle leerkrachten achter de vooropgestelde zelfevaluatiedoelen te scharen; om tussen leerkrachten overeenstemming te bereiken over de noodzaak van de zelfevaluatie; om tussen leerkrachten overeenstemming te bereiken over prioritair aan te pakken knelpunten en om leerkrachten te sensibiliseren voor het uitvoeren van de zelfevaluatie.

e. RESPONSIEF VERMOGEN: KWALITEITSVOLLE ZELFEVALUATIE BERUST OP HET VERMOGEN VAN SCHOLEN OM DE VRAGEN EN VERWACHTINGEN VAN DE OMGEVING TE ONDERKENNEN EN EEN PLAATS TE GEVEN IN HET ZELFEVALUATIEPROCES

Van scholen wordt verwacht dat ze tijdens zelfevaluaties inspelen op vragen en verwachtingen vanuit de verschillende soorten omgevingen waarmee ze in interactie treden. Hoger werden vier soorten omgevingen onderscheiden (1) de bestuurlijke omgeving, (2) de sociale omgeving, (3) de professionele omgeving en (4) de competitieve omgeving (Hooge, 1998). De aard van de zelfevaluatie die Vlaamse scholen in het kader van het gelijke onderwijskansenbeleid dienden uit te voeren, is illustratief voor het belang van elk van deze vier soorten omgevingen. De bestuurlijke omgeving stelt een aantal administratieve en inhoudelijke richtlijnen voorop. Het centrale thema van de zelfevaluatie kan niet worden losgekoppeld van de sociale omgeving waarin de school gevestigd is en meer bepaald van alle buurtbewoners en –organisaties die mogelijk betrokken zijn bij het schoolgebeuren. Om hun beleid vorm te geven moeten scholen verder op een gepaste manier inspelen op hun professionele omgeving (scholen waarnaar de leerlingen doorstromen, buitengewoon onderwijs, samenwerkingsverbanden tussen scholen, lokale over-

legplatforms, en dergelijke) maar ook op hun competitieve omgeving. Responsieve scholen verliezen de invloed van deze verschillende omgevingen niet uit het oog bij het uitvoeren van een zelfevaluatie. Ze informeren zich eerst over de eisen en verwachtingen van hun omgeving en tekenen daarna mogelijke responsen uit. De resultaten van een zelfevaluatie zullen dan ook wel varen bij een open en toegankelijke cultuur waarin directie en leerkrachten beschikbaar zijn voor leerlingen, ouders en andere actoren uit de schoolomgeving. Dat vergt ook de juiste posities en procedures in scholen om de betrokkenheid van de omgeving te promoten. Indien mogelijk worden bijvoorbeeld leerlingenraden, ouderraden, scholenoverleg e.d. uitgenodigd een rol te spelen in de zelfevaluatie. Ook het betrekken van externen kan overwogen worden. Het aanstellen van een zogenaamde 'kritische vriend' kan bijvoorbeeld bijdragen tot objectiviteit en steun (Macbeath, 1999). Externen kunnen ervaringen uit andere scholen en vroegere zelfevaluatieprocessen inbrengen. Hun inbreng behoort weliswaar ondersteunend en niet leidend te worden ingevuld (van Aanholt & Buis, 1990; Devos & Verhoeven, 2003).

Mogelijke indicatoren van het responsieve karakter van de zelfevaluatie zijn de mate waarin scholen er in slagen om ouders te betrekken bij de gegevensverzameling; om de contacten met de pedagogische begeleiding zo in te richten dat ze de kwaliteit van zelfevaluaties ten goede komen; om zo nodig bij externen te rade te gaan; om tijdens de zelfevaluatie rekening te houden met algemeen maatschappelijke thema's en om tijdens de zelfevaluatie actief op zoek te gaan naar de mogelijke inbreng van de lokale gemeenschap.

f. INNOVATIEF VERMOGEN: KWALITEITSVOLLE ZELFEVALUATIE BERUST OP HET VERMOGEN VAN SCHOLEN OM AFSTAND TE NEMEN VAN HET TRADITIONELE EN OPENINGEN TE CREËREN VOOR VERNIEUWING

Een door innovatie geïnspireerde zelfevaluatie berust op een aanpak die getuigt van openheid naar nieuwe ideeën en werkwijzen en van een blik die niet enkel op het huidige functioneren is gericht, maar tevens op de toekomst. Dat impliceert dat er gewerkt wordt met methodieken die de inbreng van vernieuwende invalshoeken, argumenten en ideeën mogelijk maken (van Aanholt & Buis, 1990). In een innovatieve zelfevaluatie spoort de directie de leerkrachten aan om nieuwe dingen en toekomstige ontwikkelingen in de zelfevaluatieactiviteiten op te nemen. Zulk een aanpak biedt meer kansen om te leren (Güthe, 1997) en wordt bijvoorbeeld gekenmerkt door flexibiliteit, door een onderzoekende houding, en door een bereidheid om te experimenteren en om doelen en middelen in vraag te stellen (Argyris & Schön, 1996). Bij het beschrijven en beoordelen van de onderzochte thema's durft men afwijken van de traditionele referentiekaders. De zelfevaluatie gaat evenwel gepaard met voorschriften die het vernieuwende gedrag van directie en leerkrachten leiden en die de marge aangeven waarbinnen zulke alternatieven en experimenten mogelijk zijn.

Om vanuit een zelfevaluatie tot gepaste handelingsplannen te komen, moet een school het aandurven om vernieuwingen en eventuele experimenten in het beleidsvoeringsproces als alternatief naar voren te schuiven. Ook bij het maken van deze keuzes en het vooropstellen van prioriteiten is een innoverend kader bijgevolg van belang (Schildkamp, 2007). Een kwaliteitsvolle zelfevaluatie resulteert immers in beargumenteerde keuzes omtrent op te zetten vernieuwingsactiviteiten. Het gaat daarbij niet om vernieuwingen om de vernieuwing maar om innovaties die aansluiten bij de bestaande basisvisie van de

school. Vernieuwingen maken immers geen kans zonder de steun van diegenen die de vernieuwing in de praktijk moeten realiseren (Geijsel, 2001).

Mogelijke indicatoren van het innovatieve karakter van de zelfevaluatie zijn de mate waarin scholen er in slagen om tijdens de zelfevaluatie vernieuwende invalshoeken te gebruiken, om open te staan voor nieuwe vormen van gegevensverzameling, om af te stappen van bestaande denkkaders en om te experimenteren met verschillende vormen van zelfevaluatie.

g. EEN GEÏNTEGREERD BELEID: KWALITEITSVOLLE ZELFEVALUATIE BERUST OP HET VERMOGEN VAN SCHOLEN OM (DE THEMA'S VAN) DE ZELFEVALUATIE OP EEN INTEGRALE EN GEÏNTEGREERDE MANIER TE BENADEREN

In een kwaliteitsvolle zelfevaluatie worden de te bestuderen schoolaspecten op een integrale en geïntegreerde manier benaderd. Daarmee bedoelen we dat de thema's van de zelfevaluatie bestudeerd worden in het licht van het globale schoolbeleid en dat deze thema's geëvalueerd worden in hun onderlinge samenhang met andere aspecten van het schoolgebeuren (Mace, 2002). Een kwaliteitsvolle zelfevaluatie plaatst de te bevragen thema's in een ruimer kader en benadert ze vanuit een visie die maakt dat er aandacht is voor het hele schoolgebeuren. Devos et al. (1999) concluderen uit een literatuuronderzoek dat de doelstellingen van zelfevaluatie best gerelateerd worden aan de doelen van het onderwijs in het algemeen. Zelfevaluatie is volgens hun bevindingen best geen op zichzelf staande activiteit maar behoort deel uit te maken van het gehele schoolbeleid. De deelnemers aan de zelfevaluatie zien hun inbreng dan ook best als een noodzakelijk onderdeel van een groter (integraal) geheel van onderling geïntegreerde delen. Zelfevaluaties richten zich bovendien best niet alleen op een ruime invulling van het huidige functioneren, ook de toekomstgerichte invalshoek is van belang. Dat neemt niet weg dat scholen om praktische redenen bij het uitvoeren van een zelfevaluatie een balans moeten zoeken tussen de werkbaarheid enerzijds en de mate van integraliteit en geïntegreerdheid anderzijds. Het dagelijkse werk dat een zelfevaluatie met zich meebrengt moet dan ook verdeeld worden vanuit een 'functioneel evenwicht' (Sleegers, 1991). Niet iedereen kan immers tegelijkertijd met alles bezig zijn. M.a.w. er moet tijdens de zelfevaluatie sprake zijn van een zekere vorm van specialisatie, maar tegelijkertijd moeten directie en leerkrachten en eventueel andere betrokkenen op een complementaire manier samenwerken.

Mogelijke indicatoren van het geïntegreerde karakter van de zelfevaluatie zijn de mate waarin scholen er in slagen om de doelen van de zelfevaluatie te linken aan andere initiatieven; om de thema's van de zelfevaluatie te bestuderen in relatie tot andere aspecten van het schoolbeleid; om tijdens de zelfevaluatie aandacht te besteden aan zowel beleidsmatige als onderwijskundige thema's; om tijdens de zelfevaluatie gebruik te maken van bestaande raden/werkgroepen en om de werkgroepen die bij de zelfevaluatie betrokken zijn op een complementaire manier te doen samenwerken.

h. REFLECTIEF VERMOGEN: KWALITEITSVOLLE ZELFEVALUATIE BERUST OP HET VERMOGEN VAN LEERKRACHTEN EN SCHOLEN OM ZICHZELF OP EEN SYSTEMATISCHE EN KRITISCHE MANIER IN VRAAG TE STELLEN

Het reflectief vermogen van scholen is wellicht de meest cruciale factor om tot een kwaliteitsvolle zelfevaluatie te komen. Zelfevaluatie zou immers kunnen omschreven worden als een systematische vorm van reflectie. Om een zelfevaluatie tot een goed einde te brengen, is de bereidwilligheid om in het eigen handelen zowel positieve als negatieve

elementen te ontdekken en te analyseren dan ook van essentieel belang (Voogt, 1989; Devos e.a., 1999; Hendriks e.a., 2002). Zelfevaluaties zouden tot betere resultaten leiden wanneer het stilstaan bij de eigen praktijk en het identificeren van verbeterpunten niet als bedreigend worden ervaren. De veronderstelling is immers dat de kwaliteit van zelfevaluaties samenhangt met de mate waarin een schoolteam getuigt van een reflectieve praktijk, gaande van een reflectieve benadering van het individuele handelen, van het handelen met collega's, van het handelen in team én het handelen als school. Devos et al. (1999) verwijzen echter naar onderzoek waaruit blijkt dat leerkrachten zich bedreigd en onbehaaglijk voelen bij een evaluatie. Reflecteren en - bij uitbreiding - het uitvoeren van zelfevaluaties wordt nog vaak als een plichtsoefening gepercipieerd en niet als een noodzaak om het eigen functioneren te verbeteren (Van Petegem e.a., 2005b).

Mogelijke indicatoren van het reflectieve karakter van de zelfevaluatie zijn de mate waarin scholen er in slagen om de schoolleden te overtuigen van het belang van regelmatige zelfevaluatie; om leerkrachten bereid te maken om kritisch naar hun eigen praktijk te kijken en om leerkrachten een positieve houding te doen aannemen tegenover zelfevaluatie op schoolniveau.

Hierboven werd voor de acht dragers van 'zelfevaluatie als beleidsdaad' een beschrijving gegeven. Onze verwachting is dat zelfevaluaties die sterk verschillen wat deze dragers betreft ook zullen verschillen in de mate waarop ze scoren op de kwaliteitsindicatoren. Het uitvoeren van kwaliteitsvolle zelfevaluaties vergt in die logica dus een beleidsmatige aanpak die zich laat kenmerken door de dragers van beleidsvoerend vermogen. Naast deze beleidsmatige aanpak schuiven we in dit theoretisch kader ook een onderzoeksmatige aanpak naar voren. De invalshoek van 'zelfevaluatie als onderzoeksdaad' wordt in de volgende paragraaf uitgewerkt.

2.3.3 Zelfevaluatie als onderzoeksdaad

Essentieel voor een zelfevaluatie is het systematische karakter ervan (Van Petegem, 1998). Enkel voldoende objectieve en zorgvuldig verzamelde gegevens zijn bruikbaar. Het niveau van toevallig opgedane indrukken moet bij systematische zelfevaluatie overstegen worden (Devos e.a., 1999). Niet voor niets spreken Blok et al. (2005, p. 13) van de opvattingen over uitvoering van sociaalwetenschappelijk onderzoek als bron voor kwaliteitsindicatoren voor zelfevaluaties. In het volgende presenteren we het uitvoeren van een zelfevaluatie daarom als een onderzoeksdaad. We onderscheiden daartoe de drie stappen die men doorgaans in een onderzoek onderscheidt (Janssens, 1991; Creswell, 2005): (a) het stellen van een vraag, (b) systematische dataverzameling en (c) het beantwoorden van de vraag. In de laatste paragraaf wordt toegelicht hoe de verschillende aspecten van 'zelfevaluatie als onderzoeksdaad' in het surveyonderzoek geoperationaliseerd worden.

a. HET STELLEN VAN EEN VRAAG: INHOUDELIJKE VERANKERING, RELEVANTIE EN PRECISIE

Doorgaans is de vraag naar de inhoud van een zelfevaluatie het uitgangspunt bij de vormgeving aan een zelfevaluatieproces in het algemeen en bij het kiezen van een zelfevaluatie-instrument in het bijzonder. Nevo (1995) stelt dat zowat alles het voorwerp kan zijn

van een zelfevaluatie. De verantwoording van gemaakte inhoudelijke keuzes is momenteel echter teleurstellend (Blok e.a., 2005, p. 29). Het zou met name vaak ontbreken aan een beredeneerd en beargumenteerd antwoord op de vraag naar de inhoud(en) van de zelfevaluatie. Dat is deels toe te schrijven aan onduidelijkheid omtrent welke aspecten van het schoolfunctioneren de moeite waard zijn om in een zelfevaluatie te worden opgenomen. Hier lijkt dan ook een probleem te rijzen: er bestaat immers geen algemeen aanvaarde consensus omtrent wat al dan niet en in welke mate bijdraagt tot een effectieve schoolwerking en (dus) tot een hogere kwaliteit van het verstrekte onderwijs (Van Petegem, 1998; Macbeath, 1999). Vanuit het perspectief van de meta-evaluatie is een dergelijke consensus echter niet noodzakelijk. Meta-evaluatie omhelst geen beschrijving en beoordeling van de gemaakte keuzes in se, maar richt zich op de argumentatie van de school om de gemaakte keuzes te ondersteunen. Wat ons interesseert met het oog op het voorspellen van de kwaliteit van een zelfevaluatie is de manier waarop een school tot de selectie van te evalueren themata komt. We gebruiken in het beoordelen hiervan drie criteria: inhoudelijke verankering, relevantie en precisie (Oost & Markenhof, 2002b).

- *Inhoudelijke verankering*

Inhoudelijke verankering houdt in dat de school de te evalueren thema's kan situeren in een ruimer kader: een schoolwerkplan, het pedagogisch project of een theoretisch kader. Bij dit laatste denken wij bijvoorbeeld aan de bevindingen van schooleffectiviteitsonderzoek (Van Petegem, 2005) of concepten zoals de lerende organisatie. Ook bestaande zelfevaluatie-instrumenten en elders door de overheid ontwikkelde kaders voor de (zelf)evaluatie van scholen kunnen een verscheidenheid aan goede aanknopingspunten bieden. Macbeath (1999, p. 112) benadrukt dat het niet aangewezen is bestaande kaders integraal over te nemen. Scholen kunnen bijvoorbeeld ook bij de eigen stakeholders op zoek gaan naar elementen waarvan vermoed wordt dat ze de moeite zijn om in een zelfevaluatieproces op te nemen (Macbeath, 1999). Het is uiteindelijk de school zelf die dient uit te maken wat zij verstaat onder zowel een 'goede' school en wat daar volgens haar toe bijdraagt (Van Petegem, 1998, p.112). Deze inzichten worden best gesitueerd in een gebruiksvriendelijk denkkader met bijbehorend begrippenkader (van Aanholt & Buis, 1990; Van Petegem, 1998; McBeath, 1999). Reezigt & de Jong (2001) benadrukken bijvoorbeeld dat schoolverbeteringsprocessen effectiever zijn als ze op een duidelijke (onderwijskundige) theorie zijn gebaseerd.

- *Relevantie*

Relevantie verwijst naar de mate waarin de school aannemelijk weet te maken waarom de zelfevaluatie uitgevoerd zou moeten worden. Dit houdt in dat bijvoorbeeld wordt aangetoond dat er omtrent het thema van de zelfevaluatie vragen leven bij leden van het schoolteam en dat het de moeite waard is dat deze vragen beantwoord worden. De teamleden moeten de thema's van de zelfevaluatie voldoende belangrijk vinden. Om tot een gepaste selectie van zelfevaluatiethema's te komen, worden antwoorden op de wat-vraag middels discussie en negotiatie best aan de specificiteit van de eigen schoolcontext aangepast (Blok e.a., 2005).

➢ *Precisie*

Precisie tenslotte verwijst naar de nauwkeurigheid waarmee aangegeven wordt waar men naar op zoek is. Vaagheid is uit den boze (Swanborn, 1991).

b. DE DOELMATIGE DATAVERZAMELING

Zelfevaluaties vereisen een duidelijke strategie die het mogelijk maakt om op een gerichte manier gegevens te verzamelen die de bestaande praktijken beschrijven. Methodologisch heeft een school daartoe verschillende technieken ter beschikking. Te denken valt aan vragenlijsten, functioneringsgesprekken, discussierondes, observaties, workshops en interviews (Van Petegem, 1996, p. 10). Oplossingen in de vorm van zelfevaluatie-instrumenten zijn dan ook maar één van de vele mogelijkheden om te komen tot een beschrijving van het schoolfunctioneren. Meuret & Morlaix (2003) komen tot de bevinding dat de ene methode niet systematisch te verkiezen is boven de andere. Zelfevaluatiegegevens worden evenwel niet lukraak verzameld. De gegevensverzameling dient doelmatig te zijn. Dat roept de vraag op of ze tot informatie leidt die geschikt is voor het doel waarvoor de zelfevaluatie uitgevoerd wordt. De doelmatigheid van een gegevensverzameling blijkt uit de mate waarin de beschrijving van de zelfevaluatiethema's op een betrouwbare, valide en efficiënte manier plaatsvindt.

➢ *Bewaken van de betrouwbaarheid*

Als men de zwakke punten van een organisatie wil ontdekken, dan is het noodzakelijk dat de evaluatie een objectieve weerslag is van de situatie (Albrechts & Caals, 2003, p. 24). Betrouwbaarheid is belangrijk omdat de beslissingen die naar aanleiding van de zelfevaluatie genomen worden, gebaseerd moeten zijn op gegevens die niet afhangen van allerlei toevalligheden. Dat houdt in dat de verzamelde gegevens voldoende consistent, nauwkeurig en reproduceerbaar behoren te zijn. Kortom: vrij van meetfouten, niet beïnvloed door storende niet-relevante factoren. Storende factoren kunnen betrekking hebben op het zelfevaluatie-instrument, de evaluatiecontext, de respondent of de persoon die gegevens verzamelt. Daartoe zijn in de context van een zelfevaluatie een aantal aandachtspunten van belang. Volgende vragen plaatsen een aantal kritische bedenkingen bij de betrouwbaarheid van een bepaalde zelfevaluatiemethode.

- Wat het zelfevaluatie-instrument betreft: Zijn de vragen duidelijk geformuleerd zodat ze niet dubbelzinnig geïnterpreteerd kunnen worden? Zijn de vragen voldoende eenvoudig te begrijpen? Gebuikt men triangulaire vragenlijsten (Tjio & Linssen, 2003) (vragenlijsten die zoveel mogelijk gelijke vragen bevatten voor leerkrachten, directieleden, ouders en leerlingen)?
- Wat de evaluatiesituatie en de respondent betreft: Hebben het tijdsstip en de duur van de evaluatie een invloed op de resultaten? Hoe zit het met de tijdsdruk en de evaluatieomgeving (geluidsoverlast, concentratieproblemen)? Zijn de respondenten voldoende gemotiveerd om de gegevensverzameling tot een goed einde te brengen? Is men bereid oprechte informatie te verschaffen? Wordt er zo nodig anoniem tewerk gegaan? Vertrouwelijkheid is immers noodzakelijk om te vermij-

den dat mensen zich geremd voelen in het verwoorden van opinies of dat zij zich persoonlijk het object van kritische bemerkingen voelen (Vleminckx & Van Nuland, 1998; McBeath, 1999).

- Wat de beoordelaar(s) betreft (voornamelijk in het geval waarin men kwalitatieve manieren van gegevensverzameling gebruikt): Scoren verschillende mensen een vaststelling op een gelijkaardige manier? Laten ze niet relevante persoonsinformatie buiten beschouwing? Gaat men nauwkeurig en objectief tewerk of spelen subjectieve interpretaties en vooroordelen een rol? Wordt het beschrijven en beoordelen van elkaar onderscheiden?

Een evaluatie kan niet valide zijn zonder dat ze betrouwbaar is. Betrouwbaarheid is dus een noodzakelijke maar geen voldoende voorwaarde om van validiteit te kunnen spreken (Airasian, 1994; McMillan, 2001).

➢ *Bewaken van de validiteit*

De validiteit van een evaluatiemethode doelt op de eigenschap dat een evaluatiemethode al dan niet meet wat de gebruiker ervan bedoeld heeft ermee te meten. Anders geformuleerd: men moet de juiste conclusies kunnen trekken uit de gegevens die men doorheen de zelfevaluatie verzamelt. Om dat te verzekeren moet die gegevensverzameling aan twee vormen van validiteit voldoen: Inhoudsvaliditeit en begripsvaliditeit.

- In eerste instantie moet de gegevensverzameling inhoudsvalide zijn. Dat betekent dat ze op een representatieve manier informatie oplevert over de vooropgestelde thema's van de zelfevaluatie. Dat houdt bijvoorbeeld in dat de vragenlijst een goede representatie vormt van alle mogelijke relevante gegevens bij de vooropgestelde onderzoeksvraag (Smaling, 1987, p. 280). In de evaluatieliteratuur spreekt men in dat geval van doelstellingenrepresentativiteit (Van Petegem & Vanhoof, 2002).
- Men spreekt van begripsvaliditeit als kan worden aangetoond dat het zelfevaluatie-instrument effectief de door de school beoogde thema's meet. De vraag is hier wat de items in de vragenlijst precies betekenen en of ze treffende operationalisaties zijn van het te meten construct. Hier gaat het om de vraag of de items meten we wat ze willen meten (Swanborn, 1991, p. 27).

Scholen die met een bestaand zelfevaluatie-instrument werken, moeten goed bewaken dat dit instrument voldoende aansluit bij de bedoelde inhoud van de zelfevaluatie (cf. inhoudsvaliditeit). Bestaande instrumenten hebben doorgaans wel als voordeel dat de begripsvaliditeit gewaarborgd is. Wanneer scholen met eigen instrumenten werken, geldt wellicht het omgekeerde. Er is meer flexibiliteit in het kiezen van gepaste inhouden maar de begripsvaliditeit is moeilijk te bewaken. Hieronder beschrijven we zowel een aantal aandachtpunten bij het selecteren als bij het zelf ontwikkelen van een zelfevaluatie-instrument.

- Aandachtspunten bij het selecteren van een zelfevaluatie-instrument. Het aanbod van zelfevaluatie-instrumenten is momenteel bijzonder groot (Cremers-Van Wees, Rekveld, Brandsma & Bosker, 1996; Hendriks, 2000; Duurkoop, Majoor & de Pinth, 2002). Deze instrumenten richten zich op verschillende onderwijsniveaus, hebben een algemene dan wel specifieke focus, richten zich op verscheidene aspecten van het schoolfunctioneren, nemen verschillende vormen aan en hanteren uiteenlopende methodieken. Uit die veelheid van instrumenten moet de school een keuze maken. Het is aan de school zelf om uit te maken of een bepaalde methode van zelfevaluatie inderdaad zoden aan de dijk kan helpen brengen (Gombeir, 1998, p. 13). Die keuze moet ingegeven zijn door beredeneerde argumenten; het simpelweg gebruiken van een instrument dat ter beschikking is zonder de geschiktheid ervan te beoordelen, hypothekeert de bruikbaarheid van de resultaten van de zelfevaluatie. Achter elk bestaand zelfevaluatie-instrument gaan immers visies en invalshoeken schuil (Devos, Verhoeven, Opbrouck & Verbeeck, 2000). Een school moet dus allereerst overwegen of ze de achterliggende concepten wil hanteren en of ze de in het instrument aangereikte weg ook echt wil inslaan (Gombeir, 1998). Deze afweging van de conceptuele bruikbaarheid van het instrument is dan ook een eerste belangrijke stap in het selecteren van een zelfevaluatie-instrument. Daarnaast moet het instrument dekkend zijn. Dat betekent dat het instrument inhoudelijk moet aansluiten bij de doelen van de zelfevaluatie (Blok e.a., 2005). Het moet zich richten op de te evalueren aspecten van het schoolfunctioneren. Naast inhoudelijk moet het instrument ook methodologisch dekkend zijn. Daarbij staat de vraag of de gehanteerde methodologie voor de school in kwestie het meest geschikt is om de vooropgestelde schoolaspecten te beschrijven centraal. Bijkomende criteria om de bruikbaarheid van een instrument te beoordelen zijn verder: Is de gekozen werkwijze gebruiksvriendelijk? Wordt het instrument als niet-bedreigend ervaren? Wordt de keuze van het instrument ondersteund door de betrokkenen? Sluit het instrument aan bij de externe ondersteuning?
- Aandachtspunten bij het uitwerken van een eigen werkwijze. De opstap naar zelfevaluatie hoeft niet noodzakelijk te worden gemaakt aan de hand van wetenschappelijk gevalideerde instrumenten. Zelfbevragingen kunnen op een eenvoudige manier ook door de school zelf ontwikkeld en georganiseerd worden (Gombeir, 1998). Dit heeft als voordeel dat de output minder gebonden is aan de inhoud van een instrument en dat voor de eigen school relevante aspecten van het functioneren naar voren kunnen komen. Het resultaat van dergelijk schooleigen ontwikkelingswerk moet evenwel net als elders ontleende zelfevaluatie-instrumenten aan de vooropgestelde eisen voldoen. In het bijzonder de vraag naar de begripsvaliditeit stelt zich hier.

Merk op dat een school niet noodzakelijk beroep moet doen op een instrument. Instrumenten kunnen immers niet op alle mogelijke vragen antwoorden geven. Een kwaliteitsvolle zelfevaluatie kan dus eveneens gekenmerkt worden door een beargumenteerde afwezigheid van een bestaand instrument. De school kan dan een eigen strategie voor

informatieverzameling uitwerken en een eigen beoordelingssysteem hanteren. Deze vormen van zelfevaluatie behoren evenwel niet tot de focus van deze studie.

- *Bewaken van de efficiëntie en haalbaarheid*

Het realiseren van uiterst valide en betrouwbare zelfevaluaties vergt een intensieve tijds- en kosteninvestering. Een succesvolle zelfevaluatie gaat echter ook uit van een tijdsplanning die in verhouding is tot de doelen van de zelfevaluatie (van Aanholt & Buis, 1990). Gezien de beschikbaarheid van tijd en middelen niet onuitputtelijk zijn, stellen we daarom ook de *efficiëntie* van de gegevensverzameling als aandachtspunt voorop. De keuze voor een bepaalde werkwijze of een concreet instrument wordt dus best óók ingegeven door overwegingen inzake efficiëntie en haalbaarheid. Macbeath (1999) benadrukt bijvoorbeeld dat bewaakt moet worden dat het uitvoeren van een zelfevaluatie past binnen andere activiteiten die in de school lopen. Dit kan betekenen dat een zelfevaluatie binnen aanvaardbare grenzen aan validiteit en betrouwbaarheid inboet. Maar uiteraard ook andersom: wanneer de zelfevaluatie validiteit, en betrouwbaarheid vereist, kan men niet voorbij aan een zekere tijdsinvestering. De investeringen moeten met andere woorden in verhouding zijn met de doelen die men met de zelfevaluatie wil bereiken (Oost & Markenhof, 2002a) en met de resultaten die men verwacht (Devos e.a., 2000).

c. HET BEANTWOORDEN VAN DE VRAAG - INTERPRETATIE

Los van de vraag naar de geldigheid van de verzamelde zelfevaluatiegegevens staat de vraag naar de geldigheid van de antwoorden/conclusies die getrokken worden. Op basis van de uitgevoerde gegevensverzameling reflecteert men immers over de zelfevaluatiethema's en vindt er een terugkoppeling plaats naar de oorspronkelijke (probleem)situatie. Deze analyses laten toe tot een handelingsplan te komen. Ook in deze fase is er een aantal sleutelelementen waarvan verwacht wordt dat ze bijdragen tot een kwaliteitsvolle zelfevaluatie. Wij onderscheiden er vier (Devos e.a., 2000; Van Petegem, 2005):

- In de eerste plaats is het van belang dat men de onderzochte thema's eerst beschrijft alvorens ze vervolgens ook te beoordelen. Dat houdt in dat men de resultaten van de zelfevaluatie in eerste instantie enkel omschrijft zonder er een waardeoordeel aan te koppelen. Om tot correcte interpretaties te komen, is het belangrijk dat beide fasen zoveel mogelijk gescheiden worden.
- Daarnaast is er de vraag of de interpretaties aansluiten bij de doelstellingen van de zelfevaluatie. Wanneer de interpretatie van de zelfevaluatieresultaten niet aansluit bij de doelstellingen die men bij het begin van de zelfevaluatie heeft geformuleerd dan bestaat het gevaar dat de verkeerde conclusies getrokken worden. Het beantwoorden van de vraag gebeurt dus logischerwijs best zo doelgericht mogelijk. Naast deze inhoudelijke doelgerichtheid, dient ook het voorkomen van valkuilen in het interpretatieproces benadrukt te worden. Zo moet men zich er goed van bewust zijn dat de zelfevaluatiegegevens enkel een beeld geven van de onderzochte onderwer-

pen en van de onderzochte respondenten. Correcte interpretaties houden met andere woorden rekening met het gevaar van foutieve veralgemeningen.

- In het interpretatieproces spelen ook de vergelijkingspunten die gehanteerd worden een cruciale rol. Van Petegem & Vanhoof (2004) pleiten voor een gecombineerd gebruik van een normgericht, criteriumgericht en schoolgericht referentiekader. Bij het beantwoorden van de vragen die aan de zelfevaluatie voorafgaan is het van belang dat de keuze van te gebruiken vergelijkingsbasis goed beargumenteerd wordt.
- De interpretatie van de zelfevaluatiegegevens is er doorgaans op gericht om ook verklaringen te vinden voor de waargenomen fenomenen. In het beantwoorden van de zelfevaluatievragen wordt dus best voldoende aandacht besteed aan de oorzaken of redenen die een rol kunnen spelen in het verklaren van de vaststellingen. Op die manier komt men tot interpretaties die een goede basis kunnen vormen voor het identificeren van actiepunten.

d. DE OPERATIONALISATIE VAN 'ZELFEVALUATIE ALS ONDERZOEKSDAAD'

De onderstaande lijst stelt voor de verschillende aspecten van 'zelfevaluatie als onderzoeksdaad' een aantal concrete items voorop. Aan de respondenten die voldoende nauw betrokken waren bij de verschillende aspecten van de zelfevaluatie, wordt de vraag gesteld in welke mate de stellingen van toepassing zijn op de zelfevaluatie die in de eigen school heeft plaatsgevonden.

Inhoudelijke verankering
De keuze van de te evalueren thema's werd goed beargumenteerd.
De inhoud van de zelfevaluatie werd gesitueerd in een ruimer kader.
Relevantie
De geëvalueerde thema's waren relevant.
Precisie
De te evalueren thema's werden zeer nauwkeurig omschreven.
Ik had een duidelijk beeld van de bedoelde inhoud van de zelfevaluatie.
Betrouwbaarheid
De gegevensverzameling werd zorgvuldig uitgevoerd.
Het tijdstip waarop de vragenlijst ingevuld werd, heeft geen invloed gehad op de resultaten.
Ik heb de vragenlijst eerlijk ingevuld.
De anonimiteit bij het invullen van de vragenlijst was voldoende gewaarborgd.
Validiteit
De verzamelde gegevens geven een goed beeld van de bevraagde thema's.
De gegevensverzameling leverde representatieve informatie op.
De vragenlijst verzamelde precies informatie over wat wij wilden weten.
Efficiëntie
Bij de start van de zelfevaluatie werd een realistische tijdsplanning vooropgesteld.
De kosten-batenanalyse van deze zelfevaluatie is positief.
Keuze instrument
De vragenlijst sloot goed aan bij de doelen van de zelfevaluatie.
De bruikbaarheid van de vragenlijst werd afgewogen tegen de bruikbaarheid van andere vragenlijsten.
De vragenlijst was gebruiksvriendelijk.

Interpretatie
De gegevens werden correct geïnterpreteerd.
Er was een passend kader om te bepalen wat goed en wat slecht was.

3 OMTRENT DE ONDERSTEUNING

Bij het uitvoeren van zelfevaluaties kunnen scholen er voor kiezen om een beroep te doen op externe ondersteuning. Daaronder verstaan we deskundigen die geen lid zijn (van het cliëntsysteem) van de schoolorganisatie als dusdanig (Verhoeven, 1986). Het gaat dus zowel om nascholings- als om begeleidingsinitiatieven. Vormen van training en ondersteuning zijn noodzakelijke voorwaarden om tot kwaliteitsvolle zelfevaluaties te komen (Nevo, 1995). We verwachten dan ook dat de mate waarin scholen een beroep doen op externe ondersteuning van belang is in het voorspellen van de kwaliteit van zelfevaluaties door scholen. We willen over de ondersteuning waarop scholen tijdens het zelfevaluatieproces een beroep doen zowel descriptieve als beoordelende uitspraken doen. Dat vergt twee afzonderlijke invalshoeken.

3.1 Het beschrijven van de ondersteuning

De ondersteuning waarop scholen een beroep doen, kan verscheidene vormen aannemen. Precies die variatie willen we in eerste instantie in kaart brengen. Het is echter niet onze bedoeling om uitspraken te doen over individuele ondersteuningsinitiatieven. Onze interesse gaat in eerste instantie uit naar de types van ondersteuning waarop scholen met het oog op het welslagen van het zelfevaluatieproces een beroep doen. Daarbij wordt niet verwezen naar concrete initiatieven. In de vragenlijst wordt in kaart gebracht of de leden van het schoolteam in het verleden al dan niet aan de volgende vormen van ondersteuning deelnamen. We presenteren de activiteiten die we onderscheiden meteen zoals ze in de vragenlijst geoperationaliseerd zullen worden (Vandenberghe, Van de Velde & Maes, 2000).

In de loop van één van de laatste vijf schooljaren:
...heb ik documentatiemateriaal gelezen over zelfevaluaties door scholen.
...woonde ik op onze school een studiedag bij die handelde over zelfevaluatie.
...woonde ik een externe studiedag bij (een halve of hele dag) over zelfevaluatie.
...heb ik meerdaagse cursussen/trainingen/workshops over zelfevaluatie gevolgd.
...ben ik betrokken geweest bij een begeleidingtraject omtrent zelfevaluatie verzorgd door de pedagogische begeleiding.
...ben ik betrokken geweest bij een begeleidingtraject omtrent zelfevaluatie verzorgd door een nascholingsinstituut.
...heb ik een studiebezoek gedaan bij andere scholen met zelfevaluatie-ervaring.

3.2 Het beoordelen van de ondersteuning

Voor twee aspecten van de ondersteuning die scholen genoten, werd een indicator uitgewerkt. Het gaat om de mate waarin de ondersteuning die scholen genoten inspeelde op

de reële behoeftes van scholen enerzijds en de mate waarin de ondersteuning een aantal doelstellingen bereikte anderzijds. De keuze voor de eerste indicator is ingegeven door de inventaris van criteria die door Peeters et al. (1996) ontwikkeld werd. Het gaat om één van de cruciale elementen waarover in de literatuur consensus bestaat dat ze de effectiviteit van ondersteuningspraktijk ten goede komen. We lichten beide criteria toe.

3.2.1 Omtrent behoeftes: Speelt de ondersteuning in op reële behoeftes van scholen?

Peeters et al. (1996) baseren dit criterium op de overweging dat ondersteuning en de middelen die eraan worden besteed slechts zin hebben wanneer ze aan een behoefte (proberen) tegemoet (te) komen. We hanteren daarbij de omschrijving van een behoefte als een tekort dat door de school die de zelfevaluatie uitvoert zelf aangevoeld en erkend wordt (Schollaert, Peeters & van Cauwenberghe, 1994). Nevo (1995, p. 173) schrijft dat ondersteuning bij zelfevaluaties 'must be adapted to the special needs of school educators'.

3.2.2 Omtrent de resultaten: In welke mate worden de vooropgestelde doelen bereikt?

De doelen van nascholing kunnen divers zijn: sensibilisering, kennisvergroting, komen tot waarneembare veranderingen. Peters et al. (1996, p. 59) schrijven dat elk doel op zich legitiem kan zijn maar dat er in het beoordelen van ondersteuningsinitiatieven toch niveauverschillen moeten gemaakt worden. Zij hanteren van laag naar hoog de volgende hiërarchie van cumulatieve ondersteuningsdoelstellingen. Daarmee wordt bedoeld dat veranderingen op een hoger niveau niet mogelijk worden geacht zonder veranderingen op de lagere niveaus.

1. *Er wordt sensibilisering nagestreefd. Men wil de deelnemers overtuigen van het belang van het uitvoeren van zelfevaluatie.*
2. *Er wordt verandering op kennisniveau nagestreefd. Men wil de kennis van de deelnemers over (het verloop van) zelfevaluaties uitbreiden.*
3. *Er wordt verandering op inzichtniveau nagestreefd. Men wil de mate waarin de deelnemers verbanden leggen, samenhangen begrijpen en aspecten van zelfevaluatie in hun eigen termen kunnen verwoorden, vergroten.*
4. *Er wordt verandering op toepassingsniveau nagestreefd. Men wil de deelnemers in staat stellen om het geleerde uit te voeren in de praktijk.*
5. *Er wordt verandering op integratieniveau nagestreefd. Men wil de deelnemers ertoe brengen om het geleerde spontaan in hun eigen praktijksituatie te gaan toepassen.*

Los van de vraag welke doelen door de ondersteuners nagestreefd werden, staat de beoordeling door de scholen zelf over de mate waarin men daarin effectief is geslaagd. Deze invalshoek staat centraal in de beoordeling van de resultaten van de ondersteuning in deze studie. Het kader voor de beoordeling van de ondersteuning, wordt aan de hand van de volgende items geoperationaliseerd.

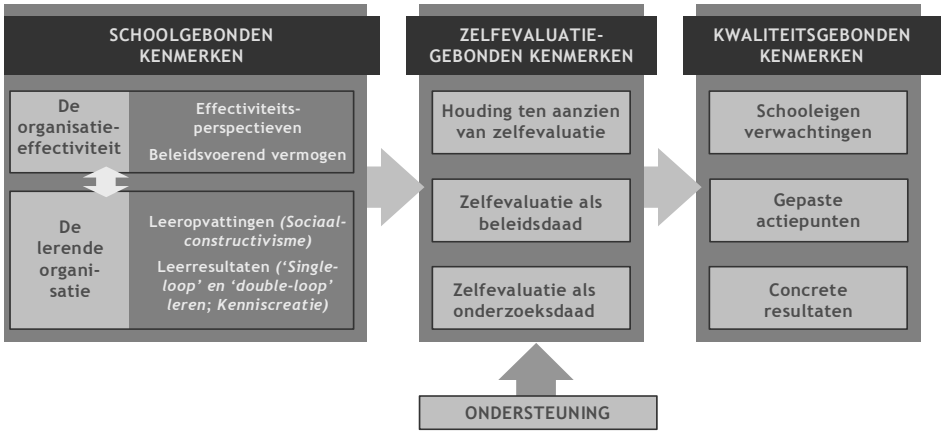
De studiedagen en/of begeleiding waaraan ik deelnam:
...speelde voldoende in op mijn behoeftes.
...hield voldoende rekening met de eigenheid van onze school.
...overtuigde mij van het belang van het uitvoeren van zelfevaluatie.
...slaagde erin mijn kennis over (het verloop van) zelfevaluaties uit te breiden.
...slaagde erin de mate waarin ik verbanden leg en samenhangen begrijp tussen aspecten van zelfevaluatie te vergroten.
...stelde mij in staat om een kwaliteitsvolle zelfevaluatie uit te voeren.
...slaagde erin mij ertoe te brengen om zelfevaluaties spontaan toe te passen.

4 AFRONDEND: TE TOETSEN ONDERZOEKSHYPOTHESEN

De probleemstelling, de voorbereidende studies en het theoretisch kader leveren vanuit verschillende invalshoeken structurerende elementen aan om naar het verloop en de kwaliteit van zelfevaluaties te kijken. Ze stellen ons in staat om beargumenteerde verwachtingen en hypotheses uit te spreken. Doordat ze 'toetsbaar' geformuleerd worden, zijn ze richtinggevend voor de toekomstige analyses. Met name de volgende hypotheses worden in het verdere verloop van dit proefschrift getoetst aan empirische evidentie. Alvorens de hypothesen te presenteren, brengen we de onderzoeksvragen en het theoretisch kader in herinnering.

1. Aan de hand van welke indicatoren kan men de kwaliteit van een zelfevaluatie operationaliseren?
2. Hoe kwaliteitsvol verlopen zelfevaluaties in Vlaamse scholen en wat zijn de verschillen tussen scholen terzake?
3. In welke mate wordt het verloop van een zelfevaluatie bepaald door (a) kenmerken van het functioneren van de school en (b) de ondersteuning die scholen genieten?
4. In welke mate worden de resultaten van een zelfevaluatie bepaald door (a) het verloop van de zelfevaluatie en (b) de ondersteuning die scholen genieten?

Figuur 10: Theoretisch kader voor meta-evaluatie



- Hypothese 1: Over de variatie in het verloop van zelfevaluatieprocessen

1. Onze eerste hypothese is dat het verloop van geïnstrumentaliseerde zelfevaluaties geen vastliggende standaard kent. Het proces wordt telkens op een specifieke manier ingevuld. Zo verwachten wij bijvoorbeeld verschillen inzake vooropgestelde functies en verwachtingen, inzake gehanteerde instrumenten en ondersteuning en inzake betrokkenheid van diverse geledingen in de verschillende fasen van het zelfevaluatieproces.

- Hypothese 2: Over de houding ten aanzien van zelfevaluatie

2a. Wij verwachten dat de houding van alle respondenten ten aanzien van zelfevaluatie gemiddeld slechts in beperkte mate positief is. Dit gemiddelde beeld verhult echter een grote variatie. Wij verwachten dat de nulhypothese die stelt dat er geen verschillen zijn tussen respondenten en scholen ten aanzien van deze houding verworpen dient te worden. Wij verwachten bijvoorbeeld dat schoolleiders zich positiever uitspreken over zelfevaluatie dan leerkrachten.

2b. Verder verwachten wij dat de houding van de respondenten ten aanzien van zelfevaluatie positiever is naarmate scholen volgens diezelfde respondenten meer organisatie-effectief zijn én meer tegemoet komen aan het beeld van de lerende organisatie.

- Hypothese 3: Over 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad'

3a. Onze hypothese is dat zelfevaluaties in de bestudeerde scholen gemiddeld slechts gematigd positief scoren op de verschillende dragers van 'zelfevaluatie als beleidsdaad' en van 'zelfevaluatie als onderzoeksdaad'. We verwachten ook hier echter noemenswaardige verschillen tussen scholen. Zelfevaluaties komen dus in verschillende mate tegemoet aan (de dragers van) beide concepten. We verwachten ook dat er noemenswaardige verschillen zijn tussen respondenten binnen dezelfde school.

3b. De invulling van de 'zelfevaluatie als beleidsdaad' en van de 'zelfevaluatie als onderzoeksdaad' staat niet los van de kenmerken van het algemene functioneren van de school. Scholen zullen hoger scoren op beide concepten en hun dragers naarmate ze meer organisatie-effectief zijn én meer tegemoet komen aan het beeld van de lerende organisatie.

- Hypothese 4: Over de samenhang van de zelfevaluatiegebonden kenmerken

4. Wij verwachten dat er een sterke positieve samenhang is tussen de verschillende zelfevaluatiegebonden kenmerken. Respondenten en scholen die hoog scoren op 'houding ten aanzien van zelfevaluatie', zullen volgens deze hypothese ook hoog scoren op 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad'.

- Hypothese 5: Over de 'kwaliteit van zelfevaluaties'

5a. We verwachten dat de bestudeerde zelfevaluaties volgens de respondenten gemiddeld genomen als eerder wel kwaliteitsvol worden beoordeeld. Wij verwachten echter zowel tussen als binnen scholen verschillen in de perceptie van de drie indicatoren van kwaliteitsvolle zelfevaluatie.

5b. We verwachten dat zelfevaluaties die hoog scoren op één kwaliteitsindicator ook hoog zullen scoren op de andere kwaliteitsindicatoren.

5c. De derde deelhypothese heeft betrekking op de rol van de beoordelaar op de uitgesproken kwaliteitsbeoordeling. Wij verwachten dat mensen die deel uitmaken van de schoolleiding of nauw bij het zelfevaluatiegebeuren betrokken zijn een meer positief oordeel uitspreken over de kwaliteit van de zelfevaluatie dan andere teamleden. Daarnaast is onze onderzoekshypothese dat de beoordeling door internen (bvb. schoolleiders) (in zekere mate) positief samenhangt met de beoordeling door externen (bvb. de onderwijsinspectie).

- Hypothese 6: Over het verklaren van de 'kwaliteit van zelfevaluaties'

6. Naarmate scholen hoger scoren op de zelfevaluatiegebonden kenmerken zullen zij ook hoger scoren op de verschillende kwaliteitsindicatoren. Hetzelfde geldt volgens onze verwachtingen voor scholen die hoger scoren op de schoolgebonden kenmerken. We verwachten echter dat het effect van de schoolgebonden kenmerken op de kwaliteit van zelfevaluaties loopt via de mediërende zelfevaluatiegebonden kenmerken.

- Hypothese 7: Over de impact van ondersteuning

7a. We verwachten dat de score op de zelfevaluatiegebonden kenmerken hoger zal zijn wanneer respondenten een vorm van ondersteuning genoten. We verwachten daarbij een grotere impact van ondersteuning die door de respondenten als kwaliteitsvol wordt omschreven.

7b. We verwachten dat ondersteuning ook een impact heeft op de kwaliteit van zelfevaluatie. Het effect van ondersteuning loopt echter over de zelfevaluatiegebonden kenmerken. Eenmaal daarvoor gecontroleerd wordt, verwachten we geen rechtstreeks effect meer van (de kwaliteit van) ondersteuning op de kwaliteitsindicatoren van zelfevaluatie.

- Hypothese 8: Over meta-evaluatie

8. Op basis van de in het theoretisch kader opgenomen schoolgebonden en zelfevaluatiegebonden kenmerken kunnen voorspellingen worden gedaan over de kwaliteit van zelfevaluaties. We verwachten dat na een weloverwogen selectie van relevante informatie nog steeds een behoorlijk deel van de variantie in de kwaliteit van zelfevaluatie kan verklaard worden.

Deze hypotheses worden in het vervolg van deze studie getoetst aan empirische evidentie. Dat doen we eerst aan de hand van kwantitatieve analyses, daarna aan de hand van gevalstudies.

DEEL 5

VOORBEREIDING VAN DE KWANTITATIEVE ANALYSES

Dit deel gaat aan de bespreking van de resultaten van de kwantitatieve analyses vooraf. We lichten toe waarom er sprake is van kwantitatieve analyses en welke gegevens we daarbij aanwenden (zie par. 1). In de analyses zal immers sprake zijn van een geïntegreerd gebruik van drie gegevensverzamelingen. Paragraaf 2 gaat dieper in op de achtergrond van deze databronnen. Hier staan we stil bij aspecten van steekproeftrekking, inhouden, gerealiseerde respons en dergelijke. Paragraaf 3 tenslotte bevat de voorbereidende analyses. We focussen hier op het ontwikkelen van de schalen, de indicatoren en de clusters van scholen en respondenten die in het verdere onderzoek gebruikt zullen worden.

1 (WAAROM) VERSCHILLENDE VORMEN VAN KWANTITATIEVE ANALYSES

De centrale motivering voor het opnemen van een kwantitatief luik in dit onderzoek bestaat uit het streven naar veralgemeenbaarheid en representativiteit. Wat we beogen, is het genereren van een aantal verantwoorde uitspraken over hoe scholen vorm geven aan zelfevaluaties én over de mate waarin de aanpak van scholen aan een aantal procesindicatoren voldoet. We willen het proces en de kwaliteit van zelfevaluaties zowel beschrijven als verklaren. Om daartoe informatie te verzamelen moeten de juiste personen aan een systematische bevraging onderworpen worden. De data die wij in de kwantitatieve analyses zullen aanwenden zijn afkomstig van drie bronnen: twee bronnen zijn het resultaat van eigen vragenlijstonderzoek, een derde bron bestaat uit gegevens die door de Vlaamse overheid (c.q. de onderwijsinspectie) verzameld werden. In het tweede geval spreken we van secundaire data-analyse.

Het gebruik van **vragenlijsten** is een onderzoeksbenadering waarbij informatie wordt verzameld over een aantal kenmerken van een groot aantal onderzoekseenheden. Die dataverzameling gebeurt op een gestandaardiseerde manier. Het gebruik van vragenlijsten is de aangewezen weg om relatief eenvoudig en snel feiten, opinies en gedrag te kunnen vaststellen.

De keuze voor **secundaire data-analyse** is ingegeven door mogelijke tekorten van het werken met surveys. Vragenlijsten richten zich noodgedwongen op de opinies van respondenten. Ze hebben als het ware het nadeel dat de bestudeerde fenomenen telkens in kaart gebracht worden vanuit het interpretatiekader van de respondent. Dit is in de context van deze studie niet onbelangrijk. Uitspraken over de kwaliteit van zelfevaluaties worden best niet alleen gebaseerd op de perceptie van schoolleiders en leerkrachten. Een wezenlijk aspect van onderwijskwaliteit is weliswaar dat ze gepaard gaat met beoordeling, zowel objectief als subjectief. In een systeem voor kwaliteitszorg moet men echter een onderscheid maken tussen een verwachte kwaliteit, een beleefde kwaliteit en een feitelijke kwaliteit (Cuyvers, 2002, p. 26). Om een volledig beeld te krijgen van de kwaliteit van zelfevaluaties is er nood aan onderzoek dat de beleefde kwaliteit koppelt aan de verwachte en feitelijke kwaliteit. Door gebruik te maken van data van de Vlaamse overheid kunnen wij hier aan tegemoet komen. De gegevens bestaan immers uit het oordeel van een externe instantie over het verloop en de kwaliteit van zelfevaluaties in Vlaamse scholen.

2 TOELICHTING BIJ DE BESCHIKBARE DATA

In deze paragraaf geven we toelichting bij de drie gegevensverzamelingen die aan de basis liggen van de kwantitatieve analyses in deze studie. We verwijzen er respectievelijk naar als de 'GOK-zelfevaluaties', de 'GOK-inspectie' en 'Zelfevaluatie binnenstebuiten'. Voor elke vorm van gegevensverzameling staan we stil bij de gebruikte instrumenten, de respondenten, de gegevensverzameling, de representativiteit van de data en de aansluiting bij het theoretisch kader. De eerste twee beschrijvingen zijn een samenvatting van wat in andere rapporten beschreven werd (Van Petegem e.a., 2005b; MVG, 2006). De derde gegevensverzameling is nieuw en daarom uitgebreider. We ronden deze paragraaf af met een vergelijkend overzicht van de drie gegevensverzamelingen.

2.1 De GOK-zelfevaluaties

GOK staat voor het gelijke onderwijskansenbeleid (GOK-beleid) dat in september 2002 in Vlaanderen van start is gegaan (X, 2002). Ruim 2000 Vlaamse scholen - basis en secundair - krijgen naar aanleiding van dit beleid bijkomende lestijden toebedeeld om gelijke onderwijskansen te bevorderen. Of scholen al dan niet over aanvullende lestijden beschikken, hangt af van hun aandeel doelgroepleerlingen. Met de beschikbare lestijden kunnen scholen leerkrachten vrijmaken om aan gelijke onderwijskansen te werken. Deze leerkrachten worden GOK-leerkrachten genoemd.

Terwijl de beleidsruimte van scholen m.b.t. de gelijke kansenproblematiek vroeger eerder beperkt te noemen was, genieten scholen sinds het nieuwe decreet een grotere vrijheid. Eén van de weinige sturende elementen vanwege de overheid is de verplichting om zichzelf te evalueren. Tijdens twee verschillende fasen van het nieuwe GOK-beleid werd van scholen verwacht dat ze aspecten van het eigen schoolfunctioneren kritisch onder de

loep namen, beschreven en beoordeelden. Dat gebeurde in de vorm van een analyse van de beginsituatie bij de invoering van het nieuwe decreet en in de vorm van een zelfevaluatie na twee jaar van gelijke kansenbeleid. De dataverzameling die wij zullen gebruiken bracht in kaart hoe kwaliteitsvol GOK-scholen aan deze zelfevaluatie vorm hebben gegeven[5]. Daarnaast werd bestudeerd welke kenmerken van het functioneren van scholen de kwaliteit van deze zelfevaluaties beïnvloeden.

De populatie van de GOK-studie bestaat uit alle scholen die naar aanleiding van het gelijke onderwijskansendecreet een aantal aanvullende lestijden ter beschikking krijgen. Er worden in die groep drie deelpopulaties onderscheiden (basisonderwijs, 1ste graad secundair onderwijs en 2de en 3de graad secundair onderwijs). Voor elk van deze drie doelpopulaties werd een representatieve steekproef getrokken (voor details zie Van Petegem e.a. (2005b)). Dit resulteerde in de volgende steekproefgroottes.

Tabel 7: Effectieve steekproefgrootte van de verschillende doelpopulaties

Doelpopulatie	Aantal scholen in de doelpopulatie	Te realiseren steekproefgrootte	Gerealiseerde steekproefgroottes (na twee rondes)
Basisonderwijs	1801	317	265 (84%)
1ste graad SO	374	190	176 (93%)
2de & 3de graad SO	119	91	77 (85%)
Totaal[6]	2294	598	518 (87%)

De bevraging van scholen verliep in twee rondes. Telkens werd de directeur of iemand van het middenkader die nauw betrokken was bij de GOK-zelfevaluatie aan de hand van een schriftelijke vragenlijst bevraagd. De eerste bevraging vond plaats voor de GOK-zelfevaluatie werd uitgevoerd. De inhoudelijke focus lag onder meer bij de dragers van het beleidsvoerend vermogen in de GOK-scholen. De tweede bevraging vond plaats na de GOK-zelfevaluatie en richtte zich in sterke mate op de 'zelfevaluatie als beleidsdaad' en op de resultaten van de zelfevaluatie.

2.2 De GOK-inspectie

Het GOK-decreet bepaalde dat er een aparte inspectie van alle GOK-scholen moest komen. Ook de vraag of de GOK-zelfevaluaties kwaliteitsvol werden uitgevoerd was een onderdeel van deze inspectie. Kwaliteitsvol werd in de besluiten van de Vlaamse Regering onder meer omschreven als het gepaard gaan met het verzamelen en analyseren van

[5] Dit onderzoek werd uitgevoerd in opdracht van het Ministerie van Onderwijs en Vorming in het kader van OBPWO (Onderwijskundig Beleids- en Praktijkgericht Wetenschappelijk Onderzoek) (nr. 02.07). Op basis van het rapport verscheen een publicatie in boekvorm: zie Van Petegem e.a., 2005b.

[6] Deze optelsom is louter informatief. Het spreekt voor zich dat het verband tussen het totale aantal scholen in de doelpopulatie en tussen de totale steekproefgrootte verloren gaat bij een dergelijke optelling.

gegevens en als het resulteren in voorstellen voor verbetering van de eigen werking (X, 2002). De werkwijze van de GOK-inspectie was deze van een doorlichting door een team van externe beoordelaars. De gelijkgerichtheid van de aanpak van verschillende inspectieteams werd nagestreefd door overleg met alle betrokkenen partijen bij het tot stand komen van de opzet van de controle. Daarnaast werd een eerste versie van het doorlichtingsinstrument uitgeprobeerd in een selectie van scholen. Op basis van deze cases oefenden de inspectieleden zich in het gebruik van het instrument. Tezamen met hun ervaringdeskundigheid biedt deze aanpak volgens de onderwijsinspectie voldoende garanties voor de interbeoordelaarsbetrouwbaarheid van de GOK-inspectie (MVG, 2006).

In de volgende analyses richten we ons op de door de inspectie verzamelde gegevens voor zover deze betrekking hebben op aspecten van het verloop en de resultaten van de GOK-zelfevaluatie. Dat betekent dat we zowel voor basisscholen als secundaire scholen beschikken over vier indicatoren. Scholen kregen voor elke indicator een beoordeling op een schaal die bestaat uit vier niveaus: onvoldoende, aanvaardbaar, voldoende tot goed, zeer goed. Deze gegevens zijn beschikbaar voor de volledige populatie van GOK-scholen zoals beschreven in Tabel 7.

We staan tenslotte ook even stil bij de inhoud van de GOK-inspectie. Eén van de indicatoren heeft betrekking op de mate waarin scholen op basis van de zelfevaluatie komen tot voorstellen voor verbetering en bijsturing. Deze schaal is met andere woorden een operationalisatie van wat we in het theoretisch kader als 'de kwaliteit van de zelfevaluatie' bestempelden. De drie andere indicatoren hebben betrekking op het verloop van de zelfevaluatie. Omdat de gebruikte operationalisaties in het theoretisch kader niet vermeld werden, zetten we de betekenis van de verschillende scores hieronder even op een rij.

Indicator 1: Voorstellen tot verbetering en bijsturing. De school...	
0nvoldoende:	... heeft geen voorstellen tot verbetering/bijsturing geformuleerd.
Aanvaardbaar:	... heeft weinig voorstellen tot verbetering/bijsturing geformuleerd.
Voldoende tot goed:	... heeft voldoende voorstellen tot verbetering/bijsturing geformuleerd.
Zeer goed:	... heeft grondig werk gemaakt van het formuleren van voorstellen tot verbetering/bijsturing van haar goede werking.
Indicator 2: Gegevensverzameling. De school...	
Onvoldoende:	... heeft geen gegevens verzameld met betrekking tot de gekozen doelstellingen.
Aanvaardbaar:	... heeft weinig gegevens verzameld met betrekking tot de gekozen doelstellingen.
Voldoende tot goed:	... heeft voldoende gegevens verzameld m.b.t. de gekozen doelstellingen.
Zeer goed:	... heeft systematisch en op volledige wijze gegevens verzameld m.b.t. de gekozen doelstellingen.
Indicator 3: Gegevensanalyse. De school...	
Onvoldoende:	... heeft de gegevens niet geanalyseerd.
Aanvaardbaar:	... heeft de verzamelde gegevens in enige mate geanalyseerd.
Voldoende tot goed:	... heeft de verzamelde gegevens voldoende geanalyseerd.
Zeer goed:	... heeft de verzamelde gegevens grondig geanalyseerd.

Indicator 4: Betrokkenheid.	
Onvoldoende:	Het schoolteam was helemaal niet betrokken bij het zelfevaluatieproces.
Aanvaardbaar:	Een gedeelte van het schoolteam voelde zich betrokken bij het zelfevaluatieproces.
Voldoende tot goed:	Het schoolteam was voldoende betrokken bij het zelfevaluatieproces.
Zeer goed:	Het schoolteam heeft het zelfevaluatieproces grondig gedragen.

Elke GOK-school werd door het inspectieteam op elk van deze vier indicatoren beoordeeld. Op die manier beschikken we dus ook over de beoordeling van een externe instantie over het verloop en de kwaliteit van de GOK-zelfevaluaties in scholen.

2.3 Zelfevaluatie binnenstebuiten

De gegevensverzameling die we benoemen als 'zelfevaluatie binnenstebuiten' is specifiek opgezet met het oog op beantwoorden van de voorliggende onderzoeksvragen en sluit naadloos aan bij het theoretisch kader dat in deel 3 ontwikkeld werd. Het verloop van deze gegevensverzameling beschrijven we uitgebreider dan de andere. In de volgende paragrafen staan we achtereenvolgens stil bij de doelpopulatie en de steekproef (2.3.1), bij het ontwikkelen van de vragenlijst (2.3.2) en bij de kenmerken van de steekproef (2.3.3).

2.3.1 (Het tot stand komen van) de steekproef

In de volgende paragrafen bespreken we de steekproef van 'zelfevaluatie binnenstebuiten'. We eindigen met het omschrijven van de steekproef. We geven echter eerst aan hoe deze tot stand kwam. Daartoe wordt ook stilgestaan bij de opzet van de bevraging.

a. Van een doelpopulatie naar een steekproef

De doelpopulatie voor deze studie bestaat uit alle Vlaamse scholen (basis en secundair) die in één van de voorbije schooljaren een geïnstrumentaliseerde zelfevaluatie uitvoerden. Geïnstrumentaliseerd omschreven we hoger als gebruikmakend van een schriftelijke vragenlijst die door (eventueel onder meer) leerkrachten ingevuld werd. Klassieke voorbeelden van zulke vragenlijsten zijn onder meer DISO (Diagnose Instrument Secundair Onderwijs), PROZA (Projectgroep Ontwikkeling Zelfanalyse-instrument), SION (School in Ontwikkeling) en IZES (Instrument voor Zelfevaluatie). Ook scholen die met andere reeds bestaande instrumenten werkten behoren tot de doelgroep. Dat geldt bovendien ook voor scholen die zelf een vragenlijst ontwikkelden.

Het spreekt voor zich dat deze populatie niet duidelijk afgelijnd omschreven kan worden. Meer zelfs, het is (praktisch) onmogelijk om te achterhalen welke scholen precies tot deze doelpopulatie behoren. Er bestaat nergens een databank van Vlaamse scholen die aan de hand van een vragenlijst voor leerkrachten (cf. geïnstrumentaliseerd) een zelfevaluatie uitvoerden. Omdat de doelpopulatie dus ongekend is, was het logischerwijze niet mogelijk om op een systematische manier te komen tot een representatieve steekproef.

Ondanks het feit dat er geen sprake kon zijn van steekproeftrekking in de klassieke betekenis, werden toch een aantal principes in acht genomen bij het zoeken van potentiële steekproefscholen. Een eerste principe was zoveel als mogelijk scholen te bevragen met het oog op het reduceren van sampling error. Hoe groter de steekproef, des te representatiever de steekproef (Creswell, 2005). Een tweede principe was het nastreven van verwachte variatie in de doelpopulatie. Daarmee bedoelen we bijvoorbeeld variatie in onderwijsniveau, variatie in onderwijsnet, variatie in gebruikte instrumenten, variatie in vroegere ervaringen met zelfevaluatie. Ons doel was niet de (overigens onbekende) variatie in de doelpopulatie verhoudingsgewijs te representeren. Het doel was te bewaken dat de cases in de steekproef getuigden van de verschillende mogelijkheden in de populatie.

Om dit doel te bereiken werden verschillende strategieën aangewend om tot mogelijke steekproefscholen te komen. Elk van deze strategieën informeerde scholen grondig over het opzet van dit onderzoek. Om de verschillende strategieën – en de reacties van scholen – te kaderen, staan we eerst stil bij de opzet van de bevraging.

b. Opzet van de bevraging

De opzet van de bevraging is het verzamelen van opvattingen van de leden van schoolteams over het verloop en de kwaliteit van de zelfevaluatie die heeft plaatsgevonden enerzijds en over een aantal schoolgebonden kenmerken anderzijds. Dat gebeurde aan de hand van een schriftelijke vragenlijst bij verschillende geledingen van het team. Er werd een online vragenlijst afgenomen bij de schoolleider, (indien aanwezig) het middenkader en zoveel mogelijk leden van het leerkrachtenteam.

- Wat schoolleiding en middenkader betreft: wij verwachtten in elk geval de deelname van de (adjunct)directeur en het middenkader in de school (denk aan TA(C)'s, beleidsondersteuners, coördinatoren).
- Wat leerkrachten betreft: het motto was hier zoveel mogelijk met een minimum van 20, tenzij in kleinere scholen uiteraard. Daar streefden we naar de deelname van alle leerkrachten.
- Wat administratief, pedagogisch en paramedisch ondersteunend personeel betreft: scholen werden gestimuleerd om ook hen uit te nodigen om deel te nemen indien zij betrokken partij waren bij de zelfevaluatie. Indien zij te ver afstonden van het zelfevaluatieproces was hun deelname niet noodzakelijk en bepaalde de school zelf of ze hun deelname voor de feedback wenselijk achtte.

De bevraging gebeurde anoniem d.m.v. een online vragenlijst. Respondenten konden van op elke PC die toegang geeft tot het internet (op school of thuis) deelnemen. Zij surften daarvoor naar een beveiligde website. Om in te loggen gebruiken zij het persoonlijke paswoord dat op hun geïndividualiseerde uitnodiging om deel te nemen vermeld stond. De invultijd bedroeg naar schatting 30 minuten.

c. Strategieën om tot het vooropgestelde aantal deelnemende scholen te komen

Om een zo groot mogelijk bronbestand aan te leggen van scholen die tot de doelpopulatie behoorden, volgden we verschillende sporen. Het doel van de volgende sporen was te achterhalen welke scholen zich recent engageerden in een geïnstrumentaliseerde vorm van zelfevaluatie.

- Raadpleging van nascholingsinstituten. Aan drie niet netgebonden instellingen die actief zijn op het terrein van het begeleiden van zelfevaluaties werd gevraagd naar een lijst van scholen die tijdens 2004-2005 en/of 2005-2006 een geïnstrumentaliseerde zelfevaluatie uitvoerden. Deze instellingen waren Prose, EduBROn en ECEGO.
- Raadpleging van de pedagogische begeleidingsdiensten. Ook aan de netgebonden begeleidingsdiensten werd gevraagd welke scholen in de betreffende schooljaren een geïnstrumentaliseerde zelfevaluatie uitvoerden. Het kon daarbij zowel gaan om instrumenten die door de begeleidingdiensten zelf ontwikkeld werden (zoals DISO, SION) als om elders ontwikkelde of door de school zelf ontwikkelde instrumenten. Er werd een oproep gelanceerd via de afdelingshoofden van begeleidingsdiensten van het Gemeenschapsonderwijs (GO), het Vlaams Secretariaat van het Katholiek Onderwijs (VSKO), het Onderwijssecretariaat van de Steden en Gemeenten van de Vlaamse Gemeenschap (OVSG) en het Provinciaal Onderwijs Vlaanderen (POV). Daarnaast werd ook bij het Overleg Kleine onderwijsverstrekkers (OKO) geïnformeerd. Dit overlegplatform verzorgt de gezamenlijke vertegenwoordiging van onder meer de Federatie van Onafhankelijke Pluralistische Emancipatorische Methodescholen (FOPEM) en de Federatie van Rudolf Steinerscholen in Vlaanderen.

Beide sporen resulteerden in een lijst van scholen waarvan geweten was dat ze tot de doelpopulatie van 'zelfevaluatie binnenstebuiten' behoorden. Deze scholen werden in mei en juni 2006 aangeschreven met de uitnodiging om aan het onderzoek deel te nemen. Daarnaast werd zowel via de nieuwsbrief en de website van de onderzoeksgroep EduBROn als via Schooldirect een open oproep gedaan.

Vermeldenswaardig bij de strategieën om tot het vooropgestelde aantal deelnemende scholen te komen is het feedbackrapport dat we deelnemende scholen ter beschikking stelden. Elke school mocht naar aanleiding van haar medewerking een rapport met daarin de eigen resultaten verwachten. Het rapport bevatte onder meer informatie over de schoolcultuur, kenmerken van de school als lerende organisatie (opvattingen over leren, feitelijke leeractiviteiten, e.d.) en het door het schoolteam gepercipieerde verloop van de bestudeerde zelfevaluatie. Daarbij bezorgden wij scholen ook passende kaders om tot een correcte interpretatie van de data te komen.

d. Van de 'toezegging tot deelname' tot een definitief aantal deelnemende scholen

Er was een duidelijke interesse bij de aangeschreven scholen om aan 'zelfevaluatie binnenstebuiten' deel te nemen, ook bij scholen die niet tot de doelpopulatie behoorden.

Daarom werd in een eerste ronde (mei-juni 2006) telefonisch of per email gevraagd naar bewijskracht voor het gebruik van een schriftelijk instrument voor zelfevaluatie. In het geval van een positief antwoord resulteerde dit in een principiële bevestiging van deelname door de school. Er bereikte ons 110 zulke principiële toezeggingen.

In september 2006 werden de schoolleiders van de scholen die principieel toezegden, uitgenodigd om een inleidende vragenlijst in te vullen. Dit was een voorbereidende stap die door één iemand (sterk betrokken bij de zelfevaluatie) diende uitgevoerd te worden. Het doel was achtergrondinformatie en praktische gegevens in te winnen over de school en de zelfevaluatie die er had plaatsgevonden. Door deze gegevens in eerste instantie bij één iemand te verzamelen kon de eigenlijke bevraging van de andere teamleden (in de loop van oktober) enigszins ingekort worden. De inleidende vragenlijst richtte zich op vragen zoals het tijdstip, de functies en de instrumenten van de zelfevaluatie. Ook de bevraagde inhouden en genoten ondersteuning werd in kaart gebracht. Tenslotte werd gevraagd een contactpersoon voor het verdere verloop van de studie naar voren te schuiven. De taken van de interne contactpersoon waren in hoofdzaak het verzorgen van de communicatie met de onderzoekers, en het ontvangen en verdelen van paswoorden. Tijdens de bevraging bestond hun rol uit eerstelijnshulp enerzijds en informeren en motiveren anderzijds. Om de respons in goede banen te leiden, werden de contactpersonen immers gevraagd om via de bestaande communicatiekanalen ruchtbaarheid te geven aan de bevraging.

103 van de 110 scholen die uitgenodigd werden de inleidende vragenlijst in te vullen, hebben dit ook effectief gedaan. Deze uitval is te verklaren door directiewissels, het herzien van de toezegging bij het lezen van gedetailleerde achtergrondinformatie en laattijdig reageren. Uiteindelijk waren het dus 103 scholen die het nodige materiaal ontvingen om aan de eigenlijke bevraging deel te nemen. Dat gebeurde begin oktober 2006. Voor elk personeelslid in deze scholen werd een unieke uitnodiging tot deelname aangemaakt. Deze bevatte specifieke achtergrondinformatie over de schooleigen zelfevaluatie (tijdstip van de zelfevaluatie, gebruikte instrument, geëvalueerde thema's, schooleigen benaming van de zelfevaluatie) en een uniek paswoord. Ruim 5000 van dergelijke uitnodigingen werden in 103 pakketten verstuurd naar de interne contactpersonen in de verschillende scholen.

Ook in deze groep was er nog uitval. Deze uitval is in de eerste plaats te verklaren door het herzien van de toezegging bij het doorlezen van de concrete vragenlijst. Een aantal schoolleiders was van oordeel dat de zelfevaluatie niet meer voldoende recent was om op deze intensieve manier bevraagd te worden. Anderen gaven aan dit engagement niet te hebben kunnen opnemen naast (onverwachts) andere prioriteiten (zoals concrete dagelijkse problemen of een doorlichting) of door ziekte. Nog anderen gaven verder geen enkele motivering. Van de 103 scholen waren er 5 waarvan uiteindelijk niemand de vragenlijst invulde. In 2 scholen was de respons te beperkt om betrouwbare uitspraken te kunnen doen op schoolniveau. Dat betekent dat in de analyses telkens sprake zal zijn van 96 deelnemende scholen met volwaardige gegevens. In 57 van deze scholen werden ten minste 20 teamleden bevraagd. De 39 resterende scholen waren doorgaans kleinere scho-

len. In de groep van scholen met 15 of minder respondenten werden bijvoorbeeld gemiddeld vier van de vijf teamleden bevraagd. Deze groep van 96 scholen en de teamleden die er in functioneren, zullen wij in het vervolg omschrijven als de steekproef van 'zelfevaluatie binnenstebuiten'. In totaal namen 2716 respondenten deel aan de survey. Dat maakt gemiddeld ruim 28 respondenten per deelnemende school. Dat aantal varieert van 7 tot 97. In paragraaf 2.3.3 beschrijven we de steekproef.

2.3.2 Ontwikkeling van de vragenlijst

Cruciaal bij het gebruik van een survey is het beperken van meetfouten door een doordacht gebruik van goede instrumenten. In deze paragraaf staan we stil bij onze strategische stappen om de kwaliteit van de gebruikte instrumenten te bewaken. Over elke stap wordt in de volgende alinea's kort gerapporteerd.

a. Inventarisatie van de te bevragen variabelen

De eerste fase van de ontwikkeling van de vragenlijst bestond erin te inventariseren over welke aspecten data verzameld dienden te worden. Naast een beschrijving van de feitelijke zelfevaluatiepraktijken in scholen moest de vragenlijst het ook mogelijk maken om verbanden te leggen met schoolgebonden kenmerken. Daartoe werden de variabelen die in het theoretisch kader beschreven werden, geoperationaliseerd en zonodig toegepast op het uitvoeren van zelfevaluaties in scholen. De inhoud van de vragenlijsten berust dus ook in sterke mate op theorieën, eerdere empirische vaststellingen en hypothesen uit de literatuurstudie. De resultaten van het operationaliseringsproces werden geïntegreerd in het bespreken van het theoretisch kader (zie deel 4).

b. Bewaken van de inhoudsvaliditeit van de vragenlijst

Het is van essentieel belang dat de vragenlijsten de beoogde verschijnselen in scholen meten. De inhoud van de variabelen moet maximaal overeenstemmen met de inhoud van het theoretische concept dat bevraagd wordt. Om dat te garanderen hebben wij ons verdiept in de te bestuderen fenomenen. Het bijwonen van studiedagen, verkenning van instrumenten, lectuur van beleidsdocumenten en omzendbrieven, informele gesprekken met bevoorrechte getuigen en participatie aan de onderzochte verschijnselen maakten het mogelijk de achterliggende theoretische concepten te doorgronden en in hun complexiteit te analyseren. Deze theoretische en ervaringsdeskundigheid werd in deze stap gecombineerd met discussies met collega's en met het bevragen van deskundigen. Eerst werd de versie van de vragenlijst voor commentaar aan collega-onderzoekers voorgelegd. Na verwerking van hun revisies werd de stap naar experts gezet (zie pilootstudie).

c. Bewaken van de technische kwaliteit van de vragenlijst

Doorheen de vorige stappen werd de vragenlijst voortdurend aangepast op basis van de ontvangen commentaren. Deze commentaren richtten zich voornamelijk op inhoudelijke aspecten van de vragenlijst. Pas in een latere fase werd expliciet aandacht besteed aan de technische kwaliteit van de vragenlijst in het algemeen en van de formulering van de

items in het bijzonder. Aan de hand van een daartoe opgestelde lijst met criteria werd elk item kritisch bestudeerd en eventueel aangepast. Centraal daarin stonden kwaliteitseisen zoals de ééndimensionaliteit en de duidelijkheid van de items.

Bij het bewaken van de technische kwaliteit van de vragenlijst werd tevens uitvoerig stilgestaan bij de keuze van een geschikte schaal om de items in de vragenlijst te scoren. Er werd uiteindelijk voor het grootste deel geopteerd voor de schaal: 1 = 'totaal niet', 2 = 'eerder niet', 3 = 'noch niet/noch wel', 4 = 'eerder wel', 5 = 'totaal wel', en 'WN/NVT = Weet niet/niet van toepassing'. Deze schaal heeft het voordeel dat de antwoorden van respondenten die het antwoord op een vraag om gelijk welke reden schuldig (moeten) blijven uit de Likert-schaal gefilterd worden door de restcategorie 'weet niet' (Demaeyer & Vanhoof, 2005). De opname van een middenpunt was wenselijk omdat respondenten ten aanzien van de voorliggende items een neutrale houding kunnen hebben.

d. De pilootstudie

Alvorens de vragenlijsten te gebruiken voor de uiteindelijke survey werd de geschiktheid ervan afgetoetst bij de finale doelgroep. Daartoe werden de vragenlijsten inclusief de inleidende brief ter discussie voorgelegd aan een toevallige selectie van potentiële respondenten. Het doel van de pilootstudie was in de eerste plaats een kwaliteitscontrole van de vragenlijst. Daarbij stonden drie groepen van kwaliteitscriteria centraal. (1) Is de vragenlijst voldoende duidelijk? Zijn de verschillende aspecten van de vragenlijst helder en bevattelijk voor de respondenten? De aandacht werd daarbij gericht op de inleidende instructies, de aard van de vraagstellingen en het woordgebruik in vragen en items. (2) Is de vragenlijst voldoende accuraat? Sluit de vragenlijst inhoudelijk aan bij het te onderzoeken thema? Zijn de respondenten van mening dat de bevraagde thema's voldoende relevant zijn om de zelfevaluatie in hun school in kaart te brengen en zijn bij opsommingen alle mogelijke antwoorden opgenomen? (3) Is het invullen van de vragenlijst voldoende haalbaar? Is de belasting van de respondenten bij het invullen van de vragenlijst aanvaardbaar? Daarbij wordt zowel aandacht besteed aan de duur die het invullen van de vragenlijst vereist als aan het conceptuele niveau van de vragenlijst.

Bij de pilootstudie werden 6 respondenten betrokken. Zowel 2 leden van een middenkader als 4 leerkrachten werden gevraagd om te reflecteren over de verschillende kwaliteitscriteria voor de vragenlijst. De respondenten die gevraagd werden de vragenlijst in te vullen, hebben dat zonder noemenswaardige problemen gedaan. Tijdens de interviews werden desalniettemin ook een aantal bemerkingen en suggesties naar voren geschoven, die ons in staat stelden de kwaliteit van de vragenlijst te verbeteren. Denk daarbij aan het aanvullen van antwoordmogelijkheden, het concretiseren van een aantal items of het verduidelijken van de lay-out. Na het verwerken van de resultaten van de pilootstudie werd de vragenlijst gefinaliseerd.

2.3.3 Beschrijving van de steekproef

De steekproef van 'Zelfevaluatie binnenstebuiten' bestaat uit 96 scholen (zie Tabel 8). 51 hiervan zijn scholen voor basisonderwijs, 45 verzorgen secundair onderwijs. De groep

scholen uit het basisonderwijs telt 9 scholen voor lager onderwijs, 2 autonome kleuterscholen en 40 basisscholen. In het secundair onderwijs gaat het voornamelijk om scholen met een boven- én onderbouw. Er is ook 1 school voor buitengewoon secundair onderwijs (BuSO). Uit de meest rechtse kolom van Tabel 8 blijkt verder dat de meeste scholen met een bovenbouw zich enkel op ASO richten of een combinatie van onderwijsvormen aanbieden (in hoofdzaak ASO/TSO/BSO of TSO/BSO).

Tabel 8: Verdeling van de steekproefscholen (absolute aantallen en percentages)

Onderwijsniveau		Verdeling BaO		Verdeling SO: Type		Verdeling SO: Onderwijsvorm	
BaO	51 (53%)	Kleuter	2 (4%)	Boven én onderbouw	22 (49%)	Middenschool	13 (29%)
SO	45 (47%)	Lager	9 (18%)	Bovenbouw	9 (20%)	ASO	11 (24%)
Totaal	*96 (100%)*	basis	40 (78%)	Middenschool	13 (29%)	ASO/TSO/BSO	10 (22%)
		Totaal	*51 (100%)*	BuSO	1 (2%)	TSO/BSO	5 (11%)
				Totaal	*45 (100%)*	ASO/TSO	3 (7%)
						ASO/TSO/BSO/KSO	1 (2%)
						TSO/BSO/KSO	1 (2%)
						BuO	1 (2%)
						Totaal	*45 (100%)*

Iets meer dan de helft van de deelnemende scholen (56.3%) maakt deel uit van het Gesubsidieerd Vrij Onderwijs, één derde (32.3%) van het Gemeenschapsonderwijs. De resterende 11.5% van de scholen behoren tot het Gesubsidieerd Officieel Onderwijs. Of de verdelingen in de steekproef representatief zijn voor de populatie van deze studie kunnen we niet nagaan, vermits deze ongekend is.

In de 96 deelnemende scholen werden 2716 teamleden bevraagd. In de volgende alinea's beschrijven we de samenstelling van deze groep respondenten.

In eerste instantie kijken we naar de functies die de respondenten vervullen in de deelnemende scholen (zie Tabel 9). Het overgrote deel van de respondenten zijn leraren. 83% geeft aan het meeste tijd te spenderen aan zijn/haar functie als leraar. Daarnaast is 2.1% GOK-leraar. De schoolleiding maakt in totaal 7.3% uit van de steekproef. 3.9% van de respondenten is directeur of adjunct-directeur; 3.4% geeft aan hoofdzakelijk bezig te zijn met taken van het middenkader (beleidsondersteuning, technisch adviseur, ...). Tenslotte is er ook nog een noemenswaardige groep ondersteuners. 4.7% van de respondenten houdt zich bezig met administratieve of ICT- ondersteuning. Iets minder dan 3% is verantwoordelijk voor pedagogische, paramedische ondersteuning op de school die wij onderzochten.

Het rechtse deel van Tabel 9 geeft de antwoorden weer op de vraag 'of' en zo ja 'waar' de respondenten in hoofdzaak hun lesopdracht vervullen. 11.5% geeft aan geen lesopdracht te vervullen. Bijna één op de drie respondenten vervuld de lesopdracht in het basisonderwijs, waarvan 8.7% in het kleuteronderwijs en 23.5% in het lager onderwijs. 56.3% van de respondenten geeft les in het secundair onderwijs. Het grootste deel doet dit in de 3de graad (21.5%), gevolgd door de 1ste graad (18.8%) en vervolgens de 2de graad

(14.5%). Tenslotte blijkt ook 1,4% van de respondenten een lesopdracht uit te voeren die zich hoofdzakelijk in de 4de graad situeert.

Tabel 9: Verdeling van respondenten over functie en aard lesopdracht

Functie in de school				Kern van de lesopdracht		
	Aantal	%			Aantal	%
(Adjunct-)Directeur	106	3.9		Geen	313	11.5
Middenkader	92	3.4		kleuteronderwijs	237	8.7
Leraar	2254	83.0		lager onderwijs	637	23.5
GOK-leraar	57	2.1		SO 1ste graad	510	18.8
Ped.of paramedische ondersteuning	80	2.9		SO 2de graad	394	14.5
Adm. of ICT ondersteuning	127	4.7		SO 3de graad	588	21.6
Totaal	**2716**	**100 %**		SO 4de graad	37	1.4
				Totaal	**2716**	**100 %**

In totaal is 63.8% van de respondenten bevraagd als teamlid van een secundaire school. 36.2% maakt deel uit van een school voor basisonderwijs. Op schoolniveau was deze verhouding respectievelijk 53% vs. 47%. Van de respondenten uit het secundair onderwijs is de grootste groep tewerkgesteld in het ASO. Ruim de helft (51%) geeft aan voornamelijk in deze onderwijsvorm actief te zijn. Precies een kwart is actief in het beroepsonderwijs (BSO) en 22% in het technisch onderwijs. Bijna 2% van de respondenten werkt in het Kunst Secundair Onderwijs (KSO). De gemiddelde leeftijd van de respondenten is 40 jaar, met een minimum van 20 en een maximum van 64.

We hebben hoger beargumenteerd dat het niet mogelijk is in de context van deze studie tot een representatieve steekproef te komen. We wilden in de eerste plaats verzekeren dat de verwachte variatie in de doelpopulatie zich ook in de steekproef zou manifesteren. Dit blijkt het geval voor de kenmerken waarover hierboven gerapporteerd werd.

2.4 Afrondend: een vergelijking van de 3 gegevensverzamelingen

Om het overzicht te bewaken, hebben we de kerngegevens over de drie gegevensverzamelingen samengebracht in twee tabellen. Tabel 10 richt zich op de aard en context van de gegevensverzamelingen.

Tabel 10: Aard en context van de drie gegevensverzamelingen

	GOK-zelfevaluatie	GOK-inspectie	Zelfevaluatie binnenstebuiten
Uitvoerders	Universiteit Antwerpen en KULeuven n.a.v. OBPWO-onderzoek 02.05	Inspectie van het Vlaamse Onderwijs	Universiteit Antwerpen

	GOK-zelfevaluatie	GOK-inspectie	Zelfevaluatie binnenstebuiten
Bevraagde scholen	Representatieve steekproef van alle GOK-scholen	Alle GOK-scholen	Scholen die recent gebruik maakten van een zelfevaluatie-instrument
Gegevensverzameling	Vragenlijst ingevuld door de schoolleider	Doorlichting door een inspectieteam	Vragenlijst ingevuld door de schoolleiding, het middenkader en leerkrachten
Totale aantal respondenten	518 deelnemende scholen met telkens 1 respondent	Ruim 2700 scholen	96 deelnemende scholen met in totaal 2716 respondenten
Tijdstip gegevensverzameling	Ronde 1: Voorjaar 2003 Ronde 2: Voorjaar 2004	Voorjaar 2004	Najaar 2006

Tabel 11 richt zich op de inhoud van de drie gegevensverzamelingen. We geven aan waar de verschillende elementen uit het theoretisch kader aan bod komen.

Tabel 11: Inhoud van de drie gegevensverzamelingen

			GOK-zelf-evaluatie	GOK-inspectie	Zelfevaluatie binnen-stebuiten
School-gebonden kenmerken	De organisatie-effectiviteit	Effectiviteits-perspectieven			x
		Beleidsvoerend vermogen	x		
	De lerende organisatie	Leeropvattingen			x
		Leerresultaten			x
Zelfevaluatie-gebonden kenmerken	Houding t.a.v. zelfevaluatie				x
	Zelfevaluatie als beleidsdaad		x	x	x
	Zelfevaluatie als onderzoeksdaad			x	x
Ondersteuning					x
Kwaliteit van de zelfevaluatie			x	x	x

3 VOORBEREIDENDE ANALYSES

In de kwantitatieve analyses maken we vaak gebruik van schalen. Wij hebben ervoor gekozen om het beschrijven van de ontwikkeling van deze schalen los te koppelen van de eigenlijke resultaten. In dit deel met voorbereidende analyses wordt dan ook in eerste instantie gerapporteerd over de schaalconstructie. Daarnaast staan we nog stil bij de clusteranalyses die uitgevoerd werden.

3.1 Schaalconstructie

Het gebruik van meetschalen heeft als voordeel dat de betrouwbaarheid van het meetinstrument verhoogd wordt. De theoretische concepten uit het conceptuele kader werden

bij het ontwikkelen van de vragenlijst daarom geoperationaliseerd in verschillende items. Er wordt verondersteld dat deze items verschillende aspecten van hetzelfde achterliggende theoretische concept meten. Als dat het geval is spreken we van constructvaliditeit. De schaal wordt ook verondersteld intern consistent en ééndimensioneel te zijn. Hieronder beschrijven we hoe we dit onderzochten en hoe we dit zullen rapporteren in de daartoe gebruikte tabellen.

De logica binnen de gepresenteerde tabellen is de volgende. Bij het ontwikkelen van de vragenlijst werd het theoretische concept waarin we geïnteresseerd zijn, geoperationaliseerd op basis van een oorspronkelijk aantal items. In het volgende toetsen we in eerste instantie of onze theoretische verwachtingen bevestigd worden door de verzamelde gegevens. Dat doen we in hoofdzaak door middel van confirmatieve factoranalyses (door middel van structural equation modelling - SEM). Soms is het mogelijk schalen te vereenvoudigen door items te verwijderen zonder de interne consistentie noemenswaardig te verslechteren (of zelfs te verbeteren). De selectie van items gebeurde dan op basis van de gestandaardiseerde factorscores. Deze analyses leidden tot een schaal met een finaal aantal items met een eigen Cronbach's alpha-waarde.

De schaalanalyses werden uitgevoerd met het softwarepakket AMOS, bedoeld voor SEM. In het toetsen van schalen gebruiken wij de volgende vier fitindices: de chi-kwadraat-index (CMIN), de Adjusted Goodness of Fit Index (AGFI), Comparative Fit Index (CFI) en de Root Mean Square Error of Approximation (RMSEA). Een model (c.q. schaal) is goed wanneer de met CMIN geassocieerde p-waarden groter is dan 0.05. Toch wordt deze Fit-index vooral voor de volledigheid opgenomen, vermits deze bij grote steekproeven niet meer betrouwbaar is (Raykov & Marcoulides, 2000). Dat geldt niet of in mindere mate voor de andere fit-indicatoren. De Goodness of Fit Index (GFI) geeft aan hoe goed het model de variantie in de data kan verklaren (cf. R^2 in regressieanalyse). Wij gebruiken meestal de Adjusted Goodness of Fit (AGFI) omdat deze rekening houdt met het aantal parameters in het model. De Comparative Fit Index (CFI) vergelijkt het vooropgestelde model met een model waarin er geen relaties verondersteld worden (cf. het nulmodel). De bijhorende waarde geeft aan welk percentage van de misfit in rekening kan worden gebracht door van het nulmodel naar het vooropgestelde model te gaan. De richtlijnen in de literatuur geven aan dat een model goed past wanneer AGFI groter is dan 0.90 en wanneer CFI groter is dan 0.95 (Hu & Bentler, 1999). De Root Mean Square Error of Approximation (RMSEA) tenslotte geeft aan hoe goed het model aansluit bij de werkelijke situatie in de populatie, moest deze gekend zijn (Byrne, 2001). De afwijking wordt per vrijheidsgraad weergegeven en is dus best zo klein mogelijk. Een waarde onder 0.05 wijst op een goede fit, waarden onder 0.08 en 0.10 zijn middelmatig. Wij beoordelen de RMSEA door te kijken naar Pclose. Deze test toetst de nulhypothese dat RMSEA niet groter is dan .05. Als Pclose groter is dan .05, verwerpen we de nulhypothese en spreken we dus van een goede fit.

Voor de schalen die resulteren uit de SEM-analyses werden nog twee bijkomende analyses uitgevoerd. Eerst werd Cronbach's alpha berekend. Dit is een klassieke maat voor de interne consistentie van een schaal. Als vuistregel wordt gehanteerd dat een schaal kan

worden gebruikt bij een alpha van 0.70 of hoger. Om vervolgens de ééndimensionaliteit van deze schalen te controleren werd voor elke schaal tevens een exploratieve factoranalyse uitgevoerd. Interne consistentie is immers geen garantie voor ééndimensionaliteit. Factoranalyse maakt het mogelijk te achterhalen of in de items van een schaal niet meer dan één factor onderkend moet worden. Indien de schaal werkelijk ééndimensioneel is (Kaiser, 1960), zal slechts één factor weerhouden worden. Een 1 in de kolom FA (factoranalyse) wijst op de ééndimensionaliteit van de schaal.

Zoals hierboven blijkt, spitsen wij onze analyses in het toetsen van de schalen in eerste instantie toe op het niveau van de individuele respondenten. Dat impliceert niet dat wij de hiërarchische structuur van de dataset uit het oog verliezen. In de beschrijvende en verklarende analyses wordt resoluut gekozen voor het werken met de meerniveaumodellen. Op die manier geven we gevolg aan de verschillende redenen om te kiezen voor multilevel analyses (Hox, 2002; De Maeyer & Rymenans, 2004). In de fase van de schaalconstructie wordt echter enkel een algemene indicatie gegeven van de betrouwbaarheid op schoolniveau. Dat doen we door voor de geaggregeerde gegevens uit 'zelfevaluatie binnenstebuiten' over de scholen heen een Cronbach's alpha waarde te berekenen. In de tabellen wordt deze (waar relevant) tussen haakjes naast de Cronbach's alpha waarde op respondentenniveau gerapporteerd.

3.1.1 Schoolgebonden kenmerken

Deze paragraaf rapporteert over de ontwikkeling van schalen voor de schoolgebonden variabelen. Er worden vijf inhoudelijke thema's onderscheiden.

a. Dragers van beleidsvoerend vermogen

In de constructie van schalen voor het meten van het beleidsvoerend vermogen van de GOK-scholen hanteren we twee invalshoeken. Eerst bespreken we de afzonderlijke dragers om daarna het achterliggende concept 'beleidsvoerend vermogen' onder de loep te nemen. We maken van de beschrijving van deze eerste schalen gebruik om onze werkwijze toe te lichten.

- *Een schaal voor elke drager*

In eerste instantie werd voor elke drager van beleidsvoerend vermogen een eigen schaal ontwikkeld (zie Tabel 12). In het theoretisch kader lichtten we toe waarom we bij twee dragers gebruik maken van twee subschalen. Gezamenlijke doelgerichtheid valt immers uiteen in 'visie' en 'gezamenlijke doelgerichtheid'. In responsiviteit maken we een onderscheid tussen responsiviteit ten aanzien van ouders en ten aanzien van de ruimere omgeving.

Tabel 12: Schaalconstructie dragers 'beleidsvoerend vermogen' (BVV)

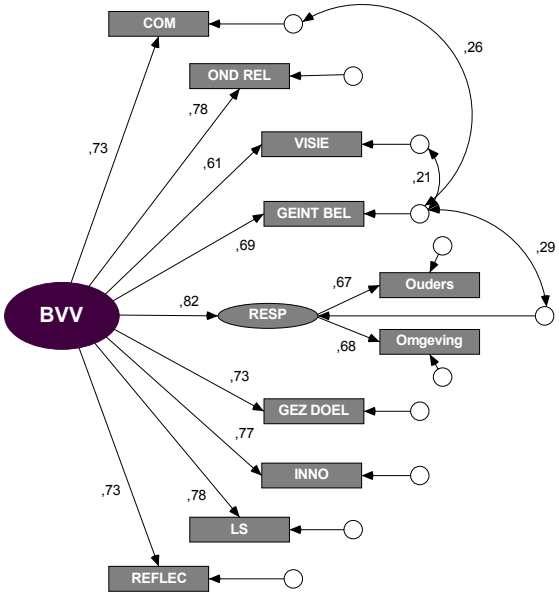
Dragers BVV	Aantal items	N	SEM						alpha	FA
			CMIN	p	AGFI	CFI	RMSEA	pclose		
Communicatie	4	501	3.89	0.14	0.98	0.99	0.044	0.46	0.83	1
Ondersteuning	7	495	24.03	0.01	0.97	0.99	0.049	0.48	0.81	1
Leiderschap	8	491	36.04	0.01	0.96	0.98	0.054	0.37	0.82	1
Visie	5	524	3.66	0.30	0.99	1.00	0.021	0.73	0.87	1
Gezamenlijke doelgerichtheid	5	495	11.75	0.02	0.96	0.99	0.063	0.26	0.81	1
Innovatief vermogen	4	442	1.05	0.31	0.99	1.00	0.011	0.53	0.87	1
Responsiviteit t.a.v. ouders	4	443	4.61	0.10	0.97	1.00	0.054	0.36	0.80	1
Responsiviteit t.a.v. omgeving	4	443	2.52	0.11	0.97	1.00	0.059	0.30	0.85	1
Geïntegreerd beleid	4	505	1.55	0.21	0.99	1.00	0.033	0.46	0.78	1
Reflectief vermogen	6	442	4.16	0.66	0.99	1.00	0.001	0.95	0.87	1

De resultaten in Tabel 12 tonen aan dat voor elk van de vooropgestelde concepten een intern consistente en ééndimensionele schaal kon ontwikkeld worden. De verwachtingen inzake constructvaliditeit worden ingelost.

- *Het concept 'beleidsvoerend vermogen'*

We hebben hierboven telkens gesproken over de dragers van beleidsvoerend vermogen. Deze terminologie gaat uit van de veronderstelling van het bestaan van een achterliggend concept 'beleidsvoerend vermogen' dat zich uit in de verschillende dragers. De gegevensverzameling in het kader van de GOK-zelfevaluatie stelt ons in staat deze veronderstelling te toetsen aan de empirie. We kunnen nagaan of het steek houdt dat de variantie in de verschillende dragers verklaard kan worden door een achterliggend concept. Figuur 11 visualiseert dit model en geeft meteen ook de bijhorende regressiecoëfficiënten weer. De in de rechthoeken weergegeven variabelen zijn de 10 schaalscores die hierboven ontwikkeld werden. We hebben het model zo opgevat dat de twee (gemeten) schalen voor responsiviteit (cf. ouders en de ruimere omgeving) beide afhangen van de meer omvattende maar latente (niet gemeten) variabele 'responsiviteit'.

Figuur 11: Meetmodel (dragers van) beleidsvoerend vermogen



De theoretische verwachtingen in Figuur 11 worden door de empirische evidentie niet tegengesproken (RMSEA: 0.04 (pclose: 0.61); AGFI: 0.97 ; CFI: 0.99)[7]. De theoretische verwachtingen blijken dus te stroken met de verzamelde gegevens. Uit de hoge ladingen blijkt dat een groot deel van de variantie in de verschillende dragers wordt verklaard door 'het beleidsvoerend vermogen' van scholen. De gestandaardiseerde regressiecoëfficiënten variëren van 0.61 tot 0.82. Scholen die - volgens de schoolleider - hoog scoren op één drager, zullen dus ook hoog scoren op de andere dragers. Scholen die ergens laag scoren, zullen ook laag scoren op de andere dragers. Dit effect loopt in sterke mate via het beleidsvoerend vermogen van scholen. Gemiddeld verklaart het beleidsvoerend vermogen 53% van de variantie in de verschillende dragers. Uitschieters zijn enerzijds de drager 'visie' met 37% en de dragers 'responsief vermogen' met 67%.

Op basis van de regressiecoëfficiënten kan voor elke school een factorscore voor het beleidsvoerend vermogen berekend worden. Deze indicator zal in het verdere verloop van de analyses onder meer gebruikt worden voor het classificeren van GOK-scholen.

[7] Wanneer het aantal respondenten te groot is om op de Chi^2-toets te vertrouwen, worden de resultaten van deze toets in de tekst niet meer systematisch gerapporteerd.

b. De effectiviteitsperspectieven in scholen

Om vlot te kunnen functioneren moet een school volgens Quinn & Rohrbaugh (1983) in voldoende mate aan vier vereisten voldoen. Deze vereisten zijn: zorgen voor duidelijke spelregels (het interne proces model), zorgen voor hechte (vergader)teams (het human relations model), zorgen voor een tijdige aanpassing van doelen (het open systeem model) en zorgen voor resultaatsgerichte inspanningen (het rationele doel model). In de vragenlijst 'zelfevaluatie binnenstebuiten' gebruikten wij een ingekorte versie van de operationalisatie van Maslowski (2001). Om de structuur in de lijst van 28 waarden te exploreren, voerden we een factoranalyse uit op respondentniveau (met Varimax rotatie). Vier factoren blijken een eigenwaarde te hebben groter dan 1 (Kaiser, 1960). Tabel 13 bevat de resultaten van deze analyse.

Tabel 13: Factorstructuur in de effectiviteitsperspectieven

	Factor 1	Factor 2	Factor 3	Factor 4
Stabiliteit	**0.59**	0.39	0.00	0.14
Continuïteit	**0.57**	0.38	0.07	0.12
Formele aanpak	**0.68**	-0.05	0.11	0.13
Rechtlijnigheid	**0.71**	0.27	-0.02	0.13
Controle	**0.67**	-0.10	0.24	0.15
Reglementering	**0.72**	-0.01	0.19	0.19
Zekerheid	**0.60**	0.33	-0.01	0.20
Engagement	0.08	**0.57**	0.41	0.14
Behulpzaamheid	0.09	**0.70**	0.37	0.12
Wederzijds vertrouwen	0.11	**0.82**	0.26	0.12
Loyaliteit	0.11	**0.79**	0.27	0.12
Wederzijds begrip	0.12	**0.81**	0.28	0.11
Aangenaam klimaat	0.15	**0.80**	0.14	0.07
Samenwerking	0.13	**0.72**	0.35	0.10
Vernieuwingsgezindheid	0.07	0.15	**0.77**	0.10
Flexibiliteit	0.07	0.23	**0.79**	0.06
Diversiteit	0.08	0.22	**0.77**	0.08
Op de omstandigheden inspelen	0.14	0.31	**0.70**	0.14
Openheid naar nieuwe ontwikkelingen	0.09	0.19	**0.81**	0.10
Veelzijdigheid	0.09	0.22	**0.79**	0.17
Aanpassingsvermogen	0.11	0.29	**0.76**	0.13
Resultaatgerichtheid	0.12	0.03	-0.01	**0.81**
Effectiviteit	0.20	0.24	0.17	**0.72**
Prestatiegerichtheid	0.13	-0.08	0.00	**0.80**
Efficiëntie	0.25	0.29	0.21	**0.67**
Doelgerichtheid	0.20	0.32	0.28	**0.64**
Succes nastreven	0.14	0.02	0.15	**0.72**
Bekwaamheid	0.14	0.42	0.21	**0.54**

Deze vier factoren sluiten duidelijk goed aan bij de vier effectiviteitsperspectieven die in eerder onderzoek gevonden werden. Zij verklaren tezamen 63% van de variantie in de verschillende items. De minst hoge lading bedraagt 0.54 en er zijn geen ernstige problemen met ambigue items. De schalen kunnen dan ook met het nodige vertrouwen gebruikt worden voor de verdere analyses (zie Tabel 14). Ze zijn intern consistent op respondent én schoolniveau en ééndimensioneel.

Tabel 14: Schaalconstructie 'Effectiviteitsperspectieven'

Schaal	Aantal items	N	SEM						alpha	FA
			CMIN	p	AGFI	CFI	RMSEA	pclose		
Interne proces model	7	2042	30.2	0.00	0.98	1.0	0.05	0.48	0.82 (0.88)	1
Human relations model	7	2042	31.9	0.00	0.98	1.0	0.04	0.80	0.93 (0.96)	1
Open systeem model	7	2042	11.8	0.02	0.99	1.0	0.03	0.93	0.92 (0.96)	1
Rationele doel model	7	2042	17.1	0.01	0.99	1.0	0.03	0.91	0.89 (0.92)	1

Op basis van de schaalscores bekijken we in de voorbereidende analyses enkel nog de relatie tussen de verschillende effectiviteitsperspectieven. We hebben daartoe in Tabel 15 de bivariate correlatiecoëfficiënten op respondent- en schoolniveau opgenomen. Daaruit blijkt dat respondenten (scholen) die hoog scoren op het ene perspectief doorgaans ook hoog scoren op de andere perspectieven. Eerder onderzoek naar deze relaties is tegenstrijdig en geeft geen duidelijkheid over de richting en de sterkte van de bestudeerde verbanden (Quinn & Rohrbaugh, 1983; Van Muijen, 1994; Maslowski, 2001).

Tabel 15: Bivariate correlaties op respondent-(en school)niveau

	Interne proces model	Human relations model	Open systeem model	Rationele doel model
Interne proces model	1			
Human relations model	.40 (.38)	1		
Open systeem model	.31 (.38)	.59 (.72)	1	
Rationele doel model	.48 (.57)	.40 (.20)	.37 (.19)	1

c. Sociaalconstructivistische visie op leren

We brachten in de vragenlijst ook de perceptie van de respondenten ten aanzien van hun eigen leren in kaart. Dat deden we door de verschillende elementen die als kenmerkend voor het sociaalconstructivisme naar voren worden geschoven, te vertalen in contrasterende zinsneden (zie theoretisch kader). Een lage score wijst op het onderschrijven van de 'klassieke' zinsnede, en een hoge score wijst op het onderschrijven van het sociaalconstructivistische gedachtegoed. Elke vraag richtte zich op een ander kenmerk van de sociaalconstructivistische visie op leren. In deze voorbereidende analyse gaan wij na of de acht items ook effectief het achterliggende concept 'sociaalconstructivistische visie op leren' meten. Dat blijkt voor een schaal met 6 items het geval te zijn. In eerste instantie lijken de fit-indicatoren in Tabel 16 te wijzen op een aanvaardbaar tot goed model. Voorzichtigheid wordt echter ingegeven door de eerder lage regressiegewichten van

de manifeste variabelen op het latente concept 'sociaalconstructivistische visie'. Ook de Cronbach's alpha geeft aan dat deze schaal met de nodige reserves moet worden gebruikt.

Tabel 16: Schaalconstructie 'Sociaalconstructivistische visie op leren'

Schaal	Aantal items	n	SEM						Cronbach's alpha	FA
			CMIN	p	AGFI	CFI	RMSEA	pclose		
Soc. construct. visie	6	2519	27.2	0.00	0.99	0.99	0.04	0.22	0.60	1

d. 'Single-loop' en 'double-loop'

Om na te gaan hoe diepgaand scholen leren werd aan de respondenten gevraagd om voor een aantal mogelijke reacties bij problemen in de school aan te geven hoe vaak deze voorkomen. Zij konden de reacties beoordelen aan de hand van een 5-puntenschaal: 1= nooit, 2= zelden, 3= soms, 4= regelmatig en 5 = altijd. Het gaat hier om ordinale gegevens zodat werken met 'Asymptotically Distribution Free' (ADF) schatting aangewezen is in plaats van Maximum likelihood (Kupek, 2005).

De items in de schaal 'reacties bij problemen' vallen vanuit theoretisch oogpunt uiteen in twee groepen. Een eerste groep verwijst naar niet constructief geachte reacties (bijvoorbeeld: het probleem niet erkennen of minimaliseren, de (echte) kern van de zaak onbesproken laten en informatie achterhouden). De tweede groep omvat de wel constructief geachte reacties (zoals de verschillende meningen van betrokkenen verzamelen, onderlinge spanningen en meningsverschillen bespreekbaar maken en objectieve gegevens verzamelen om het probleem in kaart te brengen). In een eerste stap hebben we scores van de niet constructieve items daarom gespiegeld. Een hoge score op de schaal betekent dus dat niet constructieve reacties op problemen niet frequent voorkomen en constructieve reacties wel frequent.

Twee van de oorspronkelijke 14 items blijken op basis van een te lage regressiecoëfficient best uit de schaal te worden verwijderd. De resterende schaal scoort zeer goed op de indicatoren die wij bij het construeren van schalen vooropstelden. De fit-indicatoren voor deze schaal zijn in Tabel 17 opgenomen.

Tabel 17: Schaalconstructie 'Single vs. double loop leren'

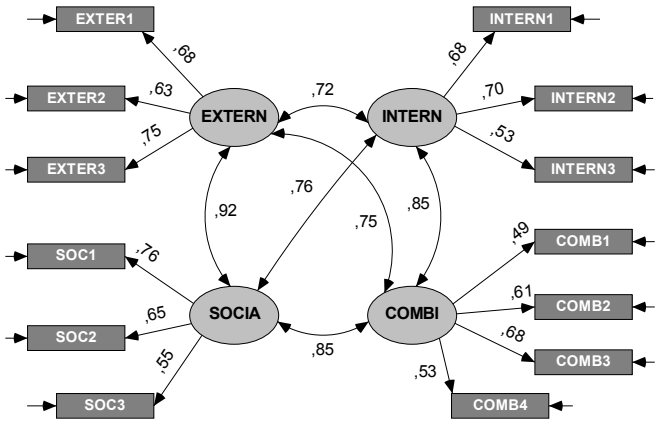
Schaal	Aantal items	n	SEM						Cronbach's alpha	FA
			CMIN	p	AGFI	CFI	RMSEA	pclose		
Single vs. double loop leren	12	2056	53.6	0.06	0.98	0.99	0.01	1.0	0.92 (0.97)	1

e. Vormen van kenniscreatie

De theorie van de kenniscreatie werd geoperationaliseerd door na te gaan in welke mate leeractiviteiten - die telkens binnen één van de vier onderdelen van de kennisspiraal

gesitueerd kunnen worden - plaatsvinden in scholen. De vier vormen zijn externaliseren, internaliseren, combineren en socialiseren. Het model in Figuur 12 brengt de gestandaardiseerde regressiecoëfficiënten van de items en de onderlinge correlaties tussen de vier vormen van kenniscreatie in beeld.

Figuur 12: Meetmodel 'vormen van kenniscreatie'



Gegeven de volgende fit-indicatoren gaat het om een goed model (RMSEA: 0.04 (pclose: 0.82); AGFI: 0.98 ; CFI: 0.97). Voor de duidelijkheid zijn de (correlaties tussen de) errortermen niet in de figuur opgenomen. De correlaties tussen de vier vormen van kenniscreatie zijn sterk. In een exploratieve factoranalyse worden de vier vormen dan ook niet zuiver teruggevonden. Onze theoriegedreven verwachtingen worden in SEM echter niet verworpen. De Cronbach's alpha-waarden voor de vier vormen wijzen op redelijk tot behoorlijke schalen. De waarden zijn 0.76 voor 'externaliseren' (3 items), 0.66 voor 'combineren' (4 items) , 0.68 voor 'internaliseren' (3 items) en 0.70 voor 'socialiseren' (3 items). Op schoolniveau bedraagt Cronbach's alpha respectievelijk 0.84, 0.77, 0.80 en 0.81. Om de lengte van de vragenlijst te bewaken werd het aantal items per vorm van kenniscreatie bewust beperkt gehouden. Dit heeft uiteraard een impact op deze scores. De Cronbach's alpha waarde voor de gehele schaal met 13 items bedraagt 0.87 (op schoolniveau 0.91). In verdere analyses zal met deze schaal gewerkt worden. Deze schaal meet 'de mate van kenniscreatie' in scholen.

3.1.2 Zelfevaluatiegebonden kenmerken

Naast de schalen die betrekking hebben op schoolgebonden kenmerken wordt in de analyses gebruik gemaakt van zelfevaluatiegebonden kenmerken. In de volgende paragrafen worden schalen beschreven om (a) de kwaliteit van zelfevaluatie in kaart te brengen, om (b) de houding van teamleden ten aanzien van zelfevaluatie te meten en om na te gaan of de bestudeerde zelfevaluaties voldoen aan de kenmerken van (c) 'zelfevaluatie als beleidsdaad' en (d) 'zelfevaluatie als onderzoeksdaad'.

a. De 'kwaliteit' van zelfevaluaties

Om het concept 'kwaliteit van de zelfevaluatie' te beschrijven, doen we een beroep op twee benaderingen: een schooleigen en een schooloverstijgende benadering.

De schooleigen benadering houdt in dat de respondenten gevraagd werd om aan te geven in welke mate de zelfevaluatie volgens hen geleid heeft tot het realiseren van de twee doelen die men zelf als het meest belangrijk bestempelde. Die twee doelen konden de respondenten selecteren uit een lijst van 10 zelfevaluatiedoelen (zie Theoretisch kader). De indicator voor de schooleigen benadering is dan de gemiddelde beoordeling voor het al dan niet realiseren van deze twee doelen. De beoordelingsschaal (5-punts) liep van 'totaal niet' tot 'totaal wel'. Beide vragen lijken consistent beantwoord te worden. Cronbach's alpha voor de schaal met deze twee items bedraagt 0.73 (op schoolniveau 0.92). We noemen deze schaal de 'schooleigen kwaliteit van de zelfevaluatie'.

De schooloverstijgende benadering bekijkt de kwaliteit van 'zelfevaluaties' als het al dan niet bereiken van een aantal resultaten waarover in de literatuur en tussen experts consensus bestaat dat ze nastrevenswaardig zijn. Deze resultaten werden geformuleerd als stellingen die al dan niet van toepassing kunnen zijn op de bestudeerde zelfevaluaties. Conceptueel maakten we in de stellingen een onderscheid tussen het al dan niet 'resulteren in passende (SMARTe) actiepunten' enerzijds en het 'resulteren in concrete resultaten' anderzijds. Een exploratieve factoranalyse leert dat de theoretische verwachtingen ondersteund worden door de data. De 3-factorenoplossing in Tabel 18 verklaart 70% van de variantie in de items. Uit deze factorstructuur blijkt dat de geroteerde data onze theoretische veronderstellingen omtrent de drie te onderscheiden kwaliteitsindicatoren niet tegenspreken[8]. De schaal 'Resulteren in passende (SMARTe) actiepunten' is intern consistent (Cronbach's alpha is 0.93 op respondentniveau en 0.97 op schoolniveau) en ééndimensioneel. Dat geldt ook voor de schaal 'Resultaten in concrete resultaten' (Cronbach's alpha is hier 0.95 en 0.98 op schoolniveau).

Ook bij de GOK-zelfevaluaties werd een schooloverstijgende kwaliteitsbenadering gebruikt. De kwaliteit van de GOK-zelfevaluatie werd hier enkel beoordeeld door de schoolleiders. Zij spraken zich uit over 6 stellingen omtrent de resultaten van de zelfevaluatie. Vijf van deze items blijken een goede schaal te vormen (CMIN: 6,26 (p=0.10); AGFI: 0.97; CFI: 0.99; RMSEA: 0.05 (pclose= 0.36). Cronbach's alpha bedraagt 0.74. De factorscore van scholen op deze indicator is gebaseerd op de volgende items: De zelfevaluatie geeft een goed zicht op de problemen die zich voordoen bij onze leerlingen, geeft een goed zicht op tekorten in het handelen van onze leerkrachten, geeft een goed zicht op de

[8] De volledigheid gebiedt ons mee te geven dat de niet geroteerde factormatrix in eerste instantie wijst op 1 factor waarop alle items behoorlijk sterk laden met daarnaast bijkomend twee relevante maar moeilijk te interpreteren factoren. Er is dus ook evidentie om te pleiten voor het gebruik van één achterliggend concept 'kwaliteit van de zelfevaluatie'. Deze aanpak maakt echter dat een noemenswaardig deel van de variantie in de items niet mee in de analyses opgenomen wordt. Bovendien hebben we goede theoretische gronden om aan het onderscheid tussen de drie kwaliteitsindicatoren vast te houden.

sterke punten in onze schoolwerking, heeft strategieën opgeleverd om GOK-problemen verder aan te pakken en zet ons ertoe aan in de toekomst meer aan zelfevaluatie te doen.

Tabel 18: Factorstructuur voor de kwaliteitsindicatoren

	Factoren		
Kwaliteitsindicator 1: de schooleigen benadering .	**1**	**2**	**3**
In welke mate heeft de zelfevaluatie het meest belangrijke doel bereikt?	.71	.28	.20
In welke mate heeft de zelfevaluatie het op één na meest belangrijke doel bereikt?	.61	.23	.24
Kwaliteitsindicator 2: Resulteren in passende actiepunten. De zelfevaluatie resulteerde...			
...in duidelijk omschreven actiepunten.	.32	.74	.31
...in actiepunten waarvan de realisatie achteraf gemakkelijk na te gaan is.	.27	.75	.33
...in actiepunten waarvoor een groot draagvlak bestond.	.25	.66	.44
...in realistische actiepunten.	.26	.74	.40
...in actiepunten met een duidelijke tijdsplanning.	.19	.65	.40
Kwaliteitsindicator 3: Resulteren in concrete resultaten. De zelfevaluatie...			
...bevorderde onze wil om kritischer tegen de zaken aan te kijken.	.23	.32	.70
...resulteerde in extra motivatie om kwaliteit te leveren.	.20	.31	.77
...heeft bijgedragen tot meer samenwerking in onze school.	.21	.31	.80
...heeft ertoe bijgedragen dat wij in onze school in sterkere mate dezelfde doelen zijn gaan nastreven.	.22	.39	.74
...versterkte de samenhang in onze school.	.22	.30	.78
...stimuleerde ons tot het bespreken van onze eigen opvattingen over goed onderwijs.	.24	.34	.68
...versterkte de betrokkenheid van teamleden bij het schoolbeleid.	.19	.34	.77

b. De houding van teamleden ten aanzien van zelfevaluatie

De houding van respondenten ten aanzien van zelfevaluatie werd in kaart gebracht aan de hand van 12 stellingen. Deze werden gepresenteerd in de vorm van twee uitersten, een negatieve variant (score 1) en een positieve variant (score 6). Tabel 19 bevat relevante informatie omtrent de constructie van deze schaal.

Tabel 19: Schaalconstructie 'Houding ten aanzien van zelfevaluatie'

Schaal	Aantal items	N	SEM						Cronbach's alpha	FA
			CMIN	p	AGFI	CFI	RMSEA	pclose		
Houding t.a.v. zelfevaluatie	12	1313	174.0	0.00	0.96	0.99	0.05	0.37	0.92 (0.94)	1

De items in het getoetste meetmodel laden allen hoog op de latente variabele 'houding ten aanzien van zelfevaluatie'. Ook de Cronbach's alpha ligt hoog, met name op 0.92 (0.94). De schaal is daarenboven ééndimensionaal en kan dus met veel vertrouwen aangewend worden binnen de beschrijvende en verklarende analyses.

c. 'Zelfevaluatie als beleidsdaad'

De acht dragers van de 'zelfevaluatie als beleidsdaad' werden in twee van de in deze studie gebruikte gegevensverzamelingen in kaart gebracht. De eerste maal bij de GOK-zelfevaluaties; de tweede maal in het kader van 'zelfevaluatie binnenstebuiten'. De formulering van de items en de aard van de bevraging waren in zekere mate verschillend. De tweede verzameling trok met name lessen uit de eerste. We moeten dan ook een onderscheid maken tussen de schaalconstructies in beide situaties.

- *'Zelfevaluatie als beleidsdaad' in GOK-zelfevaluaties*

De schalen voor de dragers van 'zelfevaluatie als beleidsdaad' werden bij het bestuderen van de GOK-zelfevaluaties ontwikkeld aan de hand van de traditionele schaaltechnieken. We rapporteren in Tabel 20 over de kenmerken van de schalen zoals deze in het oorspronkelijke rapport zijn opgenomen (Van Petegem e.a., 2005b). De analyses leren dat deze schalen met de nodige garanties kunnen aangewend worden in de toekomstige analyses.

Tabel 20: Resultaten schaalconstructie 'Zelfevaluatie als beleidsdaad' – GOK-zelfevaluatie

Schaal	Aantal items	alpha	FA
Doeltreffende communicatie	5	0.82	1
Ondersteunende relaties en samenwerking	6	0.87	1
Gezamenlijke doelgerichtheid	4	0.86	1
Gedeeld leiderschap	5	0.81	1
Responsiviteit	5	0.87	1
Geïntegreerde aanpak	6	0.78	1
Innovatie	7	0.85	1
Reflectie	6	0.90	1

- *'Zelfevaluatie als beleidsdaad' in 'Zelfevaluatie binnenstebuiten'*

Ook in de vragenlijst bij 'zelfevaluatie binnenstebuiten' werd in kaart gebracht in welke mate zelfevaluaties volgens de respondenten aan de verschillende dragers van beleidsvoerend vermogen tegemoet komen. In deze voorbereidende analyses volgen wij naar analogie met de schaalconstructie voor de dragers van beleidsvoerend vermogen twee sporen. Eerst kijken we naar de verschillende dragers afzonderlijk. Daarna bekijken we het concept 'zelfevaluatie als beleidsdaad'.

- *Een schaal voor elke drager*

De schalen die eerder in het onderzoek naar de GOK-zelfevaluaties ontwikkeld werden, worden hier op hun bruikbaarheid getoetst aan de hand van SEM-analyses. De vooropgestelde schalen blijken aan de psychometrische vereisten te voldoen. Zowel de fit-indicatoren bij de SEM-analyses als de waarden van Cronhach's alpha wijzen er op dat de schalen goed tot zeer goed zijn.

Tabel 21: Schaalconstructie dragers 'Zelfevaluatie als beleidsdaad' – 'Zelfevaluatie binnenstebuiten

Dragers BVV	Aantal items	N	SEM						alpha	FA
			CMIN	p	AGFI	CFI	RMSEA	pclose		
Communicatie	5	425	0.9	0.82	1.0	1.0	0.00	0.95	0.89 (0.96)	1
Ondersteuning & samenwerking	5	425	6.4	0.04	0.96	1.0	0.07	0.21	0.91 (0.96)	1
Gedeeld Leiderschap	6	425	3.9	0.27	0.98	1.0	0.03	0.64	0.91 (0.95)	1
Gezamenlijke doelgerichtheid	5	425	7.0	0.07	0.97	1.0	0.05	0.35	0.86 (0.93)	1
Responsief vermogen	5	425	4.1	0.25	0.98	1.0	0.03	0.62	0.79 (0.74)	1
Innovatief vermogen	5	425	6.3	0.10	0.97	0.99	0.05	0.41	0.75 (0.83)	1
Geïntegreerd beleid	5	425	2.5	0.29	0.98	1.0	0.02	0.60	0.84 (0.86)	1
Reflectief vermogen	4	425	0.1	0.75	1.0	1.0	0.00	0.85	0.82 (0.92)	1

- *Het concept 'Zelfevaluatie als beleidsdaad'*

Zonet werd een schaal ontwikkeld voor elke drager. De keuze voor de term dragers impliceert het bestaan van een achterliggend concept of achterliggende concepten die de variatie in de verschillende dragers kunnen verklaren. In de volgende voorbereidende analyse gaan we daarom na of er in de lijst van 40 items sprake is van 8 goed afgelijnde factoren dan wel van één of enkele achterliggende factor(en). We rapporteren daartoe achtereenvolgens over de factor- en SEM-analyses die wij uitvoerden.

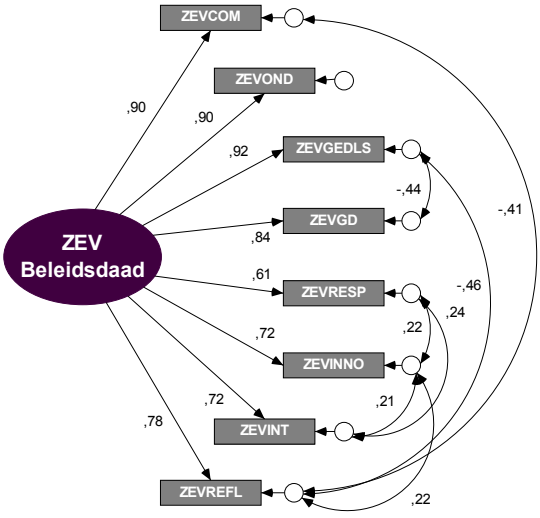
De factoranalyse (FA) gebruikten we in een exploratieve benadering. We wilden nagaan welke factoren onderscheiden kunnen worden in de lijst van 40 items. Een FA zonder restricties komt tot een 6-factoren oplossing (eigenwaardes groter dan 1). Bijna alle items laden evenwel het hoogst op de eerste factor, die 47% van de variantie verklaart. De verschillende dragers zijn – zoals verwacht overigens – niet mooi te onderscheiden op deze manier. Na de eerste factor voegen de volgende factoren verhoudingsgewijs niet veel meer toe aan het percentage verklaarde variantie in de items. De eerste factor interpreteren wij als de mate waarin een school tijdens de zelfevaluatie blijk geeft van beleidsvoerend vermogen. We hanteren hier in de analyses de naam 'zelfevaluatie als beleidsdaad' voor. Om dit concept op een snelle manier te kunnen meten, gaan we ook een verkorte versie van de vragenlijst trachten te hanteren. Daarvoor gebruiken we de 10 items met de hoogste factorladingen op de eerste factor. Dit zijn items met een factorlading van 0.78 of groter op de factor 'zelfevaluatie als beleidsdaad'. Deze worden opgelijst in Tabel 22. Deze schaal heeft een Cronbach's alpha waarde van 0.96 en is ééndimensioneel. De items in deze verkorte vragenlijst zijn afkomstig uit vier van de acht schalen: doeltreffende communicatie, gezamenlijke doelgerichtheid, ondersteunende relaties en samenwerking, en gedeeld leiderschap.

Tabel 22: Factorladingen ingekorte vragenlijst 'Zelfevaluatie als beleidsdaad'

Item	Factorlading
1. Al de nuttige expertise werd gebruikt om de zelfevaluatie tot een goed einde te brengen.	0.83
2. Er was een goede afstemming tussen de leerkrachten en de directie wat hun taken m.b.t. de zelfevaluatie betreft.	0.83
3. Wij werkten als een hecht team samen om de zelfevaluatie tot een goed einde te brengen.	0.82
4. Er werd intern goed over de zelfevaluatie gecommuniceerd.	0.81
5. Er werden tijdens het uitvoeren van de zelfevaluatie goede professionele afspraken gemaakt.	0.80
6. De verantwoordelijken voor de zelfevaluatie vonden volgens mij steun bij collega's.	0.80
7. De draagkracht van ons team was voldoende groot om deze zelfevaluatie uit te voeren.	0.80
8. De zelfevaluatie vond plaats in een sfeer van vertrouwen tussen de directie en leerkrachten.	0.79
9. Er waren voor de zelfevaluatie overzichtelijke overlegprocedures.	0.79
10. Iedereen was goed geïnformeerd over de doelen van de zelfevaluatie.	0.78

We ronden de voorbereidende analyses met betrekking tot de 'zelfevaluatie als beleidsdaad' af door te onderzoeken of en hoeveel variantie het concept in de verschillende dragers verklaart. Daartoe toetsten we het model in Figuur 13 (n= 425; CMIN: 11.6 (p = 0.56); RMSEA: 0.00 (pclose = 0.98); AGFI: 0.99; CFI: 1.0).

Figuur 13: Meetmodel 'Zelfevaluatie als beleidsdaad'



Uit de hoge ladingen blijkt dat een groot deel van de variantie in de verschillende dragers wordt verklaard door de 'zelfevaluatie als beleidsdaad'. De gestandaardiseerde regressiecoëfficiënten variëren van 0.61 tot 0.92. Respondenten die aangeven dat de zelfevaluatie in hun school hoog scoorde op één drager, geven doorgaans ook aan dat deze hoog scoorde op de andere dragers. Dit effect loopt in sterke mate over het concept 'zelfevaluatie als beleidsdaad'. Gemiddeld verklaart 'zelfevaluatie als beleidsdaad' 65% van de varian-

tie in de verschillende dragers. Uitschieters zijn enerzijds de drager 'responsiviteit' met 37% en anderzijds 'gedeeld leiderschap' met 85%.

In de verdere analyses zullen we 3 manieren hanteren om 'zelfevaluatie als beleidsdaad' te meten. In de eerste plaats is er de score op het concept 'zelfevaluatie als beleidsdaad'. Daarnaast zijn er de 8 dragers én de verkorte versie van de vragenlijst.

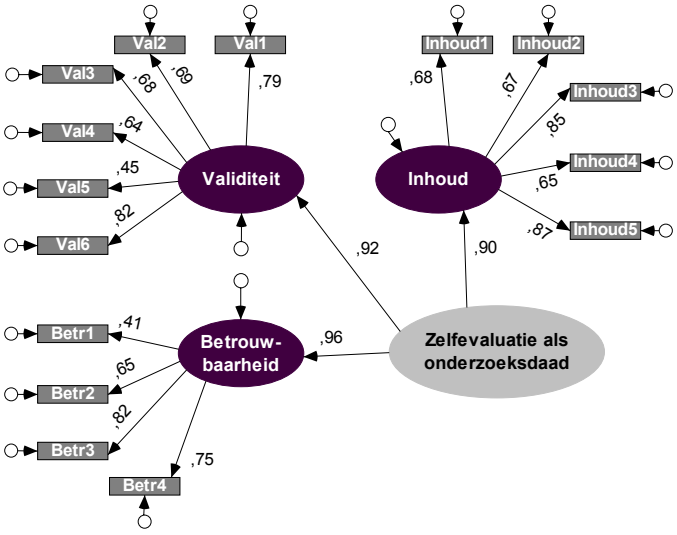
d. 'Zelfevaluatie als onderzoeksdaad'

Op verschillende fronten wordt systematiek in het uitvoeren van een zelfevaluatie als een cruciale factor naar voren geschoven. We benaderen zelfevaluaties om die reden óók als een onderzoeksdaad. Er zijn verschillende stappen in het uitvoeren van een onderzoek die relevant zijn om naar het verloop van zelfevaluaties te kijken. In het theoretisch kader werden de volgende indicatoren onderscheiden: inhoudelijke verankering, relevantie, precisie, betrouwbaarheid, validiteit, efficiëntie, de keuze van een instrument en interpretatie. Deze criteria werden geoperationaliseerd in stellingen die aan de respondenten die voldoende nauw betrokken waren bij de verschillende aspecten van de zelfevaluatie voorgelegd werden. Deze items stellen ons zowel in staat om voor een aantal dragers van 'zelfevaluatie als onderzoeksdaad' als voor het concept als dusdanig tot een passend instrument te komen.

- *Een schaal voor elke drager*

Met de stellingen kan een inhoudelijk relevant meetmodel ontwikkeld worden dat voldoende bij de vastgestelde data aansluit (zie Figuur 14). De bijhorende fit-indicatoren zijn: (n= 134; CMIN: 84.5 (p = 0.15); RMSEA: 0.04 (pclose = 0.76); AGFI: 0.93 ; CFI: 0.99).

Figuur 14: Meetmodel 'Zelfevaluatie als onderzoeksdaad'



De indicatoren worden in dit model ondergebracht in drie schalen. De schaal 'Inhoud' bevat 5 items met betrekking tot de inhoudelijke verankering, relevantie en precisie van de zelfevaluatie (Cronbach's aplha is 0.86 en op schoolniveau 0.76). De schaal 'Validiteit' gaat onder meer na of de verzamelde zelfevaluatiegegevens een goed beeld geven van de bevraagde thema's, of de gegevens correct werden geïnterpreteerd en of de vragenlijst goed aansloot bij de doelen van de zelfevaluatie. Deze schaal telt 6 items en heeft een Cronbach's alpha-waarde van 0.86 (0.83). De vier items van de schaal 'Betrouwbaarheid' richten zich op het zorgvuldig, anoniem en eerlijk uitvoeren van de gegevensverzameling. Cronbach's alpha (0.64 (0.55)) wijst erop dat deze laatste schaal met de nodige omzichtigheid gebruikt moet worden.

- *Het concept 'Zelfevaluatie als onderzoeksdaad'*

Uit Figuur 14 blijkt verder ook dat de drie indicatoren sterk positief laden op het concept 'zelfevaluatie als onderzoeksdaad'. We kunnen de factorscore gebaseerd op de drie dragers van 'zelfevaluatie als onderzoeksdaad' daarom ook als een meting van dit achterliggende concept gebruiken. De schaal met al de 15 items heeft een Cronbach's alpha-waarde van 0.90 (0.85). Ook hier willen we evenwel komen tot een verkorte versie van de vragenlijst. Uit een exploratieve factoranalyse blijkt dat de theoretisch onderscheiden factoren niet afgelijnd teruggevonden worden. Alle items laden het hoogst op de eerste factor die 48% van de variantie verklaart. De eigenwaarde van de andere factoren is niet noemenswaardig groter dan 1. De eerste factor labellen we dan ook als 'zelfevaluatie als onderzoeksdaad'. De verkorte vragenlijst bevat de 6 items (zie Tabel 23) die hoger laden dan .70. Cronbach's alpha bedraagt 0.90 op respondentniveau en 0.82 op schoolniveau.

Tabel 23: Factorladingen ingekorte vragenlijst 'Zelfevaluatie als onderzoeksdaad'

Item	Factorlading
1. Ik had een duidelijk beeld van de bedoelde inhoud van de zelfevaluatie.	.81
2. De vragenlijst sloot goed aan bij de doelen van de zelfevaluatie.	.82
3. De geëvalueerde thema's waren relevant.	.81
4. De gegevensverzameling werd zorgvuldig uitgevoerd.	.77
5. De verzamelde gegevens geven een goed beeld van de bevraagde thema's.	.72
6. De gegevensverzameling leverde representatieve informatie op.	.70

In de verdere analyses zullen we de drie bovenstaande strategieën hanteren om 'zelfevaluatie als onderzoeksdaad' te meten. Het gaat om de afzonderlijke dragers, om het concept 'zelfevaluatie als onderzoeksdaad' en om de ingekorte vragenlijst voor het meten van dit concept.

e. De samenhang tussen de zelfevaluatiegebonden kenmerken

In de eigenlijke analyses worden de drie concepten waarvoor we zonet schalen ontwikkelden (met name 'houding ten aanzien van zelfevaluatie', 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad') als afzonderlijke kenmerken van het zelfevaluatieproces naar voren geschoven. Die keuze is in de eerste plaats ingegeven door

theoretische overwegingen. De vraag is echter of een datagedreven benadering dit onderscheid bevestigt.

Tabel 24: Factorstructuur in de zelfevaluatiegebonden kenmerken

	Factoren		
Houding ten aanzien van zelfevaluatie	**1**	**2**	**3**
Vertelt ons niets nieuws/leert ons heel veel	**.68**	.11	.24
Leidt niet tot beter lesgeven/leidt tot beter lesgeven	**.74**	.10	.25
Kost veel tijd/kost weinig extra tijd	**.53**	-.00	.15
Raakt alleen enkelen/raakt iedereen	**.67**	.05	.16
Leidt niet tot beter management/leidt tot beter	**.74**	.14	.27
Is moeilijk te interpreteren/is makkelijk te begrijpen	**.63**	.21	.26
Is bij de meeste teamleden niet geliefd/is wel geliefd	**.53**	.28	-.01
Is moeilijk uit te voeren/is goed uit te voeren	**.53**	.10	.13
Is niet kosten-effectief/is kosten-effectief	**.55**	.17	.05
Is afhankelijk van het toeval/resulteert in een betrouwbaar beeld	**.72**	.24	.24
Is subjectief/is objectief	**.63**	.24	.12
Is een momentopname/geeft onze ontwikkeling goed weer	**.68**	.36	.22
Zelfevaluatie als beleidsdaad			
Al de nuttige expertise werd gebruikt om de zelfevaluatie tot een goed einde te brengen.	.06	**.62**	.41
Er was een goede afstemming tussen de leerkrachten en de directie wat hun taken m.b.t. de zelfevaluatie betreft.	.04	**.75**	.25
Wij werkten als een hecht team samen om de zelfevaluatie tot een goed einde te brengen.	.14	**.75**	.13
Er werd intern goed over de zelfevaluatie gecommuniceerd.	.21	**.62**	.22
Er werden tijdens het uitvoeren van de zelfevaluatie goede professionele afspraken gemaakt.	.23	**.73**	.16
De verantwoordelijken voor de zelfevaluatie vonden volgens mij steun bij collega's.	.16	**.68**	.09
De draagkracht van ons team was voldoende groot om deze zelfevaluatie uit te voeren.	.12	**.73**	.09
De zelfevaluatie vond plaats in een sfeer van vertrouwen tussen de directie en leerkrachten.	.22	**.65**	.18
Er waren voor de zelfevaluatie overzichtelijke overlegprocedures.	.07	**.66**	.15
Leerkrachten schaarden zich achter de vooropgestelde zelfevaluatiedoelen.	.36	**.54**	.01
Iedereen was goed geïnformeerd over de doelen van de zelfevaluatie.	.26	**.52**	.24
Zelfevaluatie als onderzoeksdaad			
Ik had een duidelijk beeld van de bedoelde inhoud van de zelfevaluatie.	.33	.24	**.77**
De vragenlijst sloot goed aan bij de doelen van de zelfevaluatie.	.17	.22	**.78**
De geëvalueerde thema's waren relevant.	.20	.26	**.77**
De gegevensverzameling werd zorgvuldig uitgevoerd.	.24	.13	**.72**
De verzamelde gegevens geven een goed beeld van de bevraagde thema's.	.31	.22	**.65**
De gegevensverzameling leverde representatieve informatie op.	.26	.26	**.55**

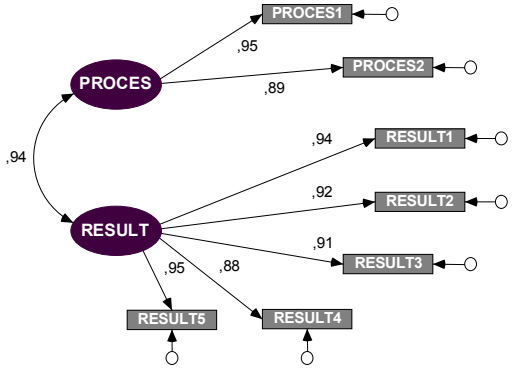
We hebben daartoe een exploratieve factoranalyse (met Varimax rotatie) uitgevoerd op de items van de ingekorte vragenlijsten voor 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad' en de items die de houding ten aanzien van zelfevaluatie in kaart brengen. De 3-factorenoplossing in Tabel 24 verklaart 53% van de variantie in de

items. Uit deze factorstructuur blijkt dat de data onze theoretische veronderstelling omtrent drie te onderscheiden zelfevaluatiegebonden kenmerken ondersteunen[9].

3.1.3 Omtrent de ondersteuning

Om tot een beoordeling te komen van de ondersteuning die scholen tijdens het zelfevaluatieproces genoten, werden in het theoretisch kader twee indicatoren vooropgesteld. Het gaat om de mate waarin de ondersteuning die scholen genoten inspeelde op de reële behoeftes van scholen enerzijds en de mate waarin de ondersteuning een aantal doelstellingen bereikte anderzijds. Er is dus een schaal die zich richt op het proces en één die focust op het product van de zelfevaluatie (zie Figuur 15).

Figuur 15: Meetmodel 'Beoordeling van de ondersteuning'



Het model zoals afgebeeld in Figuur 15 sluit aan bij de empirische vaststellingen (n= 1269; CMIN: 9.4 (p = 0.09); RMSEA: 0.03 (pclose = 0.93); AGFI: 0.99 ; CFI: 1.0). We aanzien beide schalen als een indicatie voor de kwaliteit van de ondersteuning. Cronbach's aplha bedraagt 0.91 (0.96) voor 'Proces van de begeleiding' en 0.96 (0.98) voor 'Resultaat van de begeleiding'. Uit de figuur blijkt ook dat de correlatie tussen beide schalen zeer sterk is. Omdat beide conceptueel verschillende zaken meten maken we toch het onderscheid.

[9] Ook hier wijst de niet geroteerde factormatrix in eerste instantie op 1 factor waarop alle items behoorlijk sterk laden met daarnaast bijkomend twee relevante maar moeilijk te interpreteren factoren. Er is dus ook evidentie om te pleiten voor het gebruik van één achterliggend concept. Deze aanpak maakt echter dat een noemenswaardig deel van de variantie in de items niet mee in de analyses opgenomen wordt. Bovendien hebben we goede theoretische gronden om aan het onderscheid tussen de drie zelfevaluatiegebonden kenmerken vast te houden.

3.2 Clusteranalyses

Scholen en respondenten zullen in het verdere verloop van deze studie herhaaldelijk in groepen ondergebracht worden. Dat doen we in de eerste plaats omdat het vaststellen en presenteren van effecten van types scholen en types respondenten op het verloop en de kwaliteit van zelfevaluaties een bevattelijke manier van werken is. Bovendien anticiperen we op problemen met multicollineariteit in de uit te voeren regressieanalyses. Het onderscheid tussen clusters van scholen en clusters van respondenten is ingegeven door de vaststelling dat respondenten binnen eenzelfde school sterk van mening kunnen verschillen. De gemiddelde schoolscore is dus niet noodzakelijk een goede voorspeller van de individuele score van de respondenten in de school. We beschrijven eerst de clusters op schoolniveau, daarna deze op respondentenniveau.

3.2.1 Clusters van scholen

Scholen kunnen op basis van verschillende kenmerken in clusters onderverdeeld worden. Achtereenvolgens presenteren we de resultaten van clusteranalyses naar aanleiding van het beleidsvoerend vermogen, van de effectiviteitsperspectieven, van de 'zelfevaluatie als beleidsdaad', van de 'zelfevaluatie als onderzoeksdaad' en van de score van scholen op de criteria van de GOK-inspectie.

a. Clusters van scholen op basis van beleidsvoerend vermogen (BVV)

Op basis van de gegevens die we verzameld hebben over de GOK-scholen kunnen verschillende wegen bewandeld worden om scholen in clusters onder te brengen. Wij hebben er voor gekozen om een clusteranalyse uit te voeren op basis van de verschillende dragers van beleidsvoerend vermogen. Omdat de dataset 'slechts' 396 cases met volledige data bevat, kozen wij voor hiërarchische clustering op basis van de gekwadrateerde Euclidische afstand (Ward's Method). De 3 clusters-oplossing bleek het meest eenvoudig interpreteerbaar en voor verdere analyse het meest bruikbaar.

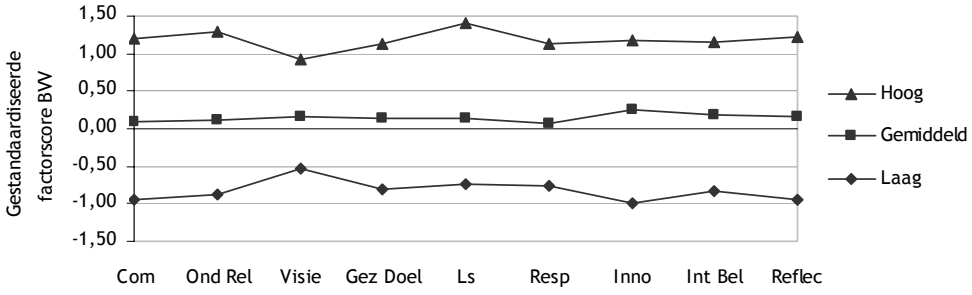
Tabel 25: Verdeling van scholen over de clusters van beleidsvoerend vermogen

3 Clusters	Aantal scholen	%	Dragers van beleidsvoerend vermogen (schaalscores)								
			COM	OND REL	VISIE	GEZ DOEL	LS	RESP	INNO	INT BEL	REFLEC
Hoog	112	28.3%	4.68	4.46	4.47	4.39	4.58	4.11	4.34	4.35	4.11
Gemiddeld	222	56.1%	4.07	3.85	4.04	3.76	4.00	3.58	3.70	3.81	3.43
Laag	62	15.6%	3.52	3.31	3.61	3.20	3.60	3.07	2.86	3.20	2.70
Totaal	396	100%	4.01	3.80	3.99	3.70	3.98	3.52	3.56	3.72	3.34

Tabel 25 geeft onder meer de verdeling van scholen over de clusters weer. De drie clusters blijken niet even groot. Cluster 2 omvat iets meer dan de helft van de bevraagde GOK-scholen (56%). Uit Figuur 16 blijkt dat we deze cluster zouden kunnen omschrijven als 'de middelste groep'. De andere twee clusters nemen op een consistente manier tel-

kens een van het gemiddelde afwijkende positie in. De scholen in cluster 1 scoren telkens het laagst, deze in cluster 3 het hoogst. Post hoc analyses (Bonferroni) bij het uitvoeren van variantie-analyses leren dat de verschillen tussen de clusters allemaal statistisch significant zijn. Op basis van deze bevindingen zullen we in het vervolg spreken van scholen met een sterk, gemiddeld en zwak beleidsvoerend vermogen.

Figuur 16: Gestandaardiseerde factorscores 'dragers van BVV' voor de scholenclusters



b. Clusters van scholen op basis van de effectiviteitsperspectieven

Een hiërarchische clustering voor de 96 scholen waarover binnen 'zelfevaluatie binnenstebuiten' gegevens werden verzameld resulteerde in een 3-clusteroplossing. Het onderscheid tussen deze drie clusters is gebaseerd op de gemiddelde standaardscores van alle respondenten in de school op de vier effectiviteitsperspectieven.

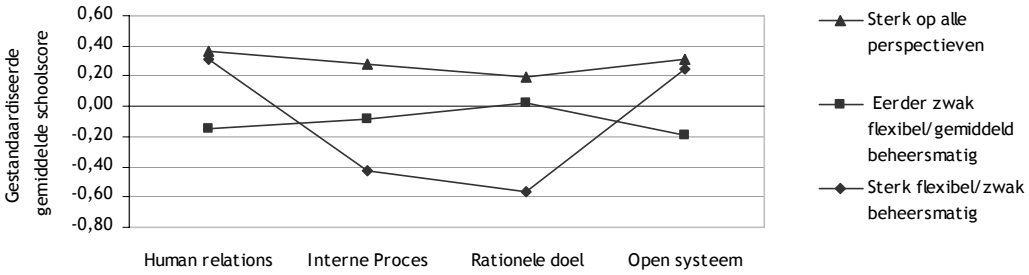
Tabel 26: Verdeling van scholen over de clusters o.b.v. 'effectiviteitsperspectieven'

3 Clusters	Aantal scholen	%	Effectiviteitsperspectieven (schaalscores)			
			HR	IP	RD	OS
Cluster 1: Sterk flexibel/zwak beheersmatig	8	8.3%	4.34	3.69	3.43	3.88
Cluster 2: Eerder zwak flexibel/ gemiddeld beheersmatig	49	51.1%	4.04	3.91	3.82	3.57
Cluster 3: Sterk op alle perspectieven	39	40.6%	4.42	4.11	3.92	3.95
Totaal	**96**	**100%**	**4.22**	**3.97**	**3.83**	**3.75**

De drie clusters die we onderscheiden zijn 'Sterk op alle perspectieven', 'Eerder zwak flexibel/gemiddeld beheersmatig' en 'Sterk flexibel/zwak beheersmatig'. Deze groepen zijn respectievelijk goed voor 40.6%, 51% en 8.3% van het totale aantal scholen. Figuur 17 visualiseert de gestandaardiseerde resultaten. Bij de interpretatie van de gestandaardiseerde scores mag men niet uit het oog verliezen dat het hier om een normgerichte benadering gaat. Daarom hebben we in Tabel 26 ook de schaalscores voor de verschillende clusters opgenomen. Dat geeft een idee van de absolute scores op de vier perspectieven.

Al de gevonden verschillen in de schaalscores van de drie clusters zijn statistisch significant.

Figuur 17: Scores van de 3 scholenclusters op de effectiviteitsperspectieven



c. Clusters van scholen op basis van 'zelfevaluatie als beleidsdaad'

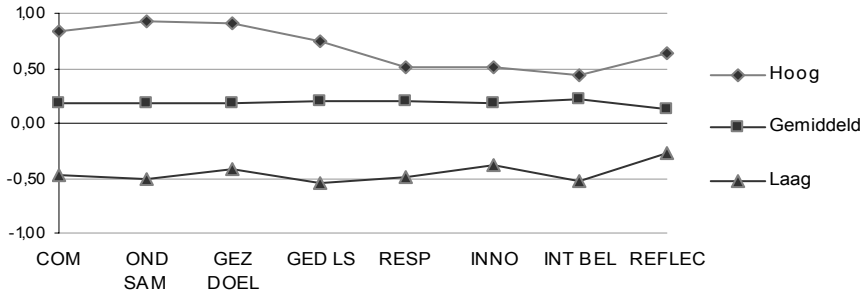
In het theoretisch kader werden acht dragers van 'zelfevaluatie als beleidsdaad' onderscheiden. Op basis van deze dragers kunnen de bevraagde scholen onderverdeeld worden in 3 clusters. Het gaat hierbij om scholen die in relatieve betekenis hoog, gemiddeld en laag scoren op de verschillende dragers. Iets meer dan de helft (55.3%) van de scholen behoort tot de groep 'gemiddeld'. De groep die verhoudingsgewijs laag scoort op de dragers van zelfevaluatie als beleidsdaad omvat bijna één school op drie. De groep scholen met een hoge score op elk van de dragers omvat 14.9% van al de scholen. De schaalscores in Tabel 27 geven aan dat er sprake is van noemenswaardige verschillen. Dat vertaalt zich duidelijk in de resultaten van een variantieanalyses. De verschillen tussen de drie clusters zijn sterk statistisch significant.

Tabel 27: Verdeling van scholen over clusters 'zelfevaluatie als beleidsdaad'

3 Clusters	Aantal scholen	%	Dragers van zelfevaluatie als beleidsdaad (schaalscores)							
			COM	OND REL	GEZ DOEL	GED LS	RESP	INNO	INT BEL	REFLEC
Hoog	14	14.9%	3.95	4.08	3.94	4.01	3.17	3.52	3.75	3.88
Gemiddeld	52	55.3%	3.38	3.48	3.38	3.50	2.84	3.30	3.58	3.51
Laag	28	29.8%	2.88	2.90	2.96	2.95	2.38	2.98	3.17	3.26
Totaal	94	100%	3.32	3.39	3.34	3.41	2.75	3.24	3.48	3.49

De verschillen tussen de clusters worden gevisualiseerd in Figuur 18. Anders dan de bovenstaande tabel is deze figuur gebaseerd op de gestandaardiseerde scores van de respondenten in de verschillende scholen. Uit de figuur blijkt dat de clusters consequent verschillend zijn over de 8 dragers van zelfevaluatie als beleidsdaad. De cluster die 'laag' scoort, situeert zich systematisch (minstens) 1 standaarddeviatie lager dan de cluster die 'hoog' scoort. De grootte van deze verschillen maakt het interessant om het effect van het clusterlidmaatschap op de kwaliteit van zelfevaluaties te bestuderen.

Figuur 18: Gemiddelde scores (gestandaardiseerd) van de 3 scholenclusters op de dragers van 'zelfevaluatie als beleidsdaad'



d. Clusters van scholen op basis van 'zelfevaluatie als onderzoeksdaad'

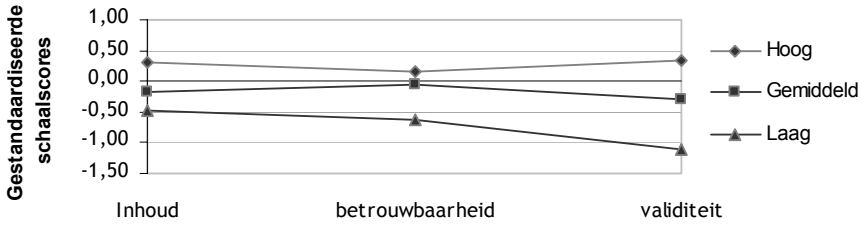
Ook op basis van de drie dragers van 'zelfevaluatie als onderzoeksdaad' voerden we een clusteranalyse uit. Dit resulteerde ook hier in een oplossing met drie groepen die we opnieuw laag, gemiddeld en hoog labellen. Tabel 28 beschrijft de verdeling van de scholen over de drie indicatoren.

Tabel 28: Verdeling van scholen over clusters 'Zelfevaluatie als onderzoeksdaad'

3 clusters	Aantal scholen	%	Inhoud	Betrouwbaarheid	Validiteit
Hoog	51	53.6%	3.96	4.24	4.08
Gemiddeld	39	41.1%	3.59	4.09	3.66
Laag	5	5.3%	3.36	3.73	3.12
Totaal	95	100%	3.78	4.15	3.86

In Figuur 19 visualiseren we de verschillen tussen scholen aan de hand van de gestandaardiseerde schaalscores. Scholen uit de cluster 'laag' scoren minder goed op de drie dragers van 'zelfevaluatie als beleidsdaad' dan de scholen die 'gemiddeld' of 'hoog' scoren. Dat geldt voornamelijk voor de validiteit van de zelfevaluatie.

Figuur 19: Gemiddelde scores van de 3 scholenclusters op de dragers van 'Zelfevaluatie als onderzoeksdaad'



e. Clusters van scholen op basis van de criteria van de GOK-inspectie

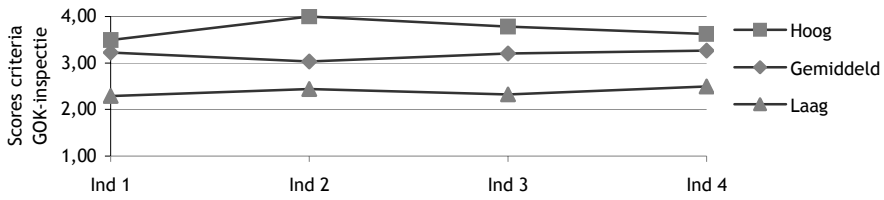
We schuiven in de analyses verschillende indicatoren voor de kwaliteit van zelfevaluaties naar voren. De aanpak van de onderwijsinspectie levert ook één zulke indicator op. We willen de verschillende criteria die in deze benadering gebruikt werden integreren. Daartoe werd ook hier een clusteranalyse uitgevoerd.

Tabel 29: Verdeling van scholen over clusters 'GOK-inspectie'

	Steekproef	
Cluster	Aantal scholen	%
Laag	184	34.9%
Gemiddeld	181	34.4%
Hoog	162	30.7%
Totaal	**527**	**100%**

De clusteranalyse verdeelt scholen in grosso modo drie gelijke groepen. Om betekenis te geven aan de drie clusters geeft Figuur 20 voor elke cluster de gemiddelde scores weer voor de vier indicatoren die de onderwijsinspectie hanteerde bij het beoordelen van de kwaliteit van de GOK-zelfevaluaties: resultaten van de zelfevaluatie (Ind 1); gegevensverzameling (Ind 2), gegevensanalyse (ind 3) en betrokkenheid bij de zelfevaluatie (Ind 4). De cluster 'Gemiddeld' neemt telkens de middenpositie in, cluster 'Hoog' scoort het best en cluster 'Laag' het minst goed. Post hoc analyses bij het uitvoeren van variatieanalyses leren ook hier dat de verschillen tussen de clusters allemaal statistisch significant zijn. In het vervolg van deze studie zullen wij dan ook spreken van een gemiddelde, sterke en zwakke inspectiebeoordeling.

Figuur 20: Gemiddelde scores van de scholenclusters op de indicatoren voor de GOK-inspectie



3.2.2 Clusters van respondenten

Ook respondenten kunnen op een verhelderende manier in clusters ondergebracht worden. Hoewel de schoolscore gebaseerd is op het gemiddelde van al de individuele beoordelingen in de school, zouden wij door enkel gebruik te maken van schoolscores voorbij gaan aan de verschillen die er tussen teamleden kunnen zijn. Respondenten die samen in een cluster zitten hebben een gelijkaardig oordeel over de bestudeerde elementen. Dat

betekent echter niet noodzakelijk dat hun scholen een gelijkaardige schoolscore hebben. Om die reden willen we ook komen tot clusters van respondenten voor de effectiviteitsperspectieven enerzijds en voor de dragers van 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad'.

a. Clusters van respondenten op basis van de effectiviteitsperspectieven

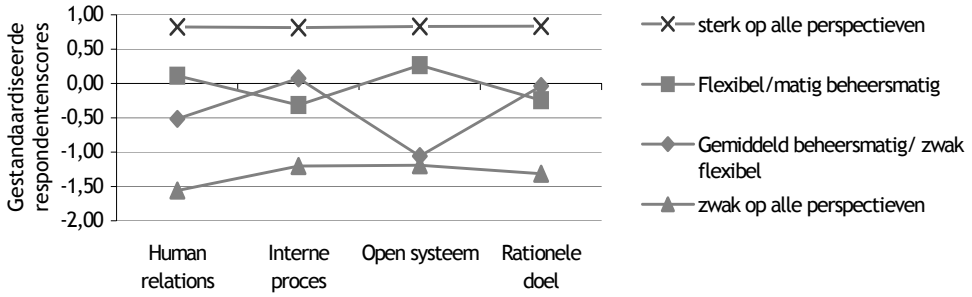
De respondenten hebben zich in de vragenlijst uitgesproken over hoe belangrijk anderen in hun school een aantal waarden vinden. Deze waarden gaan terug op vier effectiviteitsperspectieven. Op basis van de score van individuen op deze schalen werd een clusteranalyses uitgevoerd (K-means Cluster Analysis). Een vierclusteroplossing (Tabel 26) bleek het meest aangewezen in ons streven naar variatie enerzijds en interpreteerbaarheid anderzijds.

Tabel 30: Verdeling van respondenten over clusters 'effectiviteitsperspectieven'

4 Clusters	Aantal respondenten	%	Effectiviteitsperspectieven			
			HR	IP	RD	OS
Sterk op alle perspectieven	745	29.1%	4.75	4.43	4.32	4.33
Flexibel/matig beheersmatig	1012	39.6%	4.26	3.77	3.67	3.93
Gemiddeld beheersmatig/ zwak flexibel	579	22.6%	3.83	4.00	3.82	2.97
Zwak op alle perspectieven	222	8.7%	2.91	3.16	2.91	2.77
Totaal	2558	100%	4.18	3.96	3.83	3.72

De twee uiterste clusters zijn eenvoudig te interpreteren. 29.1% van de respondenten behoort tot de cluster die sterk scoort op alle perspectieven; 8.7% van de respondenten behoort tot de cluster 'zwak op alle perspectieven'. De twee andere clusters scoren tussen beide in. 39.6% scoort bovengemiddeld flexibel maar slechts matig beheersmatig (dit is de cluster 'Flexibel/matig beheersmatig'). De cluster met respondenten die aangeven dat hun school gemiddeld beheersmatig is en zwak scoort inzake flexibiliteit bevat 22.6% van de respondenten.

Figuur 21: Scores van de respondentenclusters op de effectiviteitsperspectieven



Figuur 21 visualiseert de gestandaardiseerde resultaten. Post hoc analyses bij het uitvoeren van variatieanalyses leren dat de verschillen tussen de clusters allemaal statistisch significant zijn.

b. Clusters van respondenten op basis van de dragers van 'zelfevaluatie als beleidsdaad'

Wat 'zelfevaluatie als beleidsdaad' betreft, kunnen respondenten best onderverdeeld worden in drie clusters. Deze clusters interpreteren we als 'hoog', 'gemiddeld' en 'laag'. Ze hebben betrekking op de beoordeling door de respondenten van de mate waarin de zelfevaluatie in hun school tegemoet kwam aan de dragers van de zelfevaluatie als beleidsdaad. Tabel 31 geeft de verdeling van de respondenten over de clusters en hun gemiddelde schaalscores op de 8 dragers weer.

Tabel 31: Verdeling van respondenten over clusters 'zelfevaluatie als beleidsdaad'

3 Clusters	Aantal Respondenten	%	Dragers van zelfevaluatie als beleidsdaad (schaalscores)							
			COM	OND REL	GEZ DOEL	GED LS	RESP	INNO	INT BEL	REFLEC
Hoog	751	39.5%	4.05	4.07	3.89	4.11	3.40	3.63	3.98	3.92
Gemiddeld	805	42.3%	3.17	3.23	3.17	3.30	2.54	3.11	3.40	3.33
Laag	347	18.2%	2.15	2.13	2.30	2.07	1.76	2.33	2.52	2.64
Totaal	1903	100%	3.33	3.36	3.29	3.40	2.74	3.17	3.47	3.44

We ronden de beschrijving van deze respondentenclusters af met een tabel die de herkomst van respondenten weergeeft in de drie clusters die we ten aanzien van de dragers van zelfevaluatie als beleidsdaad op schoolniveau onderscheidden (zie Tabel 32). Daaruit blijkt dat de cluster waartoe een school behoort geen uitsluitsel geeft over de cluster waartoe de respondenten behoren. Dat wijst er op dat er behoorlijk wat variatie is binnen scholen. We doen er dus inderdaad goed aan in de analyses een onderscheid te maken tussen een gemiddelde schoolscore en de individuele beoordeling van de verschillende teamleden in een school.

Tabel 32: Variatie binnen schoolclusters 'zelfevaluatie als beleidsdaad'

	Clusters respondenten			
Clusters scholen	Hoog	Gemiddeld	Laag	Totaal
Hoog	**74.7%**	23.6%	1.7%	100%
Gemiddeld	45.6%	**43.5%**	10.9%	100%
Laag	18.1%	45.9%	**36.0%**	100%

c. Clusters van respondenten op basis van de dragers van 'zelfevaluatie als onderzoeksdaad'

Tenslotte werden de respondenten ook in clusters ondergebracht op basis van hun perceptie van het onderzoeksmatige verloop van de zelfevaluatie. Ook hier blijkt een 3-

clusteroplossing aangewezen. Opnieuw is het onderscheid te interpreteren in termen van een hoge, gemiddelde en lage score op de dragers. 36.6% van de respondenten behoort tot de cluster 'hoog', 16.8% tot de cluster 'laag'. Tabel 33 bevat de gemiddelde resultaten van de drie respondentenclusters voor de drie dragers van 'zelfevaluatie als beleidsdaad'. De verschillen zijn statistisch significant, zo blijkt uit een variantieanalyse.

Tabel 33: Verdeling van respondenten over clusters 'zelfevaluatie als onderzoeksdaad'

3 clusters	Aantal respondenten	%	Inhoud	Betrouw-baarheid	Validiteit
Hoog	324	36.6%	4.41	4.71	4.41
Gemiddeld	412	46.6%	3.83	4.06	3.76
Laag	149	16.8%	2.96	3.66	2.97
Totaal	885	100%	3.89	4.23	3.87

4 AFRONDEND

Dit vijfde deel van het onderzoeksrapport rapporteerde over de gegevensverzameling en de voorbereidingen van de eigenlijke analyse van de kwantitatieve data. In de volgende delen rapporteren we over deze analyses. Daarbij zal veelvuldig verwezen worden naar de schalen en de clusters van scholen en respondenten die hierboven ontwikkeld werden.

DEEL 6

ZELFEVALUATIES IN DE BESTUDEERDE SCHOLEN – EEN BESCHRIJVING

Alvorens ons te richten op de verbanden tussen de concepten uit het theoretisch kader geven we een impressie van hoe zelfevaluaties in de bestudeerde scholen verlopen. We willen hiermee een kernachtige beschrijving geven van de variatie in de zelfevaluatieprocessen van Vlaamse scholen. Het gaat dan ook om de kernkenmerken van zelfevaluaties zoals we deze in het theoretisch kader presenteerden. Wie neemt het initiatief om de zelfevaluatie uit te voeren? Welke functies stelt men voorop? Welke instrumenten gebruikt men? Wie wordt er betrokken tijdens het proces? Naar welke ondersteuning gaat men op zoek? We baseren ons bij het beantwoorden van deze vragen op de gegevensverzamelingen die in het kader van 'zelfevaluatie binnenstebuiten' uitgevoerd werden. Het gaat om de inleidende vragenlijst die voorgelegd werd aan ruim 100 leiders van scholen die recent een zelfevaluatie uitvoerden. Daarnaast zijn er de resultaten van de bevraging van 2716 respondenten in 96 van deze scholen. De beschrijving richt zich dus soms op scholen, dan weer op respondenten. We benadrukken dat het hier louter gaat om de door ons bevraagde scholen, en dat we geen garanties hebben omtrent de representativiteit van het beeld dat opgehangen wordt.

1 INITIATIEF, FUNCTIES EN VERWACHTINGEN

De definitie van zelfevaluatie gaat uit van een proces dat door de school zelf geïnitieerd wordt. Dat blijkt voor de meeste zelfevaluaties inderdaad het geval te zijn. 94% van de schoolleiders geeft in de inleidende vragenlijst aan dat het in de eerste plaats een initiatief was van de school zelf. Telkens 2% geeft aan dat de zelfevaluatie er in eerste instantie kwam naar aanleiding van een vraag vanwege de scholengemeenschap, de overheid of het schoolbestuur/Inrichtende Macht. Vermits we enkel in kaart brachten wie de voornaamste initiatiefnemer was, is het evenwel best mogelijk dat er voor één zelfevaluatie verschillende aanleidingen zijn.

Ongeacht de oorsprong van het initiatief kan een zelfevaluatie om verschillende redenen uitgevoerd worden. Tabel 34 geeft aan dat scholen volgens hun schoolleider voornamelijk

een spiegelfunctie en/of een beleidsvoorbereidende functie meegeven aan de zelfevaluatie. In 87% van de scholen ligt de klemtoon op interne procesontwikkeling (cf. de spiegelfunctie). Dat betekent dat men aan de hand van feedback over de eigen situatie zicht probeert te krijgen op het eigen functioneren en dat van collega's. In 76% van de deelnemende scholen werd aan de zelfevaluatie een beleidsvoorbereidende functie toegedicht. In die scholen zijn de resultaten van de zelfevaluatie bedoeld om bij te dragen tot de totstandkoming van nieuw beleid en tot de verbetering van het bestaande beleid. Zelfevaluaties hebben verder in heel wat scholen een legitimatiefunctie. In een kwart van de scholen ligt de klemtoon op verantwoording naar buiten toe. Aan de hand van het uitvoeren van een zelfevaluatie legt men in deze scholen naar externen verantwoording af voor de activiteiten die men onderneemt. Ook combinaties tussen de spiegel-, legitimatie- en beleidsvoorbereidende functie komen voor. Verder blijkt ook dat de transferfunctie slechts sporadisch voorkomt. Slechts 2% van de scholen geeft immers aan dat de zelfevaluatie opgezet werd om te leren van anderen.

Tabel 34: Functies van de zelfevaluatie

Functies	% scholen
Een spiegelfunctie	87%
Een legitimatiefunctie	24%
Een transferfunctie	2%
Een beleidsvoorbereidende functie	76%

In het theoretisch kader omschreven wij kwaliteit als het voldoen aan verwachtingen. In de schooleigen benadering van kwaliteit wordt daarom eerst in kaart gebracht welke de verwachtingen zijn die de respondenten koesteren ten aanzien van zelfevaluatie. De resultaten in Tabel 35 geven aan hoe vaak een aantal specifieke verwachtingen door de respondenten uit een lijstje als het meest belangrijk of als het op één na meest belangrijk geselecteerd werd. Bijna een kwart van de totale groep respondenten stelt 'komen tot schoolgerelateerde verbeterpunten' als belangrijkste doel voorop. 'Zicht krijgen op sterke en zwakke punten in het handelen van leerkrachten' is voor 1 op 5 respondenten een belangrijke verwachting. 'Het opsporen en delen van goede voorbeelden' enerzijds en 'het oplossen van concrete problemen' blijken het minst belangrijk in deze lijst.

Hoewel de resultaten grotendeels gelijklopen blijkt uit een chikwadraattoets dat er toch verschillen zijn tussen het basisonderwijs en het secundair onderwijs (Pearson X^2 (9, 2360) = 44.1, $p < 0.001$). Het meest opvallend is dat respondenten uit het secundair onderwijs - in vergelijking met het basisonderwijs - dubbel zo vaak aangeven de zelfevaluatie uit te voeren om in orde zijn met de voorschriften van de overheid (8% vs. 4.3%). Er zijn verder verschillen tussen respondenten die deel uitmaken van de schoolleiding (SL) en andere teamleden (TL) in de school (Pearson X^2 (9, 2360) = 76.1, $p < 0.001$). Andere teamleden (leraars en ondersteunend personeel) geven vaker dan schoolleiders aan dat in orde zijn met de overheidsregulatie en voorbereid zijn op een doorlichting een doel was van de zelfevaluatie. Schoolleiders verwijzen duidelijk meer naar verwachtingen zoals

'bijdragen tot een grotere betrokkenheid van leerkrachten op het schoolbeleid' en 'komen tot schoolgerelateerde verbeterpunten'.

Tabel 35: Belang van verwachtingen bij de start van een zelfevaluatie

		Onderwijs-niveau		Functie	
	Totaal %	BaO %	SO %	SL %	TL %
Komen tot schoolgerelateerde verbeterpunten.	23	26	21	34	22
Zicht krijgen op sterke en zwakke punten in het handelen van leerkrachten.	19	20	19	14	20
Bijdragen tot een prettiger werkklimaat in de school.	12	12	12	10	12
Zicht krijgen op sterke en zwakke punten in het handelen van de schoolleiding.	11	10	11	12	10
Bijdragen tot een grotere betrokkenheid van leerkrachten op het schoolbeleid.	9	9	9	16	9
Voorbereid zijn op een controle door de inspectie.	9	8	9	4	9
In orde zijn met de voorschriften van de overheid.	7	4	8	3	7
Komen tot klasgerelateerde verbeterpunten.	6	6	6	5	6
Een concreet probleem oplossen.	3	3	3	1	3
Het opsporen en delen van goede voorbeelden.	2	2	3	1	2
Totaal	100	100	100	100	100

2 INHOUD EN INSTRUMENTKEUZE

In de inleidende vragenlijst werd schoolleiders gevraagd welke thema's er in de vragenlijst die zij gebruikten tijdens de zelfevaluatie aan bod kwamen. In eerste instantie konden zij aangeven of de thema's in Tabel 36 al dan niet tot de focus behoorden.

Tabel 36: De inhoud van zelfevaluaties

	Totaal
De onderwijskundige aanpak in de school	89%
Het functioneren als team	86%
Het leiderschap in de school	83%
Het leren van leerlingen	73%
De pedagogisch/levensbeschouwelijke aanpak in de school	72%
De relaties met ouders/externen	59%
Het nascholingsbeleid	51%
De infrastructuur/accommodatie	51%

De meest bevraagde thema's blijken de onderwijskundige aanpak, het functioneren van het team en het leiderschap in de school. Deze thema's maken in meer dan vier op de vijf scholen deel uit van de inhoud van de zelfevaluatie. Ook het leren van leerlingen en de pedagogisch/levensbeschouwelijke aanpak in de school zijn veelbevraagde thema's. Zij maken in ruim 70% van de scholen deel uit van de zelfevaluatie. De relaties met ou-

ders en externen worden in 60% bestudeerd. Het nascholingsbeleid en de infrastructuur/accommodatie in de helft van de bevraagde scholen.

Schoolleiders die in het bovenstaande lijstje (één van) de belangrijkste inhouden van hun zelfevaluatie niet terug konden vinden, konden deze alsnog noteren. Dat leverde deels meer specifieke invullingen op van het bovenstaande (zoals communicatie op school, de GOK-werking, evaluerend vermogen van de school, het welbevinden van leerkrachten en leerlingen) maar ook een verruiming (zoals de schoolvisie en de doelen van de school).

De inhoud van zelfevaluaties is doorgaans gerelateerd aan het instrument dat scholen gebruiken voor de gegevensverzameling. Vlaamse scholen gebruiken een variëteit aan instrumenten. De door ons bevraagde schoolleiders geven aan voornamelijk gebruik te hebben gemaakt van IZES (met name 52%). Verhoudingsgewijs is het gebruik van DISO, BIOS, SION en PROZA dus beperkt in de onderzochte scholen. Een kwart van de scholen werkte met een zelf ontwikkelde vragenlijst. Het gaat dan bijvoorbeeld om instrumenten ontwikkeld door de schoolleiding, de GOK-leerkracht of de ICT-coördinator.

Tabel 37: De gebruikte zelfevaluatie-instrumenten

	Totaal
DISO – Diagnose Instrument Secundair Onderwijs	6%
PROZA – Projectgroep Ontwikkeling Zelfanalyse-instrument	1%
SION – School in Ontwikkeling	2%
IZES – Instrument voor Zelfevaluatie	52%
BIOS – Buitengewoon Onderwijs Instrument voor de Ontwikkeling van scholen	5%
Andere instrumenten	15%
Een zelfontwikkeld instrument	24%

De groep 'andere instrumenten' omvat onder meer CAF (het Common Assessment Framework), instrumenten afkomstig van de pedagogische begeleiding (bijvoorbeeld de effectenmeter, de Interne Schoolanalyse (ISA)) of de scholengemeenschap, maar ook instrumenten uit het bedrijfsleven, al dan niet aangepast voor onderwijsorganisaties (de tevredenheidsmeter, SWOT-analyse). Een beperkt aantal scholen gebruikte een combinatie van twee instrumenten.

In de meeste scholen – met name 68% - bestond de gegevensverzameling in het kader van de zelfevaluatie enkel uit het afnemen van een vragenlijst. In de andere scholen werd de schriftelijke bevraging aangevuld met focusgroep- of verhelderingsgesprekken (15%), een analyse van beschikbare cijfergegevens (13,5%) of individuele interviews met betrokkenen (9,6%). In de bestudeerde scholen is er dus doorgaans geen sprake van een kwalitatief luik bij de gegevensverzameling. Het grote gewicht van het gebruik van schriftelijke vragenlijsten is uiteraard een gevolg van de selectie van scholen. Gewerkt hebben met zulk een vragenlijst was immers een voorwaarde om aan deze studie te kunnen deelnemen.

Het gebruik van schriftelijke vragenlijsten bij het uitvoeren van een zelfevaluatie is voor veel scholen een relatief nieuw fenomeen. Van de bevraagde scholen geeft immers 62,5%

aan dat de bestudeerde zelfevaluatie de eerste keer was dat zij zich in een dergelijk zelfevaluatieproces engageerden. Voor ruim 20% ging om een tweede keer. Daarnaast is ook een noemenswaardige groep scholen die reeds 2 of meer keer een zelfevaluatie uitvoerde aan de hand van een schriftelijke vragenlijst. Zij zijn goed voor 16% van de scholen in de inleidende vragenlijst. Deze verhoudingen zijn hier beschreven voor de totale groep scholen maar blijken zowel te gelden voor de basisscholen als voor secundaire scholen.

3 Betrokkenheid bij de zelfevaluatie

De wie-vraag stelt zich tijdens verschillende fasen van het zelfevaluatieproces en kan doorheen het proces verschillende antwoorden krijgen. Dat blijkt tenminste uit de bevraging van betrokkenheid bij de zelfevaluatie. Zowel de inleidende vragenlijst als de eigenlijke bevraging besteedden hier aandacht aan.

Uit de inleidende bevraging (zie Tabel 38) blijkt dat de directeur in quasi alle scholen betrokken werd bij de voorbereiding van de zelfevaluatie. Die voorbereiding is verder ook een aangelegenheid van het middenkader[10] en de leerkrachten. In 21% van de scholen speelt ook de pedagogische begeleidingdienst een rol in de voorbereidende activiteiten. Leerlingen, ouders, het schoolbestuur/de inrichtende macht en het CLB worden zelden betrokken in de voorbereiding. De gegevensverzameling richt zich in sterke mate op de leerkrachten. In 93% van de zelfevaluaties wordt via hen gegevens verzameld. Ook de directeur (75%) en het middenkader (46%) worden regelmatig betrokken tijdens de gegevensverzameling. In één op de drie scholen bevraagt men ook het administratief personeel. Bij respectievelijk 14% en 11% van de zelfevaluaties die wij bestudeerden werden (ook) de ouders en leerlingen bevraagd. Het bespreken van de resultaten is voornamelijk een activiteit van de directeur en leerkrachten (respectievelijk 97% en 93%). Het middenkader wordt in de helft van de scholen betrokken, het administratief personeel in 31% van de scholen. Verder blijkt dat ook ouders, het schoolbestuur/inrichtende macht, het CLB en de pedagogische begeleidingdienst in een behoorlijk aantal scholen betrokken worden bij het bespreken van de resultaten. Het bespreken van zelfevaluatiegegevens met leerlingen is dan weer duidelijk geen gangbare praktijk in de bevraagde scholen.

[10] Er zijn ook scholen die niet beschikken over een middenkader. De gerapporteerde gegevens voor het middenkader zijn dus geen perfect valide indicatoren voor de betrokkenheid van het middenkader in de verschillende fasen van de zelfevaluatie. De werkelijke percentages voor de scholen met een middenkader liggen dus wellicht hoger.

Tabel 38: Betrokkenheid bij de verschillende fasen van de zelfevaluatie – Perspectief schoolleiders

	Voorbereiding	Gegevensverzameling	Bespreken van de resultaten
De directeur	99%	75%	97%
Het middenkader	54%	46%	53%
Leerkrachten	42%	93%	93%
Administratief personeel	9%	33%	31%
Leerlingen	0%	11%	5%
Ouders	3%	14%	17%
Schoolbestuur/ Inrichtende Macht	7%	1%	19%
CLB	6%	3%	10%
Pedagogische begeleidingdienst	21%	10%	24%

Naast het bevragen van schoolleiders over wie er betrokken werd bij de zelfevaluatie in hun scholen, werd in de eigenlijke vragenlijst aan de teamleden van deze scholen gevraagd of zij betrokken werden bij de verschillende fasen van de zelfevaluatie. Tabel 39 zet de resultaten van deze invalshoek op een rij. Zoals blijkt is de verdeling in fasen van de zelfevaluatie wel enigszins verschillend. Het gaat om: het beslissen over de doelen en/of de inhoud van de zelfevaluatie; het kiezen en/of ontwikkelen van de gebruikte vragenlijst; het invullen van de vragenlijst; en het interpreteren van de zelfevaluatiegegevens en/of het komen tot actiepunten.

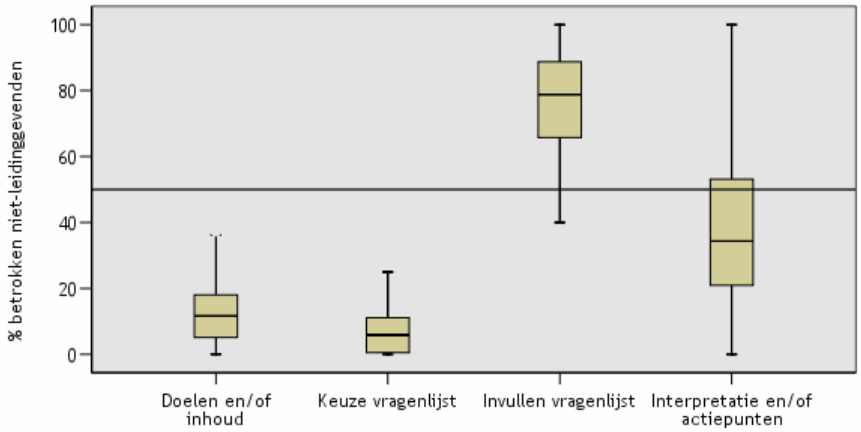
Tabel 39: Betrokkenheid bij de verschillende fasen van de zelfevaluatie – Perspectief alle respondenten

Functie van de respondenten	Doelen en/of inhoud	Keuze vragenlijst	Invullen vragenlijst	Interpretatie en/of actiepunten
(Adjunct-)Directeur	73.6%	69.8%	83.0%	81.1%
Middenkader	42.2%	34.8%	79.3%	66.3%
Leraar	11.1%	6.7%	74.6%	32.4%
GOK-leraar	15.8%	10.5%	70.2%	33.3%
Pedagogische, paramedische ondersteuning	12.7%	8.9%	69.9%	31.6%
Administratieve of ICT ondersteuning	6.3%	4.8%	70.6%	20.6%
Onderwijsniveau school				
Basisonderwijs	12.6%	9.5%	71.6%	26.8%
Secundair onderwijs	18.0%	11.4%	80.0%	49.2%
Totaal	*14.6%*	*10.2%*	*74.6%*	*34.9%*

Uit Tabel 39 blijkt duidelijk dat de respondenten voornamelijk betrokken werden bij het invullen van de vragenlijst. Afhankelijk van de functie die de respondent in de school bekleedt, geeft immers 70% à 83% aan de vragenlijst te hebben ingevuld. Deze verschillen zijn niet statistisch significant. De verschillen bij de andere fasen van de zelfevaluatie zijn dat wel. 'Beslissen over de doelen en/of de inhoud van de zelfevaluatie' en 'kiezen en/of ontwikkelen van een vragenlijst' blijken voornamelijk een activiteit voor de (ad-

junct-)directeur en het middenkader. Toch geeft respectievelijk maar 42% en 35% van de leden van het middenkader aan bij deze fasen betrokken te zijn geweest. Uit de (eerder) lage percentages leren we dat deze fasen het werk zijn van een vaak beperkte groep in scholen. Ook voor 'het interpreteren van de zelfevaluatiegegevens en/of het komen tot actiepunten' is er een statistisch significant effect van de functie in de school. De betrokkenheid van de niet-leidinggevenden in de school mag in deze fase dan wel groter zijn, ze blijft duidelijk lager dan deze van de leidinggevenden (Pearson X^2 (5, 2711) = 157.6; $p < 0.001$). In totaal geeft evenwel 35% van de bevraagde respondenten aan bij deze fase betrokken te zijn geweest. Tabel 39 geeft verder ook aan dat er verschillen zijn tussen de onderwijsniveaus. Enkel voor 'kiezen en/of ontwikkelen van een vragenlijst' zijn deze niet statistisch significant. We stelden ook vast dat respondenten uit basisscholen vaker aangeven betrokken te zijn geweest bij het interpreteren van de zelfevaluatiegegevens en/of het komen tot actiepunten dan respondenten uit het secundair onderwijs (49% vs. 27%; Pearson X^2 (1, 2711) = 137.9; $p < 0.001$).

Figuur 22: Percentage niet-leidinggevenden dat in de scholen (n=96) betrokken werd



Om af te ronden richten we ons op de vraag of er verschillen zijn tussen scholen inzake betrokkenheid. Daarom berekenden we per school het aantal niet-leidinggevende respondenten dat in de vier fasen betrokken werd. De gemiddelde cijfers in Figuur 22 weerspiegelen de trends in Tabel 39. De boxplots maken wel duidelijk dat er tussen scholen soms grote verschillen zijn wat de betrokkenheid van de niet-leidinggevenden betreft. Voornamelijk wat de betrokkenheid bij 'het interpreteren van de zelfevaluatiegegevens en/of het komen tot actiepunten' betreft blijken scholen zeer sterk van elkaar te verschillen. De eerder beperkte betrokkenheid bij 'beslissen over de doelen en/of de inhoud van de zelfevaluatie' en 'kiezen en/of ontwikkelen van een vragenlijst' blijkt zich dan weer in de meeste scholen te manifesteren.

4 Ondersteuning

Scholen kunnen er bij het uitvoeren van een zelfevaluatie voor kiezen om een beroep doen op externe ondersteuning. Het gaat dan om deskundigen die geen lid zijn van de eigen schoolorganisatie. Ook de aard van de ondersteuning waarop scholen een beroep doen, kan verscheidene vormen aannemen. Precies die variatie willen we in eerste instantie in kaart brengen.

Uit de resultaten van de inleidende bevraging van schoolleiders blijkt dat 29% van de scholen voor het uitvoeren van de zelfevaluatie een beroep deed op de pedagogische begeleiding. 31% deed een beroep op een nascholingsinstituut. Overleggen met andere scholen met ervaring inzake zelfevaluatie blijkt minder vaak voor te komen. Slechts 6% van de schoolleiders geeft aan dat deze ondersteuningsstrategie in hun school gevolgd werd. Scholen die voor ondersteuning kiezen, kiezen dus doorgaans voor een nascholingsinstituut of voor de pedagogische begeleiding. Niet meer dan 4% van de scholen combineerde meerdere ondersteuningsvormen. In bijna 40% van de scholen geeft de schoolleider aan dat de zelfevaluatie zonder enige vorm van ondersteuning verlopen is.

De vraag naar de gebruikte ondersteuning werd ook aan de individuele respondenten voorgelegd. Er werd hen gevraagd aan welke vormen van ondersteuning (met betrekking tot het uitvoeren van zelfevaluaties) zij de voorbije vijf schooljaren deelnamen. Daartoe werd het lijstje onderaan Tabel 40 gebruikt. De totale aantallen geven meteen aan dat een grote groep respondenten geen ondersteuning genoot naar aanleiding van de zelfevaluatie. Van de bevraagde ondersteuningsvormen blijken het lezen van documentatiemateriaal en het bijwonen van studiedagen binnen de school het meest vaak plaats te vinden. 17% van de totale groep respondenten geeft aan dat deze op hen van toepassing zijn. Ook dat aantal ligt evenwel laag, wetende dat de respondenten allemaal werkzaam zijn in een school waar recent een zelfevaluatie plaatsvond. De andere ondersteuningsvormen werden door een zeer beperkte groep respondenten bijgewoond. Studiebezoeken bij andere scholen met zelfevaluatie-ervaring blijken zelfs zo goed als onbestaande. Doorgaans blijken het enkel de (adjunct)directeur en leden van het middenkader te zijn die aan ondersteuning deelnemen. Een noemenswaardig deel van hen heeft documentatiemateriaal gelezen en woonde een externe studiedag bij over zelfevaluatie. Het zijn ook zij die het meest vaak betrokken worden bij begeleidingtrajecten verzorgd door de pedagogische begeleiding en nascholingsinstituten. GOK-leraren blijken in zekere mate meer ondersteuning te hebben genoten dan andere leraren. Tabel 40 rapporteert ook over het onderscheid tussen basisscholen en secundaire scholen. Over de onderwijsniveaus heen zijn er geen statistisch significante verschillen.

Tabel 40: Percentage respondenten dat gebruik maakte van specifieke ondersteuning

Functie van de respondenten							
(Adjunct)directeur	65.1%	17.9%	50.0%	34.0%	20.8%	20.8%	6.6%
Middenkader	37.0%	18.5%	23.9%	20.7%	10.9%	12.0%	1.1%
Leraar	14.4%	17.4%	4.8%	1.7%	3.7%	2.3%	1.3%
GOK-leraar	17.5%	14.0%	12.3%	5.3%	8.8%	5.3%	0%
Pedagogische, paramedische ondersteuning	13.8%	6.3%	3.8%	1.3%	3.8%	5.0%	3.8%
Administratieve of ICT ondersteuning	9.5%	15.0%	3.2%	3.2%	1.6%	3.2%	0.8%
Onderwijsniveau school							
Basisonderwijs	15.0%	16.6%	5.9%	3.6%	4.9%	5.0%	1.2%
Secundair onderwijs	18.1%	17.1%	8.0%	3.9%	4.5%	2.7%	1.7%
Totaal	*17.0%*	*16.9%*	*7.2%*	*3.8%*	*4.6%*	*3.5%*	*1.5%*
	Ik heb documentatiemateriaal gelezen over zelfevaluaties door scholen.	Ik heb op onze school een studiedag bijgewoond die handelde over zelfevaluatie.	Ik heb een externe studiedag bijgewoond (een halve of hele dag) over zelfevaluatie.	Ik heb meerdaagse cursussen/trainingen/ workshops gevolgd.	Ik ben betrokken geweest bij een begeleidingtraject verzorgd door de ped. beg.	Ik ben betrokken geweest bij een begeleidingtraject verzorgd door een nascholingsinstituut.	Ik heb een studiebezoek gedaan bij andere scholen met zelfevaluatie-ervaring.

5 De kernkenmerken van zelfevaluaties

De gegevens die hierboven gepresenteerd werden maken het mogelijk de kernkenmerken van de 'gemiddelde zelfevaluatie' te beschrijven. We hanteren daartoe het profiel dat we in het theoretisch kader samenstelden (Tabel 41). Dit profiel is een bondig overzicht van de informatie die in dit deel beschreven werd. We kunnen in principe voor alle scholen die aan deze studie participeerden een eigen zelfevaluatieprofiel uittekenen. Om de variatie in zelfevaluaties te illustreren doen we dat hier voor twee scholen. Hun profielen worden geïntegreerd in Tabel 42. School 1 is een basisschool, school 2 een school voor secundair onderwijs. In de tabel is telkens sprake van 'school 1/school 2'. Omdat het over concrete scholen gaat is er vaak sprake van Ja/nee-antwoorden in plaats van percentages. Beide scholen hebben een vergelijkbaar aantal leerlingen (ruim 350) en aantal personeelsleden (om en bij 35). Ze behoren beide tot de cluster 'gemiddeld' voor 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad'. Wat hun zelfevaluatieprofiel betreft wijken de twee scholen evenwel op een aantal punten van elkaar af. Ze verschillen inzake het initiatief, de functies van de zelfevaluatie, de inhouden, het gebruikte instrument en de betrokkenheid van de verschillende betrokkenen in de verschillende fasen van de zelfevaluatie. Dit is een illustratie die onze verwachting onderschrijft dat zelfevaluaties sterk kunnen variëren wat hun kernkenmerken betreft en toch gelijkaardig kunnen scoren op andere concepten uit het theoretisch kader.

Tabel 41: De kernkenmerken van de 'gemiddelde zelfevaluatie'

Geïnitieerd door…	94%	**De school zelf**			6%	**Een externe instantie**		
Evaluatiefunctie(s)	97%	**Spiegel**	24%	**Legitimatie**	2%	**Transfer**	76%	**Beleidsvoorbereiding**
Participanten aan de zelfevaluatie (betrokken tijdens de planning, de dataverzameling en/of de bespreking van de zelfevaluatieresultaten)			**Planning**		**Gegevensverzameling**		**Bespreking**	
		De directeur	99%		75%		97%	
		Middenkader	54%		46%		53%	
		Leerkrachten	42%		93%		93%	
		Adm. personeel	9%		33%		31%	
		Leerlingen	0%		11%		5%	
		Ouders	3%		14%		17%	
		Schoolbest./IM	7%		1%		19%	
		CLB	6%		3%		10%	
		Ped. beg.	21%		10%		24%	
Methodologische aanpak	100%	Kwantitatief: Survey bij leerkrachten			32%	Kwalitatief: focusgroep- of verhelderingsgesprekken, een analyse van beschikbare cijfergegevens of individuele interviews		
	78%	Bestaand instrument(en): IZES, DISO, BIOS, SION, PROZA en andere			24%	Eigen instrument(en)		
Geëvalueerde thema's	De onderwijskundige aanpak in de school, het functioneren als team, het leiderschap in de school, het leren van leerlingen en de pedagogisch/levensbeschouwelijke aanpak in de school.							

Tabel 42: De kernkenmerken van twee zelfevaluaties geïllustreerd

Geïnitieerd door…	J/J	**De school zelf**			J/N	**Een externe overheid**		
Evaluatiefunctie(s)	J/J	**Spiegel**	N/J	**Legitimatie**	N/N	**Transfer**	J/J	**Beleidsvoorbereiding**
Participanten aan de zelfevaluatie (betrokken tijdens de planning, de dataverzameling en/of de bespreking van de zelfevaluatieresultaten)			**planning**		**gegevensverzameling**		**bespreking**	
		De directeur	Ja/ja		Ja/nee		Ja/ja	
		Middenkader	Ja/nee		Ja/nee		Ja/nee	
		Leerkrachten	nee/ja		Ja/ja		Ja/ja	
		Adm. personeel	nee/nee		Ja/nee		Ja/nee	
		Leerlingen	Nee/nee		Nee/nee		Nee/nee	
		Ouders	Nee/nee		Nee/nee		Nee/nee	
		Schoolbest./IM	Nee/nee		Nee/nee		Nee/nee	
		CLB	Nee/nee		Nee/nee		Nee/nee	
		Ped. beg.	Ja/nee		Nee/nee		Nee/nee	
Methodologische aanpak	J/J	(In beide scholen) Kwantitatief: Survey bij leerkrachten			J/J	(In beide scholen) Kwalitatief: focusgroep- of verhelderingsgesprekken		
	J/J	(School 1) Bestaand instrument(en): IZES (School 2) Bestaand instrument(en): DISO			N/N	(In beide scholen) Eigen instrument(en):Nee		
Geëvalueerde thema's	(School 1) Schooldoelen, schoolprofiel, zelfstandigheidbevordering bij leerlingen, evaluatiebeleid (School 2) Het leren van leerlingen, het leiderschap in de school, het functioneren als team, de onderwijskundige aanpak in de school							

6 Afrondend

De resultaten in dit deel presenteerden wij met het oog op het beschrijven van verschillen tussen (de kernkenmerken van) zelfevaluaties. Onze eerste onderzoekshypothese nam als uitgangspunt dat het verloop van geïnstrumentaliseerde zelfevaluaties geen vastliggende standaard kent. De bovenstaande resultaten tonen aan dat zelfevaluatieprocessen inderdaad telkens op een specifieke manier ingevuld worden.

In het volgende deel richten we ons op de andere onderzoekshypothesen. We proberen met name de verschillen in de kwaliteit van zelfevaluaties en in de zelfevaluatiegebonden kenmerken te beschrijven en te verklaren.

DEEL 7

ZELFEVALUATIES IN DE BESTUDEERDE SCHOLEN – VIA HYPOTHESETOETSING NAAR MODELVORMING

In dit deel ligt de nadruk op relaties tussen de schoolgebonden kenmerken, de zelfevaluatiegebonden kenmerken, de ondersteuning en de kwaliteit van zelfevaluaties. In paragraaf 1 toetsen we de verschillende hypothesen die we aan het einde van het theoretisch kader vooropstelden. Dat doen we om inzicht te krijgen in hoe zelfevaluaties verschillen en hoe die verschillen kunnen verklaard worden. Daarnaast is het doel te komen tot een selectie van indicatoren met betrekking tot de zelfevaluatie en de school waarin deze plaats vindt. Deze selectie moet het mogelijk maken een zo spaarzaam mogelijk model te ontwikkelen voor het verklaren van verschillen in de kwaliteit van zelfevaluaties. Dit model ontwikkelen en toetsen we in paragraaf 2. We ronden af met een aantal conclusies (zie paragraaf 3).

1 TOETSEN VAN DE ONDERZOEKSHYPOTHESEN

De beschrijving van het theoretisch kader dat aan de grondslag ligt van het kwantitatieve luik van deze studie eindigde met het opsommen van een aantal onderzoekshypothesen. Doorheen de volgende paragrafen zullen deze hypothesen getoetst worden. We hanteren daarbij de volgorde waarin de hypotheses geformuleerd werden (zie p. 114).

1.1 De houding ten aanzien van zelfevaluatie

In deze paragraaf richten we ons op de houding van de respondenten ten aanzien van zelfevaluatie. We presenteren eerst een aantal beschrijvende resultaten. Daarna toetsen we de impact van een aantal school- en respondentgebonden kenmerken op de houding ten aanzien van zelfevaluatie. Dat doen we aan de hand van meerniveau-analyses.

1.1.1 Beschrijvende resultaten

Tabel 43 geeft aan hoe de bevraagde respondenten tegenover zelfevaluatie staan. De frequentieverdeling geeft het standpunt van de respondenten weer op een continuüm dat zich tussen twee inhoudelijke posities in bevindt. Respondenten die 1 scoren leunen aan

bij de linkse variant van de stelling (bijvoorbeeld zelfevaluatie is subjectief); respondenten die 6 scoren bij de rechtse (bijvoorbeeld zelfevaluatie is objectief). Uit de tabel blijkt dat men zich gemiddeld genomen eerder positief uitspreekt over een aantal wenselijke resultaten van zelfevaluatie. Ruim drie kwart van de respondenten onderschrijft met name dat men er veel uit leert, dat het leidt tot beter lesgeven en tot beter management. Ondanks deze positieve verwachtingen is zelfevaluatie echter niet geliefd. De vraag of zelfevaluatie bij de meeste teamleden al dan niet geliefd is, heeft de laagste gemiddelde score uit het hele rijtje (Gem. = 2.75). Deze negatieve houding lijkt niet zozeer toe te schrijven aan een te beperkte resultaatsverwachting. Het zijn met name items die betrekking hebben op proceskenmerken die eerder laag scoren. De respondenten geven doorgaans immers aan dat zelfevaluatie veel tijd kost (Gem. = 3.27), dat het een subjectief gebeuren is (Gem. = 3.39) en dat het moeilijk uit te voeren is (Gem. = 3.56). Deze items scoren gemiddeld onder of op het middenpunt van de beoordelingsschaal (i.e. 3.5).

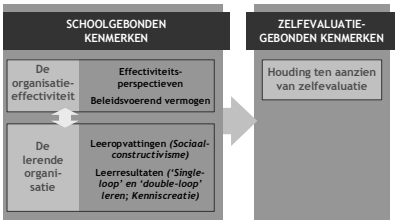
Tabel 43: Frequentieverdeling (%), gemiddelde score en standaarddeviatie voor de items m.b.t. de houding t.a.v. zelfevaluatie (n = 2686)

Zelfevaluatie...	1	2	3	4	5	6	Gem.	SD
...vertelt ons niets nieuws/leert ons heel veel	4	8	13	31	29	17	4.24	1.29
...leidt niet tot beter lesgeven/leidt tot beter lesgeven	4	8	12	28	31	16	4.22	1.32
...kost veel tijd/kost weinig extra tijd	11	22	23	24	16	5	3.27	1.38
...raakt alleen enkelen/raakt iedereen	6	16	20	24	21	12	3.75	1.43
...leidt niet tot beter management/leidt tot beter management	6	10	11	24	34	16	4.18	1.40
...is moeilijk te interpreteren/is makkelijk te begrijpen	7	14	22	30	21	6	3.64	1.31
...is bij de meeste teamleden niet geliefd/is wel geliefd	18	29	24	21	7	1	2.75	1.24
...is moeilijk uit te voeren/is goed uit te voeren	6	17	24	27	19	7	3.56	1.33
...is niet kosten-effectief/is kosten-effectief	8	12	22	29	21	6	3.62	1.34
...hangt af van het toeval/resulteert in een betrouwbaar beeld	4	9	20	33	28	5	3.89	1.18
...is subjectief/is objectief	10	19	22	27	19	4	3.39	1.35
...is een momentopname/geeft onze ontwikkeling goed weer	8	14	18	30	26	6	3.69	1.34
Schaalscore 'Houding ten aanzien van zelfevaluatie'							*3.68*	*0.95*

Zowel uit de frequentieverdeling als uit de standaarddeviatie bij de gemiddelde scores (zie Tabel 43) blijkt dat er bij elk item een aanzienlijke variatie is in de beoordeling door de respondenten. We analyseren de variantie echter enkel op schaalniveau. De schaal 'houding ten aanzien van zelfevaluatie' kent een gemiddelde van 3.68 en een standaarddeviatie van 0.95. Deze variantie blijkt samen te hangen met een aantal achtergrondkenmerken van scholen en respondenten. Dat blijkt tenminste uit de meerniveauanalyses.

1.1.2 Meerniveau-analyses

Alvorens de concrete resultaten voor de houding ten aanzien van zelfevaluatie te beschrijven, staan we stil bij onze manier van rapporteren over de meerniveau-analyses.



In de tabellen met outputgegevens (zie bijvoorbeeld Tabel 44) zijn in eerste instantie de geschatte parameters voor het *basismodel* (of nulmodel) opgenomen. Het intercept geeft informatie over de gemiddelde score voor de betreffende afhankelijke variabele. Daarnaast verkrijgen we informatie over de variantie op school- en respondentniveau. Daarna wordt informatie gegeven over het *brutomodel*. Dit model geeft het effect weer van elk van de kenmerken op school- en respondentniveau zonder dat er gecontroleerd wordt voor de andere kenmerken. Los van de vraag waaraan de gevonden verschillen moeten toegeschreven worden, krijgen we op basis van de brutomodellen bijvoorbeeld een idee van verschillen tussen onderwijsniveaus, tussen schoolleiders en teamleden, en tussen scholen met en zonder ondersteuning. Dat zijn op zich reeds relevante vaststellingen. Hoewel ze tezamen gerapporteerd worden, gaat het in de brutomodellen in wezen telkens om een andere analyse. Tenslotte is er het *nettomodel* waarin zowel de relevante school- als respondentkenmerken opgenomen zijn. Op die manier krijgen we een beeld van het netto-effect, dat wil zeggen na controle voor de mogelijke effecten van de andere variabelen in het model (De Maeyer & Rymenans, 2004). In het nettomodel worden in eerste instantie al de verklarende variabelen opgenomen waarvoor we van voldoende respondenten over gegevens beschikken. Door bijvoorbeeld ook de onderwijsvorm (enkel in het secundair onderwijs) of de beoordeling van de ondersteuning (enkel ingevuld door de respondenten die ondersteuning genoten) op te nemen, zou de analyse te zeer aan statistische power inboeten. Het aantal respondenten daalt dan immers sterk. Variabelen die we niet in het nettostartmodel opnamen, duiden we aan met 'niet opgenomen' (NO). Uit het nettostartmodel worden de variabelen verwijderd als deze verwijdering niet voor een statistisch significante toename van de deviatie zorgt (gegeven dat het verschil in deviatie een X^2-verdeling volgt, mag het verschil bij 1 vrijheidsgraad niet groter zijn dan 3.9) (Luke, 2004). Het finale model dat het resultaat is van deze analyse wordt in de tabel gepresenteerd. Per regressiecoëfficiënt rapporteren we de schatting en de standaardfout[11] (tussen haakjes). Statistisch significante effecten worden vet afgedrukt. Voor de interpretatie is het tenslotte nog van belang om mee te geven dat al de variabelen in de meerniveau-modellen gestandaardiseerd zijn[12].

---

[11] Hoe groter de standaardfout, hoe groter de kans dat de geschatte parameterwaarde afwijkt van de echte (maar ongekende) waarde in de populatie. De standaardfout wordt gebruikt bij de berekening van het betrouwbaarheidsinterval. Als het betrouwbaarheidsinterval (i.e. de geschatte waarde +/- 1.96 keer de standaardfout) de waarde 0 niet bevat, kunnen we met 95% zekerheid concluderen dat de geschatte parameter in de populatie afwijkt van 0 en dus een statistisch significante impact heeft.

[12] Gestandaardiseerde variabelen worden verkregen door de oorspronkelijke variabele om te zetten naar een variabele met een gemiddelde van 0 en een standaarddeviatie van 1. Dat maakt het moge-

Tabel 44: Meerniveau-analyses 'Houding ten aanzien van zelfevaluatie'

	Basismodel	Brutomodel	Nettomodel
VAST	Schatting (SF)	Schatting (SF)	Schatting (SF)
Intercept	.025 (.034)		-.113 (.039)
Voorspellers: info over de school			
Basisonderwijs (vs. secundair onderwijs)		.124 (.067)	
Vrij gesubsidieerd onderwijs (vs. GO)		.056 (.069)	
Officieel gesubsidieerd onderwijs (vs. GO)		-.088 (.109)	
Geen ervaring men zelfevaluatie (vs. wel ervaring)		-.026 (.070)	
Geen ondersteuning (vs. wel ondersteuning)		-.008 (.071)	
Schoolgrootte (aantal leerkrachten)		-.049 (.038)	
Schoolcultuur cluster 1 (vs. cluster2)		-.101 (.129)	
Schoolcultuur cluster 3 (vs. cluster2)		**.316 (.062)**	**.131 (.057)**
Voorspellers: info over de (percepties van de) respondent			
Schoolleiding (vs. teamleden)		**.589 (.072)**	**.453 (.068)**
Voornamelijk werkzaam in TSO (vs. ASO)		.027 (.067)	NO
Voornamelijk werkzaam in BSO (vs. ASO)		.103 (.064)	NO
Wel ondersteuning (vs. geen ondersteuning)		**.213 (.040)**	**.108 (.041)**
Perceptie kwaliteit proces ondersteuning		**.347 (.031)**	NO
Perceptie kwaliteit resultaat ondersteuning		**.500 (.032)**	NO
Leeftijd		-.028 (.019)	
Perceptie van 'human relations'		**.171 (.022)**	**-.061 (.029)**
Perceptie van 'interne proces'		**.139 (.022)**	-.020 (.025)
Perceptie van 'open systeem'		**.192 (.021)**	**.057 (.027)**
Perceptie van 'rationele doel'		**.169 (.021)**	**.064 (.025)**
Perceptie van het eigen leren		**.274 (.020)**	**.154 (.021)**
Perceptie van diepgaand leren in de school		**.359 (.026)**	**.210 (.025)**
Perceptie van mate van kenniscreatie in de school		**.326 (.022)**	**.170 (.025)**
VARIABEL			
Variantie op schoolniveau	.071 (0.016)		.031 (.010)
Variantie op respondentenniveau	.936 (0.026)		.655 (.022)
R^2 - Reductie van variantie op schoolniveau			30%
R^2 - Reductie van variantie op respondentenniveau			32%
MODEL FIT			
Deviatie	7548		4531
Aantal scholen	96		96
Aantal respondenten	2686		1852

lijk om de scores van respondenten op verschillende schalen te vergelijken. Omdat de variabelen normaal verdeeld zijn weten we bovendien dat 68% van de respondenten tussen 1 standaarddeviatie onder of boven het gemiddelde scoort; 95% scoort binnen twee standaarddeviaties van het gemiddelde.

We staan bij de interpretatie van de resultaten achtereenvolgens stil bij het basismodel, het brutomodel en het nettomodel[13].

Uit het basismodel onthouden we in eerste instantie dat een meerniveau-analyse inderdaad aangewezen is. De varianties op school- en op respondentniveau zijn beide significant verschillend van 0. Respectievelijk 7% van de totale variantie in de houding van de respondenten ten aanzien van zelfevaluatie is toe te schrijven aan verschillen tussen scholen. De resterende 93% moet toegeschreven worden aan verschillen tussen respondenten binnen scholen. Dat is op zich reeds een belangrijke vaststelling. De verschillen binnen scholen zijn veel groter dan de verschillen tussen scholen.

Als we ons concentreren op het brutomodel stellen we voornamelijk statistisch significante effecten vast voor respondentgebonden kenmerken. Op schoolniveau vinden we enkel een verschil tussen twee scholenclusters met betrekking tot de effectiviteitsperspectieven. Respondenten uit scholen die sterk scoren op de verschillende effectiviteitsperspectieven scoren 0.32 SD hoger op de houding ten aanzien van zelfevaluatie dan scholen die laag scoren op de beheersgerichte effectiviteitsperspectieven[14]. De andere verschillen tussen deze scholenclusters zijn niet statistisch significant. Op respondentniveau blijkt voornamelijk de positie die men inneemt in de school van groot belang. Teamleden die betrokken zijn bij het leiden van scholen ((adjunct)directeurs en leden van het middenkader) scoren 0.59 SD hoger dan overige teamleden ((GOK)leerkrachten en ondersteunend personeel). Ook de indicatoren van het gebruik en de beoordeling van ondersteuning hebben een statistisch significante impact. De houding van respondenten ten aanzien van zelfevaluatie is positiever als ze tenminste één vorm van ondersteuning genoten hebben en naarmate ze zich positiever uitspreken over het proces en de resultaten van die ondersteuning. Tenslotte gaat ook van elk van de opgenomen schoolgebonden indicatoren een statistisch significant effect uit op de houding ten aanzien van zelfevaluatie. De meest sterke effecten vinden we voor de perceptie van respondenten over de mate van diepgaand leren en kenniscreatie in hun school. Per verschil van één standaardafwijking

---

[13] Omdat dit het eerste meerniveau-model is dat we in dit proefschrift schatten, geven we even aan dat de controle op de achterliggende assumpties geen problemen aan het licht bracht. De residuen zijn immers normaal verdeeld. Hun gemiddelde ligt duidelijk op 0. Er blijken bovendien geen problemen te zijn met uitbijters. De gestandaardiseerde residuen zijn in elke school gecentreerd rond 0 en er zijn geen ernstige problemen inzake heteroscedasticiteit. Dat blijkt ook niet het geval wanneer de residuen afgezet worden tegen het vaste deel van de predictie volgens het nettomodel. De variabiliteit van de residuen is grofweg hetzelfde voor de verschillende uitkomsten van de predictie. Het getoetste model is dus adequaat. Tenzij er expliciet op ingegaan wordt, is dit ook het geval voor de modellen waarover we verderop rapporteren. Dat geldt ook voor het toetsen van de noodzaak om regressiecoëfficiënten te laten variëren over scholen en om interactie-effecten op te nemen in het model. De nodige analyses daartoe werden uitgevoerd maar wanneer we hierover niet rapporteren betekent dit dat verdere complicatie van het gerapporteerde model niet tot een statistisch significant betere voorspelling leidt.

[14] Wanneer respondenten uit deze laatste groep bijvoorbeeld gemiddeld 0.15 scoren op 'houding ten aanzien van zelfevaluatie' dan betekent dit dat respondenten uit scholen die sterk scoren op de verschillende effectiviteitsperspectieven gemiddeld 0.47 scoren. Vermits we met gestandaardiseerde scores werken is de standaarddeviatie immers gelijk aan 1.

tussen respondenten op deze indicatoren, zullen zij gemiddeld respectievelijk 0.36 SD en 0.33 SD verschillend scoren op de houding ten aanzien van zelfevaluatie. Belangrijke vaststellingen zijn ook dat verschillende achtergrondkenmerken van scholen en respondenten niet statistisch significant samenhangen met de houding van respondenten ten aanzien van zelfevaluatie. Het gaat onder meer om het onderwijsniveau, het onderwijsnet, de onderwijsvorm[15] waarin men hoofdzakelijk actief is (voor respondenten uit het secundair onderwijs), het als school al dan niet ervaring hebben met zelfevaluatie, het als school al dan niet een beroep gedaan hebben op ondersteuning bij de zelfevaluatie, de schoolgrootte uitgedrukt in aantal leerkrachten en de leeftijd van de respondent.

We geven bij de interpretatie van het nettomodel eerst mee dat dit model de voorspellende kracht in vergelijking met het nulmodel met 30% verbetert op schoolniveau en met 32% op respondentniveau[16]. De eerste inhoudelijke vaststelling uit het nettomodel is dat de positie die men inneemt in de school ook een zeer belangrijke rol blijkt te spelen nadat voor andere kenmerken gecontroleerd wordt. Los van de vraag hoe zij denken over leren en hoe zij de schoolcultuur percipiëren, scoren leden van de schoolleiding 0.45 SD hoger dan de overige teamleden. Ook het gebruik van een vorm van ondersteuning hangt samen met een meer positieve houding ten aanzien van zelfevaluatie. De schalen in verband met de beoordeling van de genoten ondersteuning werden niet in de analyse opgenomen. In een latere analyse zullen we ons echter specifiek op de rol van de ondersteuning toeleggen. Tenslotte leren we uit het nettomodel dat de meeste schoolgebonden indicatoren ook na controle voor andere kenmerken een statistisch significant effect hebben op de houding ten aanzien van zelfevaluatie. De regressiecoëfficiënten liggen weliswaar lager dan in het brutomodel. De grootste impact gaat uit van 'diepgaand leren' en 'de mate van kenniscreatie'. Opvallend is dat de perceptie van het belang van 'menselijke relaties' na controle een statistisch significant negatief effect heeft. Na controle voor de andere school- en respondentkenmerken blijken respondenten dus een minder positieve houding te hebben ten aanzien van zelfevaluatie naarmate ze meer van mening zijn dat hun schoolcultuur meer gekenmerkt wordt door een sterke nadruk op dit effectiviteitsperspectief (cf. interne flexibiliteit).

1.1.3 Conclusies bij de vooropgestelde onderzoekshypothesen

De bevraagde respondenten koesteren positieve verwachtingen ten aanzien van zelfevaluatie. Toch gaat het niet om een geliefde activiteit. De oorzaak van deze negatieve houding lijkt eerder bij het proces dan bij het product van zelfevaluatie te liggen. Men vindt dat zelfevaluatie veel tijd kost, dat het een subjectief gebeuren is en dat het moei-

---

[15] De onderwijsvorm blijkt systematisch geen impact te hebben op de afhankelijke variabelen in de meer-niveauanalyses die we zullen presenteren. Hoewel het onderzocht werd, wordt er verder dus niet over gerapporteerd.

[16] De R^2-waarden werden berekend volgens de benadering van Snijders & Bosker (1999) en moeten anders geïnterpreteerd worden dan R^2 in de klassieke regressieanalyse. We spreken over een verhoudingsgewijze reductie van error bij de schatting van de afhankelijke variabele.

lijk uit te voeren is. Zoals verwacht is de globale houding van respondenten ten aanzien van zelfevaluatie dan ook slechts in beperkte mate positief (cf. hypothese 2a). Ook uit de meerniveau-analyses met betrekking tot de houding ten aanzien van zelfevaluatie concluderen we dat onze theoretische verwachtingen grotendeels bevestigd worden. De houding van respondenten is zoals verwacht duidelijk positiever naarmate de schoolgebonden kenmerken (organisatie-effectiviteit en lerende organisatie) meer van kracht zijn (cf. hypothese 2b). Enkel het beperkte netto-effect van 'human relations' wijst in een andere richting. De onderzoekshypothese over de verschillen tussen schoolleiders en leerkrachten wordt ook door de data bevestigd. De andere onderzochte achtergrondkenmerken van de school en de respondent blijken echter niet statistisch significant. We onthouden verder ook de grote variatie tussen respondenten binnen scholen. De verschillen binnen de scholen zijn groter dan de verschillen tussen scholen. Toch zijn ook de verschillen tussen scholen statistisch significant.

1.2 Het verloop van zelfevaluaties: zelfevaluatie als beleidsdaad

Het uitgangspunt in het theoretisch kader was het uitvoeren van een zelfevaluatie te aanschouwen als een specifieke beleidsdaad. Het verloop van zelfevaluaties werd vanuit die invalshoek in twee verschillende contexten bevraagd. We bespreken eerst kort de bevindingen met betrekking tot de GOK-zelfevaluaties, daarna richten we ons op de resultaten van 'zelfevaluatie binnenstebuiten'.

1.2.1 De dragers van 'zelfevaluatie als beleidsdaad' binnen de GOK-zelfevaluaties

In het onderzoek naar de GOK-zelfevaluatie werd in kaart gebracht in welke mate scholen volgens schoolleiders (n=518) tegemoet kwamen aan de dragers van 'zelfevaluatie als beleidsdaad'. Voor elke drager van de 'zelfevaluatie als beleidsdaad' worden in Tabel 34 de beschrijvende resultaten van deze bevraging weergegeven. Het meest opvallende resultaat is de lage gemiddelde score op de schaal responsiviteit van de zelfevaluatie. De schoolleiders geven aan dat hun school er in zeer beperkte mate in slaagde om externen te betrekken bij de GOK-zelfevaluatie en om aandacht te besteden aan de inbreng van bijvoorbeeld ouders. Op de andere schalen scoren zelfevaluaties volgens de schoolleiders beduidend hoger. De gemiddelde scores situeren zich tussen 3.56 en 3.80. Schoolleiders spreken zich dus eerder positief uit over de andere dragers van 'zelfevaluatie als beleidsdaad' in hun school (GEM=3.5).

Tabel 45: Beschrijvende statistieken voor dragers van en het concept 'zelfevaluatie als beleidsdaad'

Schaal	GEM.	min	max	SD
Communicatief vermogen	3.80	1.20	5.00	.69
Ondersteuning & samenwerking	3.60	1.00	5.00	.72
Gezamenlijke doelgerichtheid	3.59	1.00	5.00	.74
Gedeeld leiderschap	3.82	1.00	5.00	.74
Responsief vermogen	1.91	1.00	5.00	.90
Innovatief vermogen	3.80	1.86	5.00	.57
Geïntegreerde aanpak	3.62	1.40	5.00	.65
Reflectief vermogen	3.56	1.00	5.00	.74
'Zelfevaluatie als beleidsdaad'	*3.50*	*1.65*	*4.88*	*.56*

De gemiddelde scores verbergen een aanzienlijke variatie tussen schoolleiders. Dat blijkt onder meer uit de standaarddeviaties. Om die variantie te analyseren onderzochten we in eerste instantie welke de unieke bijdrage is van een aantal achtergrondkenmerken van scholen op de perceptie van schoolleiders van de dragers van 'zelfevaluatie als beleidsdaad'. Naast onderwijsnet en onderwijsniveau werden daarom ook de schoolgrootte en het al dan niet hebben van ervaring met overheidsinitiatieven inzake gelijke onderwijskansen in een multiple regressieanalyse opgenomen.

Tabel 46: Regressieanalyses voor de dragers van én het concept 'zelfevaluatie als beleidsdaad'

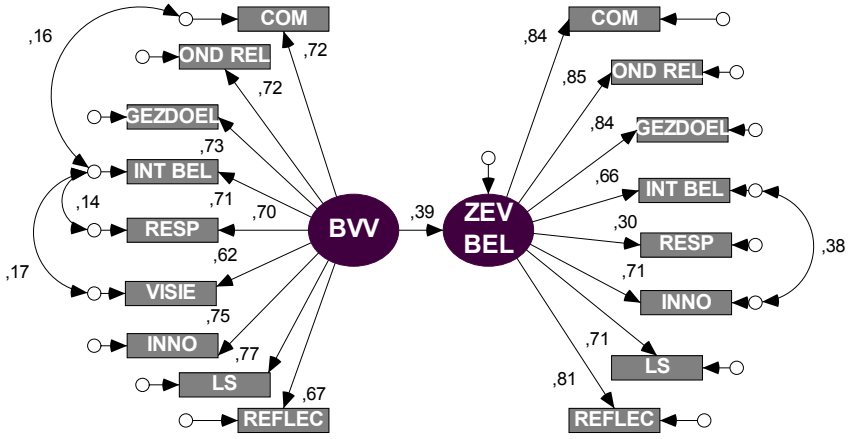
	Communicatie	Ondersteuning samenwerking	Gez. doelgerichtheid	Gedeeld leiderschap	Responsiviteit	Innovatie	Geïntegr. aanpak	Reflectie	Zelfevaluatie als beleidsdaad
Constante	3.94	3.95	3.82	3.95		3.95		3.79	3.65
SO (1)	-.49	-.41	-.49	-.27		-.31		-.47	-.30
SO(2-3)	-.38	-.40	-.53	-.31		-.30		-.51	-.34
OGO									
GO									
Wel ervaring									
Schoolgrootte		-.001							
Verklaarde variantie (R^2)	11.3	8.6	11.3	3.6	nvt	7.2	nvt	10.6	7.5
	De coëfficiënten zijn niet gestandaardiseerd. p is telkens kleiner dan 0.001								

Uit de resultaten hiervan (in Tabel 46) blijkt dat schoolleiders uit het basisonderwijs doorgaans meer dan hun collega's uit het secundair onderwijs aangeven dat hun school slaagt in het realiseren van de verschillende dragers van 'zelfevaluatie als beleidsdaad'. Op de mate van responsiviteit en de geïntegreerde aanpak van de zelfevaluatie na, scoren basisscholen telkens statistisch significant hoger dan secundaire scholen. Dat geldt

ook voor hun score op het concept 'zelfevaluatie als beleidsdaad'. Naast het onderwijsniveau is er slechts één onderzocht achtergrondkenmerk waarvan een effect uitgaat op een aspect van 'zelfevaluatie als beleidsdaad'. De mate van ondersteuning en samenwerking tijdens de GOK-zelfevaluatie blijkt groter naarmate de scholen kleiner zijn. Per 100 leerlingen extra daalt de score op de schaal 'professionele ondersteuning' gemiddeld met een tiende punt.

Eén jaar voor de eigenlijke zelfevaluatie zou plaatsvinden werden de schoolleiders van de GOK-scholen bevraagd over het beleidsvoerend vermogen van hun school. Hieronder worden de dragers van 'zelfevaluatie als beleidsdaad' en de dragers van 'beleidsvoerend vermogen' aan elkaar gekoppeld. In eerste instantie hebben we gekeken of we een statistisch significant verband vinden tussen de concepten 'beleidsvoerend vermogen' en 'zelfevaluatie als beleidsdaad' die schuil gaan achter de respectievelijke dragers. Om tot een passend model te komen moet dergelijke relatie inderdaad verondersteld worden (zie Figuur 23). De correlatie tussen beide concepten bedraagt 0.39 ($p < 0.001$). Het SEM-model heeft een behoorlijke fit (n=234; RMSEA: 0.04; pclose: 0.83; GFI: 0.93 ; CFI: 0.98). Onze hypothese dat een hoge score op beleidsvoerend vermogen samenhangt met een sterke invulling van de zelfevaluatie als beleidsdaad wordt hiermee bevestigd.

Figuur 23: Verband tussen 'beleidsvoerend vermogen' en 'zelfevaluatie als beleidsdaad'



We willen de link tussen beide concepten evenwel ook op het niveau van de verschillende dragers onderzoeken. Op die manier willen we in kaart brengen welke dragers de grootste impact hebben. We beschouwen daartoe elk van de dragers van de 'zelfevaluatie als beleidsdaad' even als een afhankelijke variabele in een klassieke multiple regressie. In Tabel 47 presenteren we de resultaten van de analyses waarin telkens (stapsgewijs) gezocht wordt naar het meest geschikte model om de variantie in deze dragers van 'zelfevaluatie als beleidsdaad' te verklaren. Elke rij representeert een regressieanalyse. Als voorspellers werden in eerste instantie de dragers van beleidsvoerend vermogen - zoals

deze een jaar voor de zelfevaluatie door de schoolleiders beoordeeld werden - opgenomen.

Tabel 47: Resultaten multiple regressieanalyses – dragers 'Zelfevaluatie als beleidsdaad'

DRAGERS ZELFEVALUATIE ALS BELEIDSDAAD	Cte	DRAGERS BELEIDSVOEREND VERMOGEN									R^2	Max CI
		COM	ONDSAM	VISIE	GEZDOEL	LS	RESP	INNO	INT BEL	REFLEC		
COM	2.21					.17	.15				.07	21
ONDSAM	1.75	.17			.22						.12	17
GEZDOEL	1.99				.21		.15				.09	15
GEDLS	2.99	.17									.03	13
RESP	0.63		.19								.04	14
INNO	2.45					.14				.22	.10	22
INT BEL	2.46		.13							.16	.06	19
REFLEC	1.88		.17							.20	.10	19
ZEVBELDAAD	1.74		.19		.22						.14	20

De coëfficiënten zijn gestandaardiseerd. p is telkens kleiner dan 0.01

Uit Tabel 47 blijkt meteen dat de verklaarde variantie (R^2) in elk van de regressieanalyses beperkt is. De waarde varieert voor de afzonderlijke dragers van 3% tot 12%. Voor het concept 'zelfevaluatie als beleidsdaad' gaat het om 14%. Dit betekent dat op basis van de verschillende dragers van beleidsvoerend vermogen blijkbaar maar een beperkt deel van de variantie in (de dragers van) de zelfevaluatie als beleidsdaad kan verklaard worden. De tabel geeft wel aan welke dragers hierin een rol spelen. 'Reflectief vermogen', 'ondersteuning en samenwerking' en 'gezamenlijke doelgerichtheid' komen het meest vaak voor en hebben de hoogste regressiecoëfficiënten. Daarnaast zijn echter ook 'doeltreffende communicatie', 'leiderschap' en 'responsief vermogen' statistisch significante voorspellers. We stellen ook vast dat concrete dragers van beleidsvoerend vermogen niet altijd de beste voorspellers zijn voor hun tegenhanger in 'zelfevaluatie als beleidsdaad'. We moeten echter voorzichtig zijn met de gerapporteerde parameters in Tabel 47. Een condition index (CI) die groter is dan 15 wijst op mogelijke problemen inzake multicollineariteit. Een waarde groter dan 30 wijst op ernstige problemen. We staan om deze reden hier ook stil bij de vraag of er sprake is van systematische verschillen in het tegemoet komen aan de dragers van 'zelfevaluatie als beleidsdaad' voor de drie clusters van scholen naargelang hun beleidsvoerend vermogen.

De vraag is of scholen met een laag, gemiddeld of hoog beleidsvoerend vermogen (volgens de perceptie van de schoolleider) verschillen op de dragers van 'zelfevaluatie als beleidsdaad'. Dat blijkt voor alle dragers het geval te zijn. De bij de variatieanalyse gebruikte F-waarde bedraagt minimaal 5.47 ((2,270); p = 0.005) en dit bij de schaal respon-

sief vermogen. Tabel 48 rapporteert over de post hoc variantieanalyses (Tukey) voor de drie clusters met betrekking tot het beleidsvoerend vermogen van scholen. Enkel verschillen met een significantieniveau kleiner dan 0.05 worden als statistisch significant verschillend gepresenteerd.

Tabel 48: Variatie in de dragers van 'zelfevaluatie als beleidsdaad' naar clusters van BVV

	Scholenclusters BVV				
Dragers van zelfevaluatie als beleidsdaad	Laag	Gemiddeld	Hoog	ANOVA (post hoc)	Eta²
Communicatief vermogen	3.64	3.76	4.17	Laag = Gem < Hoog	0,06
Ondersteuning & samenwerking	3.34	3.63	4.05	Laag < Gem < Hoog	0,09
Gezamenlijke doelgerichtheid	3.33	3.57	3.98	Laag < Gem < Hoog	0,08
Gedeeld leiderschap	3.59	3.80	4.10	Laag = Gem < Hoog	0,05
Responsief vermogen	1.78	1.88	2.30	Laag = Gem < Hoog	0,04
Innovatief vermogen	3.66	3.78	4.21	Laag = Gem < Hoog	0,09
Geïntegreerde aanpak	3.54	3.56	3.99	Laag = Gem < Hoog	0,06
Reflectief vermogen	3.31	3.55	4.04	Laag < Gem < Hoog	0,10
Zelfevaluatie als beleidsdaad	3.32	3.47	3.88	Laag < Gem < Hoog	0.09

Scholen die tot de cluster 'laag' of 'gemiddeld' beleidsvoerend vermogen behoren, scoren voor elke drager van 'zelfevaluatie als beleidsdaad' statistisch significant minder goed dan scholen die tot de cluster 'hoog' beleidsvoerend vermogen behoren. Voor drie dragers verschillen ook de scholen met een laag en gemiddeld beleidsvoerend vermogen van elkaar. Het gaat hier om de schalen 'ondersteuning en samenwerking', 'gezamenlijke doelgerichtheid' en 'reflectief vermogen' tijdens de zelfevaluatie. Ook voor het concept 'zelfevaluatie als beleidsdaad' zijn al de verschillen tussen de drie clusters statistisch significant. Eta kwadraat bedraagt maximaal 0.10[17] voor de dragers en is 0.09 voor het concept 'zelfevaluatie als beleidsdaad'. Scholen met een sterk beleidsvoerend vermogen scoren dus hoger op de dragers van en het concept 'zelfevaluatie als beleidsdaad'. Als we kijken naar de dragers geldt dit in hoofdzaak voor wat betreft het 'reflectief vermogen', de 'ondersteuning en samenwerking', het 'innovatief vermogen' en de 'gezamenlijke doelgerichtheid' tijdens de zelfevaluatie.

1.2.2 De dragers van 'zelfevaluatie als beleidsdaad' binnen 'zelfevaluatie binnenstebuiten'

In het bovenstaande werd enkel gerapporteerd over de perceptie van schoolleiders over de 'zelfevaluatie als beleidsdaad' in hun school. De gegevens die in het kader van 'Zelfe-

[17] De waarden 0.01, 0.06 en 0.14 worden geïnterpreteerd als klein, middelmatig en grote effectgroottes (Green & Salkind, 2004).

valuatie binnenstebuiten' verzameld werden, maken meerniveau-analyses mogelijk. Hieronder rapporteren we over deze analyses. We bekijken daarbij zowel de verschillende dragers als het achterliggende concept 'zelfevaluatie als beleidsdaad'.

a. BESCHRIJVENDE RESULTATEN

Omdat het concept 'zelfevaluatie als beleidsdaad' uit verschillende schalen bestaat, gaan we in de rapportage van de resultaten niet in op de afzonderlijke itemscores. We beperken ons tot de kengetallen over de schaalscores (zie Tabel 49). Daaruit blijkt in eerste instantie dat zelfevaluaties volgens de respondenten gemiddeld slechts licht positief scoren op de verschillende dragers. Net als bij de GOK-zelfevaluaties is het 'responsief vermogen' ook in deze gegevensverzameling de drager waarop scholen het minst goed scoren. Dit is de enige drager die gemiddeld lager scoort dan het middelpunt van de schaal. De andere dragers scoren gemiddeld genomen tussen 3.21 en 3.47. Het achterliggende concept 'zelfevaluatie als beleidsdaad' heeft een gemiddelde score van 3.23.

Tabel 49: Variatie in de dragers van 'zelfevaluatie als beleidsdaad'

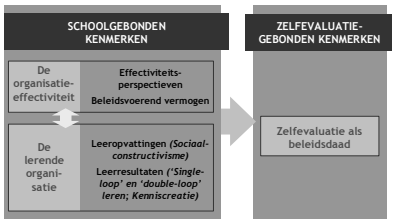
	Alle respondenten				Enkel schoolleiders			
Schaal	GEM	min	max	SD	GEM	min	max	SD
Communicatief vermogen	3.23	1	5	0.94	3.88	1.80	5.00	0.76
Ondersteuning & samenwerking	3.30	1	5	0.92	3.74	1.80	5.00	0.73
Gezamenlijke doelgerichtheid	3.25	1	5	0.86	3.59	1.00	5.00	0.74
Gedeeld leiderschap	3.33	1	5	0.96	3.79	1.00	5.00	0.80
Responsief vermogen	2.70	1	5	0.99	2.91	1.00	4.75	0.79
Innovatief vermogen	3.21	1	5	0.76	3.27	1.60	4.80	0.65
Geïntegreerde aanpak	3.47	1	5	0.85	3.61	1.33	5.00	0.81
Reflectief vermogen	3.42	1	5	0.75	3.55	1.50	5.00	0.72
Zelfevaluatie als beleidsdaad	3.23	1	5	0.67	3.54	1.99	4.67	0.54

In vergelijking met de resultaten van de GOK-zelfevaluatie (Tabel 34) liggen de scores voor de verschillende dragers lager[18]. De variatie in de antwoorden is dan weer duidelijk groter. De grote standaarddeviaties geven aan dat respondenten sterk van mening verschillen. Het is te verwachten dat de verschillen in gemiddelde scores én standaarddeviaties voor de twee gegevensverzamelingen in eerste instantie toe te schrijven zijn aan het feit dat binnen 'zelfevaluatie als beleidsdaad' naast schoolleiders ook andere teamleden bevraagd werden. Om toch een indicatie te geven voor mogelijke verschillen en gelijkenissen hebben we daarom rechts in Tabel 49 afzonderlijk de scores voor schoolleiders opgenomen. Daaruit blijkt dat de gemiddelde scores op de vergelijkbare schalen voor de

[18] De vergelijking over de gegevensverzamelingen heen kan niet gemaakt worden voor de schalen 'responsief vermogen' en 'innovatief vermogen'. Deze schalen zijn in beide gevallen immers niet identiek.

GOK-zelfevaluaties (Tabel 45) en de zelfevaluaties binnen 'zelfevaluatie binnenstebuiten' (Tabel 49) zeer dicht bij elkaar liggen. Het grootste schaalverschil vinden we terug bij 'ondersteuning & samenwerking' en bedraagt in absolute termen 0.14. Enkel voor deze schaal is het verschil statistisch significant[19], weliswaar ook hier met een eerder kleine effectgrootte (met name 0.19) (Green & Salkind, 2004). We kunnen dus niet concluderen dat de schoolleiders (sterk) significant verschillend scoorden over de twee gegevensverzamelingen heen.

b. MEERNIVEAU-ANALYSES



De rapportering over de meerniveau-analyses gebeurt in twee stappen. Eerst bestuderen we het achterliggende concept 'zelfevaluatie als beleidsdaad'. Daarna bekijken we de verschillende dragers die binnen het concept onderscheiden worden.

➢ *Focus op het concept 'zelfevaluatie als beleidsdaad'*

Tabel 50 geeft de resultaten weer van de meerniveau-analyses waarbij de samenhang tussen relevante school- en respondentkenmerken met het concept 'zelfevaluatie als beleidsdaad' onderzocht wordt. Het basismodel leert ons dat er zowel variatie is op het niveau van de school als op het niveau van de respondenten. Beide zijn respectievelijk goed voor 20% en 80% van de totale variantie. Het percentage dat verklaard wordt door het schoolniveau ligt dus hoger dan bij de houding ten aanzien van zelfevaluatie. Dat is niet onlogisch omdat de respondenten hier gevraagd werden om zich uit te spreken over éénzelfde concreet gebeuren in de school, namelijk de zelfevaluatie. Desalniettemin wordt evenwel nog steeds 80% van de variantie verklaard door verschillen tussen de respondenten. Eénzelfde zelfevaluatieproces leidt dus duidelijk niet tot dezelfde beoordeling door de verschillende personen die erbij betrokken werden.

Een significant deel van de variatie in de score van respondenten op het concept 'zelfevaluatie als beleidsdaad' kan verklaard worden door kenmerken van de school en van de respondent in kwestie. Op respondentniveau blijken voornamelijk de functie van de respondent, de perceptie van het proces en het product van de ondersteuning, en zowel de perceptie van de mate van diepgaand leren als van kenniscreatie een brutorol te spelen. Van al de statistisch significante effecten in het respondentendeel van het model hebben deze de hoogste regressiecoëfficiënten. Verder hebben ook het op individueel niveau gebruik maken ondersteuning, de score op de vier effectiviteitsperspectieven en de perceptie van het eigen leren elk op zich een statistisch significant effect. Wat de schoolkenmerken betreft blijkt er een statistisch significant verband tussen het concept 'zelfe-

[19] We toetsten de verschillen aan de hand van One-sample t tests. Daarbij gingen we na of de score van de schoolleiders binnen 'zelfevaluatie binnenstebuiten' statistisch significant afwijkt van de gemiddelde score van schoolleiders bij de GOK-zelfevaluaties.

valuatie als beleidsdaad' en het onderwijsniveau, het genieten van ondersteuning, de schoolgrootte en de schoolcultuur. De score van een respondent op 'zelfevaluatie als beleidsdaad' zal hoger zijn wanneer het om een school gaat die sterker dan gemiddeld scoort op de verschillende effectiviteitsperspectieven, wanneer het om een basisschool gaat en wanneer de school ondersteuning genoot tijdens de zelfevaluatie. Daarnaast blijkt de schoolgrootte een statistisch significant effect te hebben. Hoe groter het aantal leerkrachten in de school, des te minder respondenten de zelfevaluatie in hun school inschatten als sterk in termen van 'zelfevaluatie als beleidsdaad'. Net zoals een aantal andere schooleffecten verdwijnt het negatieve effect van de schoolgrootte echter in het nettomodel. Dat is wellicht te verklaren door het feit dat respondenten die in grote scholen werkzaam zijn meer geneigd zijn aan te geven dat hun school niet tegemoet komt aan de kenmerken van organisatie-effectiviteit en van de lerende organisatie.

Tabel 50: Meerniveau-analyses voor het concept 'zelfevaluatie als beleidsdaad'

	Basismodel	**Brutomodel**	**Nettomodel**
VAST	Schatting (SF)	Schatting (SF)	Schatting (SF)
Intercept	.087 (.050)		-.135 (.043)
Voorspellers: info over de school			
Basisonderwijs (vs. secundair onderwijs)		**.292 (.094)**	
Vrij gesubsidieerd onderwijs (vs. GO)		.089 (.100)	
Officieel gesubsidieerd onderwijs (vs. GO)		-.023 (.157)	
Geen ervaring met zelfevaluatie (vs. wel ervaring)		.042 (.101)	
Geen ondersteuning (vs. wel ondersteuning)		**-.238 (.101)**	
Schoolgrootte (aantal leerkrachten)		**-.139 (.055)**	
Schoolcultuur cluster 1 (vs. cluster2)		.095 (.181)	
Schoolcultuur cluster 3 (vs. cluster2)		**.484 (.088)**	**.201 (.064)**
Voorspellers: info over de (percepties van de) respondent			
Schoolleiding (vs. teamleden)		**.466 (.067)**	**.259 (.058)**
Wel ondersteuning (vs. geen ondersteuning)		**.345 (.038)**	**.203 (.035)**
Perceptie kwaliteit proces ondersteuning		**.411 (.028)**	NO
Perceptie kwaliteit resultaat ondersteuning		**.546 (.028)**	NO
Leeftijd		-.028 (.020)	
Perceptie van 'human relations'		**.264 (.019)**	
Perceptie van 'interne proces'		**.159 (.019)**	
Perceptie van 'open systeem'		**.275 (.019)**	**.074 (.020)**
Perceptie van 'rationele doel'		**.223 (.019)**	**.059 (.019)**
Perceptie van het eigen leren		**.135 (.020)**	
Perceptie van diepgaand leren in de school		**.486 (.020)**	**.297 (.021)**
Perceptie van mate van kenniscreatie in de school		**.462 (.020)**	**.267 (.021)**
VARIABEL			
Variantie op schoolniveau	.195 (.034)		.062 (.013)
Variantie op respondentenniveau	.774 (.023)		.483 (.016)
R^2 - Reductie van variantie op schoolniveau			38%

R^2 - Reductie van variantie op respondentenniveau			44%
MODEL FIT			
Deviatie	6197		3990
Aantal scholen	96		96
Aantal respondenten	2331		1839

Het nettomodel is een sterk model. In vergelijking met het nulmodel worden fouten in de voorspelling van de scores op het concept 'zelfevaluatie als beleidsdaad' gereduceerd met 38% op schoolniveau en met 44% op respondentniveau. Het nettomodel bevat 7 predictoren waarvan één schoolkenmerk, namelijk 'op schoolniveau beter dan gemiddeld scoren op de effectiviteitsperspectieven' (d.i. schoolcultuur cluster 3). Respondenten in zulke scholen scoren statistisch significant hoger. Dat geldt ook voor schoolleiders (in vergelijking met teamleden) en voor respondenten die een vorm van ondersteuning hebben genoten in het kader van zelfevaluatie. De mate waarin respondenten aangeven dat het 'open systeem'- en het 'rationele doel'-perspectief in hun school van belang zijn enerzijds en de perceptie van de mate van diepgaand leren en van kenniscreatie in de school anderzijds ronden het rijtje van statistisch significante effecten af. Gegeven dit nettomodel scoren respondenten die positief (negatief) scoren op de relevante dummyvariabelen en die op alle voorspellers 1 standaarddeviatie boven (onder) het gemiddelde scoren, gemiddeld 1.36 standaarddeviatie boven (onder) het gemiddelde op het concept 'zelfevaluatie als beleidsdaad'. Interacties werden geëxploreerd maar bleken niet tot een noemenswaardig beter model te leiden. Hetzelfde geldt voor het laten variëren van de regressiecoëfficiënten over scholen.

- *Focus op de dragers van 'zelfevaluatie als beleidsdaad'*

De analyses die uitgevoerd werden voor het concept 'zelfevaluatie als beleidsdaad' kunnen in principe voor elke drager afzonderlijk uitgevoerd worden. Wij zijn evenwel in hoofdzaak geïnteresseerd in de vraag naar de netto-effecten voor de verschillende dragers. We rapporteren in Tabel 51 om die reden ook enkel over het basismodel en nettomodel[20] van de meerniveau-analyses waarbij elke drager afzonderlijk bestudeerd wordt.

[20] Hoewel alle voorspellers uit Tabel 50 in de analyse opgenomen werden, bevat Tabel 51 enkel de regressiecoëfficiënten van de voorspellers met een statistisch significant effect bij minstens één van de dragers. Om de rapportering te vereenvoudigen werden ook de standaardfouten van de parameterschattingen niet opgenomen.

Tabel 51: Meerniveau-analyses voor de dragers van 'zelfevaluatie als beleidsdaad'

	COM	ONDSAM	GEZDOEL	GEDLS	RESP	INNO	INT BEL	REFLEC
VAST								
Intercept	-.20	-.20	-.27	-.25	-.24	-.16	-30	-.10
Voorspellers: info over de school								
Basisonderwijs (vs. secundair onderwijs)			.23		.19	.13		
Vrij gesubsidieerd onderwijs (vs. GO)							.15	
Schoolgrootte (aantal leerkrachten)								-.10
Schoolcultuur cluster 1 (vs. cluster2)	.26	.30		.29			.37	
Schoolcultuur cluster 3 (vs. cluster2)	.12	.24	.21	.18			.17	.18
Voorspellers: info over de (percepties van de) respondent								
Schoolleiding (vs. teamleden)	.49	.22	.31	.25				.25
Wel ondersteuning (vs. geen ondersteuning)	.31	.21	.17	.24	.27	.23	.20	
Perceptie van 'human relations'		.09					.08	
Perceptie van 'interne proces'		.05						
Perceptie van 'open systeem'						.18	-.10	.22
Perceptie van 'rationele doel'	.06	.05	.07	.06			.11	
Perceptie van diepgaand leren	.27	.29	.26	.32	.22	.21	.18	.19
Perceptie van mate van kenniscreatie	.20	.22	.24	.20	.21	.21	.32	.27
VARIABEL								
Basismodel: Variantie schoolniveau	.20	.27	.19	.20	.19	.11	.16	.15
Basismodel: Variantie respondentenniveau	.82	.75	.81	.81	.82	.83	.84	.80
Nettomodel:Variantie schoolniveau	.08	.08	.05	.06	.10	.02	.04	.02
Nettomodel:Variantie respondentenniveau	.59	.51	.60	.56	.70	.66	.63	.58
R^2 - Reductie van variantie op schoolniveau	29%	33%	27%	32%	15%	21%	26%	28%
R^2 - Reductie van variantie op respondentenniveau	34%	42%	35%	39%	21%	28%	33%	37%

Het aandeel van de variantie op schoolniveau verschilt sterk over de verschillende dragers. Zo wordt 26% van de variantie in 'ondersteuning en samenwerking' verklaard door verschillen tussen scholen. Voor 'innovatie' gaat het slechts om 11%. De nettomodellen in Tabel 51 verbeteren de voorspelling met 15% (voor responsiviteit) tot 33% (voor ondersteuning en samenwerking) in vergelijking met het basismodel. Voor het respondentniveau situeert R^2 zich tussen 21% en 42%.

In vergelijking met de resultaten van de meerniveau-analyse voor het concept 'zelfevaluatie als beleidsdaad' zien we voornamelijk gelijkenissen maar ook enkele verschillen. We focussen hier op de sterke effecten. Wat de schoolkenmerken betreft valt op dat basisscholen het statistisch significant beter doen dan secundaire scholen inzake 'gezamenlijke doelgerichtheid', 'responsiviteit' en 'innovatie'. Scholen met een groot aantal leerkrachten blijken minder goed te scoren op 'reflectie'. Ook de schoolcultuur speelt een rol als schoolkenmerk. In vergelijking met scholen die 'Sterk flexibel/zwak beheersmatig' zijn (cluster 2), scoren scholen die sterk zijn op alle perspectieven (cluster 3) en scholen

die 'gemiddeld beheersmatig/eerder zwak flexibel' zijn (cluster 1) doorgaans statistisch significant beter. Enkel voor 'responsiviteit' en 'innovatie' vinden we geen effect van de schoolcultuur, maar voor de andere dragers blijkt beheersmatigheid dus bij te dragen tot hogere scores. Wat de respondentenkenmerken betreft, blijken bij de afzonderlijke dragers doorgaans dezelfde voorspellers een rol te spelen als bij het concept 'zelfevaluatie als beleidsdaad'. Het gaat daarbij om de vraag of al dan niet deelgenomen werd aan een vorm van ondersteuning en om de functie die de respondent in de school inneemt. Schoolleiders scoren bijvoorbeeld .49 hoger op de drager 'communicatie tijdens de zelfevaluatie' dan overige teamleden. Toch zijn er ook dragers waar het effect van de functie zich niet manifesteert (met name responsiviteit, innovatie en geïntegreerd beleid). Voor alle dragers is er verder een sterk statistisch significant effect van de perceptie van diepgaand leren en van kenniscreatie in de school. Ook de perceptie van de schoolcultuur door de respondent hangt samen met de score op de dragers van 'zelfevaluatie als beleidsdaad'. Dat geldt voornamelijk voor de perceptie van de perspectieven 'open systeem' en 'rationele doel'.

Het moet gezegd dat de bovenstaande analyse voorbij gaat aan de samenhang die er tussen de dragers van 'zelfevaluatie als beleidsdaad' kan zijn. In het volgende exploreren we deze samenhang door gebruik te maken van een multivariate meerniveau-analyse[21]. Op die manier kunnen we in kaart brengen of de scores van scholen en respondenten op de verschillende dragers onderling samenhangen. De resultaten van deze analyse werden samengebracht in Tabel 52.

Tabel 52: Correlaties voor de dragers van 'zelfevaluatie als beleidsdaad'

	COM	ONDSAM	GEZDOEL	GEDLS	RESP	INNO	INT BEL	REFLEC	
COM		.68	.59	.68	.38	.37	.41	.37	Correlaties op respondentniveau
ONDSAM	.95		.58	.69	.41	.47	.46	.50	
GEZDOEL	.90	.94		.53	.39	.44	.39	.47	
GEDLS	.96	.99	.89		.42	.44	.48	.42	
RESP	.63	.66	.59	.75		.41	.40	.36	
INNO	.69	.87	.78	.83	.75		.49	.55	
INT BEL	.90	.78	.74	.86	.70	.57		.40	
REFLEC	.76	.90	.91	.84	.62	.93	.57		
	Correlaties op schoolniveau								

[21] Het is in principe mogelijk om in een multivariate meerniveau-analyse tegelijkertijd het effect van de verschillende voorspellers op de verschillende dragers te analyseren. De veelheid van afhankelijke en onafhankelijke variabelen maakt echter dat dit tot te complexe modellen zou leiden.

Onder de grijze vlakken is aangegeven in welke mate de schoolscores op de ene drager samenhangen met de schoolscores op de andere dragers. Uit de correlaties blijkt duidelijk dat de schoolscores sterk samenhangen. De schoolscore op 'communicatie' correleert bijvoorbeeld .90 of hoger met 'ondersteuning en samenwerking', 'gezamenlijke doelgerichtheid', 'gedeeld leiderschap' en 'geïntegreerd beleid'. De score van de school op 'communicatie' geeft dus meteen ook informatie over de score van de school op deze andere dragers. 'Reflectief vermogen' geeft op schoolniveau meteen ook informatie over 'innovatief vermogen' (R=.93). Dat maakt dat enkel responsief vermogen overblijft. Het kan daarom volstaan om in toekomstige analyses op schoolniveau enkel communicatie, responsief vermogen en reflectief vermogen als voorspellers op te nemen. Informatie over deze dragers geeft immers meteen ook informatie over de andere dragers.

Boven de grijze vakken zijn de correlaties op respondentniveau opgenomen. Deze liggen beduidend lager dan deze op schoolniveau. Toch blijkt de score op één drager nog steeds heel wat informatie te geven over de score op andere dragers. De correlaties variëren van .36 tot .69. Er blijkt zich hier een gelijkaardige patroon voor de te doen. De sterkste correlaties worden ook hier gevonden tussen op 'communicatie' en 'ondersteuning en samenwerking', 'gezamenlijke doelgerichtheid', 'gedeeld leiderschap'. 'Reflectief vermogen' correleert ook hier – zij het in mindere mate - met 'innovatief vermogen' (R=.55). Op respondentniveau resteren naast 'communicatie' en 'reflectief vermogen' dan nog 'responsief vermogen' en 'geïntegreerde aanpak'.

1.2.3 Conclusies bij de vooropgestelde onderzoekshypothesen

We concluderen uit de analyses met betrekking tot de 'zelfevaluatie als beleidsdaad' dat de respondenten zich gemiddeld genomen slechts licht positief uitspreken over de mate waarin de dragers van 'zelfevaluatie als beleidsdaad' van toepassing zijn op de zelfevaluatie die in hun school plaatsvond (cf. hypothese 3a). Dit geldt evenwel niet voor de drager 'responsief vermogen', die doorgaans niet blijkt gerealiseerd te zijn. Er zijn voor alle dragers evenwel grote verschillen tussen scholen en respondenten. Eénzelfde zelfevaluatieproces leidt duidelijk niet tot dezelfde beoordeling door de verschillende personen die erbij betrokken werden. Voornamelijk de functie van respondenten blijkt een sterke impact te hebben. Schoolleiders spreken zich veel positiever uit over de dragers dan andere teamleden. Andere achtergrondkenmerken van scholen en respondenten blijken niet meer statistisch significant wanneer voor de schoolgebonden kenmerken gecontroleerd wordt. Uit de analyses bleek dan ook dat de invulling van 'zelfevaluatie als beleidsdaad' duidelijk niet los staat van de organisatie-effectiviteit en het lerende karakter van de school. Hypothese 3b, die verwachtingen in die zin uitsprak, wordt dus bevestigd.

1.3 Het verloop van zelfevaluaties: zelfevaluatie als onderzoeksdaad

Zelfevaluatie wordt in deze studie niet alleen als een beleidsdaad benaderd maar ook als een onderzoeksdaad. In de voorbereidende analyses beschreven we dat zelfevaluaties niet alleen gescoord kunnen worden op het concept 'zelfevaluatie als onderzoeksdaad'

maar ook op drie dragers van dit concept. Er wordt hier zodus eenzelfde logica toegepast als bij de 'zelfevaluatie als beleidsdaad'. De drie dragers zijn inhoud, validiteit en betrouwbaarheid. Hieronder geven we eerst de beschrijvende resultaten voor de dragers en het achterliggende concept. Daarna rapporteren we over de uitgevoerde meerniveau-analyses.

1.3.1 Beschrijvende resultaten

De operationalisering van 'zelfevaluatie als onderzoeksdaad' en van de drie dragers van het concept berust op 15 items. Tabel 53 zet de beschrijvende resultaten voor de items en de schalen op een rij. Het is goed om weten dat deze items enkel voorgelegd werden aan respondenten die in hun school betrokken waren bij de bevraagde fase van de zelfevaluatie. Sommige items werden dan ook door een grote groep respondenten gescoord (bijvoorbeeld 1930 respondenten bij 'de vragenlijst was gebruiksvriendelijk') terwijl andere enkel door een minderheid ingevuld werden (bijvoorbeeld 280 bij 'De vragenlijst sloot goed aan bij de doelen van de zelfevaluatie').

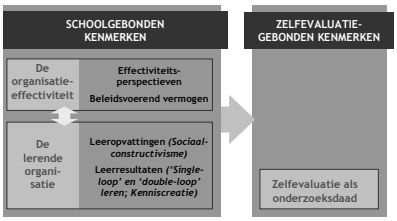
Tabel 53: Frequentieverdeling (%), gemiddelde score en standaarddeviatie voor de items en schalen m.b.t. 'zelfevaluatie als onderzoeksdaad'

Inhoud	**1**	**2**	**3**	**4**	**5**	**Gem.**	**SD**
Ik had een duidelijk beeld van de bedoelde inhoud van de zelfevaluatie.	2	11	22	52	13	3.64	.91
De keuze van de te evalueren thema's werd goed beargumenteerd.	1	6	12	63	18	3.89	.81
De inhoud van de zelfevaluatie werd gesitueerd in een ruimer kader.	1	4	12	63	20	4.00	.72
De geëvalueerde thema's waren relevant.	1	4	19	60	16	3.88	.73
De te evalueren thema's werden zeer nauwkeurig omschreven.	1	6	16	58	19	3.90	.80
Schaalscore 'Inhoud'						*3.72*	*.76*
Validiteit							
De gegevens geven een goed beeld van de bevraagde thema's.	1	3	16	55	25	4.02	.76
De gegevensverzameling leverde representatieve informatie op.	1	3	17	58	21	3.97	.75
De vragenlijst verzamelde precies informatie over wat wij wilden weten.	1	5	29	51	14	3.72	.80
De vragenlijst sloot goed aan bij de doelen van de zelfevaluatie.	1	2	16	58	23	4.01	.74
De gegevens werden correct geïnterpreteerd.	1	4	18	58	19	3.91	.75
Er was een passend kader om te bepalen wat goed en wat slecht was.	1	7	21	54	17	3.79	.83
Schaalscore 'Validiteit'						*3.86*	*.66*
Betrouwbaarheid							
De anonimiteit van de vragenlijst was voldoende gewaarborgd.	7	6	10	29	48	4.06	1.2
De gegevensverzameling werd zorgvuldig uitgevoerd.	1	4	13	57	25	4.04	.76
Ik heb de vragenlijst eerlijk ingevuld.	1	1	3	27	68	4.60	.70
De vragenlijst was gebruiksvriendelijk.	2	8	17	52	21	3.83	.90
Schaalscore 'Betrouwbaarheid'						***4.1***	***.64***
Schaalscore concept 'zelfevaluatie als onderzoeksdaad'						***3.90***	***.61***

De respondenten geven in grote lijnen aan dat de zelfevaluatie in hun school behoorlijk scoorde op de dragers van 'zelfevaluatie als onderzoeksdaad'. De gemiddelde scores voor de schalen liggen duidelijk boven het middenpunt van de beoordelingschaal. Het grootste

percentage respondenten dat het (eerder) niet eens was met de uitspraken vinden we bij (1) 'Ik had een duidelijk beeld van de bedoelde inhoud van de zelfevaluatie', (2) 'De anonimiteit van de vragenlijst was voldoende gewaarborgd' en (3) De vragenlijst was gebruiksvriendelijk. Dit zijn meteen ook de stellingen met de grootste standaarddeviaties. Er is dus behoorlijk wat variantie in de items en in de schaalscores. In de meerniveau-analyses analyseren we deze variantie.

1.3.2 Meerniveau-analyses



In de rapportering over de meerniveau-analyses onderscheiden we twee stappen. Eerst bestuderen we het achterliggende concept 'zelfevaluatie als onderzoeksdaad'. Daarna bekijken we de verschillende dragers die binnen het concept onderscheiden worden.

a. FOCUS OP HET CONCEPT 'ZELFEVALUATIE ALS ONDERZOEKSDAAD'

We geven eerst de resultaten weer van de meerniveau-analyses waarbij de samenhang tussen relevante school- en respondentkenmerken en het concept 'zelfevaluatie als onderzoeksdaad' onderzocht wordt. Een overzicht van het basis-, bruto- en nettomodel is opgenomen in Tabel 54.

We stellen vast dat slechts 7.5% van de variantie in de score van respondenten op het concept 'zelfevaluatie als onderzoeksdaad' toe te schrijven is aan verschillen tussen scholen. Het overgrote deel van de variantie situeert zich dus op respondentenniveau. Dat is merkwaardig gezien de respondenten gevraagd werden zich uit te spreken over concrete aspecten van de zelfevaluatie in hun school. In een gelijkaardig geval, met name voor het concept 'zelfevaluatie als beleidsdaad', bedroeg de variantie op schoolniveau nog 20% van het totaal. Respondenten binnen scholen verschillen dus sterk inzake hun perceptie van het concept 'zelfevaluatie als onderzoeksdaad'.

De score op 'zelfevaluatie als onderzoeksdaad' blijkt samen te hangen met de cultuur in scholen. In de brutomodellen speelt de schoolcultuur zowel op school als op respondentniveau een statistisch significante rol. Scholen én respondenten die hoger scoren op de vier effectiviteitsperspectieven scoren ook hoger op het concept 'zelfevaluatie als onderzoeksdaad'. Op schoolniveau vinden we verder geen statistisch significante effecten. Op respondentniveau is dat wel het geval. Respondenten met een leidinggevende taak in de school scoren ook hier hoger dan de andere teamleden. Ook het belang van ondersteuning wordt bevestigd. Respondenten die deelnamen aan ondersteuning of zich positief uitspreken over het proces of de resultaten van deze ondersteuning scoren statistisch significant hoger dan zij die dat niet doen. Tenslotte hebben ook de perceptie van het eigen leren, van diepgaand leren en van kenniscreatie in de school een statistisch significant brutoeffect.

Tabel 54: Meerniveau-analyses voor het concept 'zelfevaluatie als onderzoeksdaad'

	basismodel	Bruto model	Netto model
VAST	Schatting (SF)	Schatting (SF)	Schatting (SF)
Intercept	.055 (.035)		-.009 (.033)
Voorspellers: info over de school			
Basisonderwijs (vs. secundair onderwijs)		.110 (.068)	
Vrij gesubsidieerd onderwijs (vs. GO)		.070 (.070)	
Officieel gesubsidieerd onderwijs (vs. GO)		-.004 (.112)	
Geen ervaring met zelfevaluatie (vs. wel ervaring)		.039 (.071)	
Geen ondersteuning (vs. wel ondersteuning)		-.053 (.072)	
Schoolgrootte (aantal leerkrachten)		-.059 (.038)	
Schoolcultuur cluster 1 (vs. cluster2)		-.012 (.128)	
Schoolcultuur cluster 3 (vs. cluster2)		**.237 (.066)**	
Voorspellers: info over de (percepties van de) respondent			
Schoolleiding (vs. teamleden)		**.569 (.072)**	**.453 (.072)**
Wel ondersteuning (vs. geen ondersteuning)		**.137 (.042)**	
Perceptie kwaliteit proces ondersteuning		**.383 (.033)**	NO
Perceptie kwaliteit resultaat ondersteuning		**.485 (.030)**	NO
Leeftijd		-.013 (.022)	
Perceptie van 'human relations'		**.158 (.021)**	
Perceptie van 'interne proces'		**.190 (.021)**	**.070 (.026)**
Perceptie van 'open systeem'		**.134 (.021)**	
Perceptie van 'rationele doel'		**.178 (.021)**	**.099 (.026)**
Perceptie van het eigen leren		**.198 (.022)**	**.111 (.024)**
Perceptie van diepgaand leren in de school		**.257 (.024)**	**.136 (.028)**
Perceptie van mate van kenniscreatie in de school		**.244 (.022)**	**.115 (.028)**
VARIABEL			
Variantie op schoolniveau	.068 (.016)		.051 (.014)
Variantie op respondentenniveau	.839 (.027)		.702 (.026)
R^2 - Reductie van variantie op schoolniveau			16%
R^2 - Reductie van variantie op respondentniveau			17%
MODEL FIT			
Deviatie	5534		3892
Aantal scholen	96		96
Aantal respondenten	2044		1543

In vergelijking met het basismodel vermindert het nettomodel de fouten in de voorspelling van de scores op school- en op respondentniveau met respectievelijk 16% en 17%. Het model is dus niet uitermate sterk in het verklaren van verschillen op beide niveaus. In vergelijking met 'zelfevaluatie als beleidsdaad' hangt 'zelfevaluatie als onderzoeksdaad' dus in mindere mate samen met de onderzochte voorspellers. De belangrijkste voorspeller is echter ook hier de functie die respondenten in de school waarnemen. Verder blijken de perceptie van het eigen leren, de perceptie van diepgaand leren en de perceptie van kenniscreatie ook na controle voor de andere voorspellers statistisch significante effecten

te zijn. Tenslotte vinden we dat respondenten die van mening zijn dat er in hun school veel aandacht is voor het 'rationele doel'- en het 'interne proces'-perspectief statistisch significant beter scoren op het concept 'zelfevaluatie als onderzoeksdaad'. Deze perspectieven richten zich op beheersbaarheid. Een samenhang met systematiek doorheen het zelfevaluatieproces is dus niet onverwacht.

b. FOCUS OP DE DRAGERS VAN 'ZELFEVALUATIE ALS ONDERZOEKSDAAD'

In de volgende alinea's rapporteren we over analyses waarbij de verschillende dragers van 'zelfevaluatie als onderzoeksdaad' afzonderlijk onderzocht worden. Tabel 55 bevat informatie over het basismodel en het nettomodel van de drie uitgevoerde meerniveau-analyses. Enkel de voorspellers met een statistisch significant effect zijn in de tabel opgenomen.

Tabel 55: Meerniveau-analyses voor de dragers van 'zelfevaluatie als onderzoeksdaad'

	Inhoud	Validiteit	Betrouw-baarheid
VAST			
Intercept	-.10	-.23	-.28
Voorspellers: info over de school			
Vrij gesubsidieerd onderwijs (vs. GO)			.25
Geen ervaring met zelfevaluatie (vs. wel ervaring)			.16
Schoolgrootte (aantal leerkrachten)	-.09		
Schoolcultuur cluster 1 (vs. cluster2)		.39	
Schoolcultuur cluster 3 (vs. cluster2)		.16	
Voorspellers: info over de (percepties van de) respondent			
Schoolleiding (vs. teamleden)	.46	.38	.45
Wel ondersteuning (vs. geen ondersteuning)	.11		
Leeftijd		-.07	
Perceptie van 'interne proces'			.08
Perceptie van 'rationele doel'	.07	.19	.07
Perceptie van het eigen leren	.08	.12	.12
Perceptie van diepgaand leren	.11	.21	.13
Perceptie van mate van kenniscreatie		.09	.07
VARIABEL			
Basismodel: Variantie schoolniveau	.08	.09	.09
Basismodel: Variantie respondentenniveau	.92	.91	.91
Nettomodel:Variantie schoolniveau	.04	.04	.05
Nettomodel:Variantie respondentenniveau	.78	.75	.77
R^2 - Reductie van variantie op schoolniveau	15%	18%	16%
R^2 - Reductie van variantie op respondentenniveau	18%	21%	18%

De verdeling van de variantie over de twee niveaus is gelijkaardig voor de drie dragers. Ook de voorspellende kracht van de drie nettomodellen ligt dicht bij elkaar. In beide opzichten zijn de drie modellen overigens te vergelijken met het model voor het achter-

liggende concept 'zelfevaluatie als onderzoeksdaad'. Er zijn echter ook duidelijke verschillen. Zo blijken een aantal schoolkenmerken wel een statistisch significant effect te hebben op de dragers terwijl dat voor het concept niet het geval was. Opmerkelijk is dat de betrouwbaarheid hoger is in scholen zonder ervaring inzake zelfevaluatie en in scholen van het vrij gesubsidieerd onderwijs (ten minste in vergelijking met scholen van het gemeenschapsonderwijs). Grote scholen scoren minder goed op de drager 'inhoud'. 'Inhoud' en 'betrouwbaarheid' blijken niet samen te hangen met de schoolcultuur op schoolniveau. Het effect van de schoolcultuur op het concept 'zelfevaluatie als onderzoeksdaad' blijkt dus te moeten worden toegeschreven aan de impact ervan op de drager 'validiteit'. Op respondentniveau zien we eerder een bevestiging van wat voor het achterliggende concept vastgesteld werd. Dat betekent een sterke impact van de positie die men in de school inneemt, een impact van het 'interne proces' en 'rationele doel', en een impact van de perceptie van het eigen leren, diepgaand leren in de school en kenniscreatie in de school. Nieuw is wel dat ondersteuning op respondentniveau samen gaat met betere scores op 'Inhoud' en de dat de 'validiteit' van de zelfevaluatie lager ligt volgens oudere respondenten.

We hebben tenslotte ook onderzocht hoe de drie dragers van 'zelfevaluatie als onderzoeksdaad' onderling samenhangen op school- en respondentniveau. Deze gegevens werden samengebracht in Tabel 56. Op schoolniveau blijken 'inhoud' en 'validiteit' zeer sterk samen te hangen (R=.92). De score op 'validiteit' geeft daarnaast ook behoorlijk wat informatie over de score van een school op 'betrouwbaarheid' (R=.69). Ook op respondentniveau blijkt 'validiteit' het meest sterk met de twee overige dragers samen te hangen. Met het oog op selectie van relevante voorspellers van de kwaliteit van zelfevaluatie lijkt de schaal 'validiteit' dus de meest krachtige voorspeller.

Tabel 56: Correlaties tussen de dragers van 'zelfevaluatie als onderzoeksdaad'

	Inhoud	Validiteit	Betrouwbaarheid	
Inhoud		.67	.46	Respondent niveau
Validiteit	.92		.59	
Betrouwbaarheid	.41	.69		
	schoolniveau			

1.3.3 Conclusies bij de vooropgestelde onderzoekshypothesen

We concluderen in de eerste plaats dat de respondenten zich gemiddeld genomen eerder positief uitspreken over de mate waarin de dragers van 'zelfevaluatie als onderzoeksdaad' van toepassing zijn op de zelfevaluatie die in hun school plaatsvond (cf. hypothese 3a). De gemiddelde scores liggen enigszins hoger dan bij 'zelfevaluatie als beleidsdaad'. Het is echter niet onwaarschijnlijk dat dit toe te schrijven is aan de uitval van respondenten die niet of slechts in beperkte mate bij de zelfevaluatie betrokken waren. Het overgrote deel van de variantie in de scores op het concept en de dragers situeert zich op respondentenniveau. De verschillen tussen scholen zijn beperkter wat 'zelfevaluatie als

onderzoeksdaad' betreft dan wat 'zelfevaluatie als beleidsdaad' betreft. Respondenten verschillen binnen scholen verhoudingsgewijs echter sterker van mening over de dragers van 'zelfevaluatie als onderzoeksdaad' dan van 'zelfevaluatie als beleidsdaad'. We stelden verder vast dat de nettomodellen die we presenteerden niet uitermate sterk zijn in het verklaren van verschillen tussen scholen en respondenten. Het concept en dragers van 'zelfevaluatie als onderzoeksdaad' laten zich niet zo goed voorspellen dan het concept en de dragers van 'zelfevaluatie als beleidsdaad'. In de voorspelling blijkt ook hier de functie van de respondent evenwel de belangrijkste voorspeller. De perceptie van schoolgebonden kenmerken (de 'beheersgerichte' effectiviteitsperspectieven en de lerende organisatie) hangt eveneens statistisch significant samen met het oordeel over de invulling van de onderzoeksmatige aspecten van de zelfevaluatie. Onze onderzoekshypothese hieromtrent (3b) wordt dus bevestigd. De regressiecoëfficiënten zijn evenwel kleiner dan voor de andere zelfevaluatiegebonden kenmerken.

1.4 Relaties tussen de zelfevaluatiegebonden kenmerken

Hierboven bestudeerden we achtereenvolgens de drie zelfevaluatiegebonden kenmerken uit het theoretisch kader. Tot dusver hebben we niet gerapporteerd over de samenhang tussen de 'houding van teamleden ten aanzien van zelfevaluatie', 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad'. Onze hypothese is dat een hoge score op het ene concept samenhangt met een hoge score op de andere concepten. Deze hypothese wordt in deze paragraaf op twee manieren getoetst. Eerst bekijken we de samenhang op 'concept'-niveau, daarna richten we ons op de dragers van de verschillende concepten.

Tabel 57 rapporteert over een multivariate meerniveau-analyse waarin de samenhang tussen de drie zelfevaluatiegebonden concepten in kaart wordt gebracht. Die samenhang blijkt voornamelijk op schoolniveau sterk te zijn. De correlaties variëren van .65 tot .83. Scholen die hoog scoren op één concept zullen dus doorgaans ook hoog scoren op de andere concepten. Dat geldt ook, maar in mindere mate, voor het respondentniveau. De correlaties bedragen hier .35 tot .48. Zowel voor het school- als voor het respondentniveau blijkt dat de score op het concept 'zelfevaluatie als beleidsdaad' het meeste informatie geeft over de beide andere concepten.

Tabel 57: Correlaties tussen de zelfevaluatiegebonden concepten

	Houding	ZEV als beleidsdaad	ZEV als onderzoeksdaad	
Houding		.46	.35	Respondentniveau
Zelfevaluatie als beleidsdaad	.71		.48	
Zelfevaluatie als onderzoeksdaad	.65	.83		
	Schoolniveau			

Om te achterhalen of het patroon in de samenhang tussen de concepten ook opgaat voor de verschillende dragers voerden we een gelijkaardige maar meer diepgaande analyse uit. Deze bevat de houding ten aanzien van zelfevaluatie, de dragers van 'zelfevaluatie als beleidsdaad'[22] en de dragers van 'zelfevaluatie als onderzoeksdaad'. Tabel 58 rapporteert de resultaten van deze analyse. Vermits de verbanden binnen de concepten reeds besproken werden, focussen we in de volgende conclusies op verbanden over de concepten heen[23]. We stellen vast dat de 'houding ten aanzien van zelfevaluatie' op schoolniveau sterk samenhangt met het 'reflectief' en 'innovatief vermogen' tijdens de zelfevaluatie en met de 'validiteit van de zelfevaluatie'. Sterke samenhangen zijn er op schoolniveau ook tussen 'doeltreffende communicatie' en twee dragers van zelfevaluatie als onderzoeksdaad, namelijk 'inhoud' en 'validiteit'. Ook 'reflectief' vermogen hangt sterk samen met deze twee dragers. Op respondentniveau vinden we geen correlaties tussen de dragers (over de concepten heen) die groter zijn dan .40. Er tekenen zich echter dezelfde trends af als op schoolniveau.

Tabel 58: Correlaties voor de dragers van 'zelfevaluatie als beleidsdaad'

	HOUDING	COM	RESP	INNO	REFLEC	INHOUD	BETROUW	VALIDITEIT	
HOUDING		.34	.26	.31	.38	.34	.24	.40	Respondent-niveau
COM	.52		.34	.37	.37	.45	.30	.42	
RESP	.52	.61		.40	.36	.29	.15	.31	
INNO	.72	.68	.75		.55	.29	.21	.35	
REFLEC	.89	.75	.60	.92		.32	.21	.34	
INHOUD	.62	.91	.51	.50	.78		.46	.66	
BETROUW	.46	.54	.54	.42	.47	.44		.58	
VALIDITEIT	.79	.85	.59	.67	.93	.92	.74		
	Schoolniveau								

Onze onderzoekshypothese (cf. hypothese 4) over de samenhang tussen de concepten enerzijds en tussen de dragers van de concepten anderzijds wordt bevestigd. We concluderen dan ook dat er een positieve samenhang is tussen de 'houding ten aanzien zelfevaluatie', (de dragers van) 'zelfevaluatie als beleidsdaad' en (de dragers van) 'zelfevaluatie

[22] Om het aantal te schatten parameters te beperken hebben we een aantal dragers niet opgenomen. We baseren ons voor deze selectie op de eerdere analyses. De score van een school/respondent op 'doeltreffende communicatie' geeft immers meteen ook informatie over de score op 'ondersteuning en samenwerking', 'gezamenlijke doelgerichtheid', 'gedeeld leiderschap' en 'geïntegreerd beleid'. De correlaties van 'doeltreffende communicatie' zijn - tenminste op schoolniveau - dus ook indicatief voor de andere dragers.

[23] Daarmee bedoelen we dat we niet opnieuw focussen op de onderlinge verbanden van dragers binnen één 'concept' maar wel op de verbanden tussen de dragers van verschillende concepten.

als onderzoeksdaad'. Deze samenhang is sterker op schoolniveau dan op respondentniveau.

1.5 De kwaliteit van zelfevaluaties

De beschikbaarheid van de drie gegevensbestanden maakt het mogelijk ook de kwaliteit van zelfevaluaties op verschillende manieren te beschrijven. Wij maken daarbij een onderscheid tussen de kwaliteit van de GOK-zelfevaluaties en de kwaliteit van de zelfevaluaties die binnen 'zelfevaluatie binnenstebuiten' bestudeerd werden.

1.5.1 De kwaliteit van GOK-zelfevaluaties

Wij zijn in staat om vanuit twee invalshoeken uitspraken te doen over de kwaliteit van GOK-zelfevaluaties. In eerste instantie bekijken we de perceptie van schoolleiders van de bevraagde GOK-scholen. Daarna kijken we naar de kwaliteit van dezelfde zelfevaluaties volgens de onderwijsinspectie.

a. KWALITEIT DOOR DE OGEN VAN DE SCHOOLLEIDING

De resultaten van de GOK-zelfevaluatie werden bij schoolleiders aan de hand van zeven stellingen in kaart gebracht[24]. 'Kwaliteit' is hier dus het realiseren van nastrevenswaardig geachte resultaten. Tabel 59 geeft een overzicht van de stellingen, geordend op basis van de gemiddelde mate waarin de respondenten (n=468) aangeven dat de resultaten op hun school van toepassing zijn.

Tabel 59: Frequentieverdeling (%), gemiddelde score en standaarddeviatie voor de resultaten van de GOK-zelfevaluaties

De resultaten van de GOK-zelfevaluatie...	1	2	3	4	5	GEM	SD	Eta²
…geven een goed zicht op de sterke punten in onze schoolwerking.	0	3	9	58	30	4.14	.71	.02
…leverden strategieën op om GOK-problemen verder aan te pakken.	1	4	21	54	20	3.88	.80	NS
…geven een goed zicht op de problemen bij onze leerlingen.	3	8	17	50	22	3.80	.96	NS
…geven aanknopingspunten om onze GOK-strategieën te herformuleren.	1	9	20	50	20	3.79	.91	.04
…geven een goed zicht op tekorten in het handelen van onze lkn.	1	7	23	56	14	3.75	.81	.04
…geven aanknopingspunten om de GOK-doelstellingen te herformuleren.	4	9	21	48	19	3.69	.99	.08
…zetten ons ertoe aan in de toekomst meer aan zelfevaluatie te doen.	3	6	31	47	14	3.63	.87	.02
Schaalscore						3.84	.58	

De gemiddelde score op de bevraagde zelfevaluatieresultaten is telkens hoog. Bijna 90% van de respondenten geeft aan dat de zelfevaluatie hen een (eerder) goed beeld heeft gegeven van de sterke punten in hun schoolwerking. Ongeveer 70% van de deelnemende scholen spreekt van strategieën om de GOK-problemen verder aan te pakken, van aan-

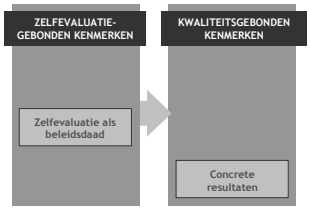
[24] De stellingen werden gescoord aan de hand van de volgende antwoordschaal: (1) volledig oneens, (2) eerder oneens, (3) noch oneens/noch eens, (4) eerder eens en (5) volledig eens.

knopingspunten om de GOK-strategieën te herformuleren en van een goed zicht op de problemen bij de leerlingen en op de tekorten in het handelen van de leerkrachten. In het algemeen spreken de schoolleiders zich dus positief uit over de resultaten van de zelfevaluatie. De gemiddelde schaalscore is 3.84. Toch reageert - uitgezonderd bij de eerste stelling - ook één vijfde van de respondenten neutraal op de stelling en geeft ongeveer 10% aan het (eerder) niet eens te zijn met de stellingen. Het gemiddelde beeld verbergt dus wel een aanzienlijk variantie binnen de totale groep deelnemende scholen. Dat blijkt ook uit de standaarddeviaties. Er zijn dus wel degelijk verschillen in de perceptie van de zelfevaluatieresultaten in de deelnemende scholen. We onderzoeken eerst of deze verschillen samenhangen met achtergrondkenmerken van de school, daarna belichten we de impact van de dragers van 'zelfevaluatie als beleidsdaad' en van het beleidsvoerend vermogen van de school.

- *Impact van achtergrondkenmerken van de school*

Het net, de schoolgrootte en het aantal GOK-uren blijken niet samen te hangen met de verschillen in de perceptie van de resultaten, het onderwijsniveau wel. Voor een aantal resultaten scoren basisscholen hoger dan secundaire scholen. De meest rechtse kolom van Tabel 59 bevat de eta^2-waarden[25] van variantieanalyses naar verschillen tussen de onderwijsniveaus (i.e. doelpopulaties) voor de afzonderlijke items. De verschillen zijn klein tot middelmatig. Het grootste verschil bedraagt .57 op de gehanteerde beoordelingsschaal en vinden we bij de stelling 'De resultaten van de GOK-zelfevaluatie geven aanknopingspunten om de GOK-doelstellingen te herformuleren' ($F=16.1$; df (2,458): $p<0.001$; $eta^2= 0.08$). Ook bij de gemiddelde schaalscore speelt enkel het onderwijsniveau een statistisch significante rol ($F=6.76$; df (2,466): $p=0.001$). Basisscholen scoren gemiddeld 3.94 terwijl secundaire scholen 1ste graad enerzijds en 2de en 3de graad anderzijds respectievelijk 3.73 en 3.77 scoren ($eta^2 = .03$).

- *Impact van de dragers van 'zelfevaluatie als beleidsdaad'*



Naast de impact van achtergrondkenmerken werd ook geanalyseerd of de dragers van 'zelfevaluatie als beleidsdaad' goede voorspellers zijn van de kwaliteit van de GOK-zelfevaluaties. Tabel 60 rapporteert over de uitgevoerde regressieanalyse. Daaruit blijkt allereerst dat de impact van het onderwijsniveau verdwijnt wanneer gecontroleerd wordt voor de dragers van 'zelfevaluatie als beleidsdaad'. Van de variantie in de resultaten van de zelfevaluatie kan 41% op basis van het regressiemodel verklaard worden ($R^2=0.41$).

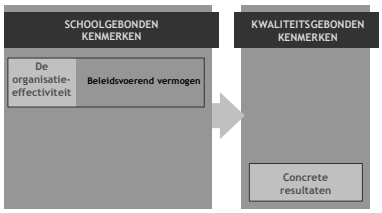
[25] De waarden 0.01, 0.06 en 0.14 worden geïnterpreteerd als klein, middelmatig en grote effectgroottes (Green & Salkind, 2004).

Tabel 60: Voorspellers van de kwaliteit van de GOK-zelfevaluatie

Voorspellers	De resultaten van de zelfevaluatie	
	B	*B*
Constante	1.17	
Achtergrondkenmerken van de school (geen significante predictoren)		
Aspecten van de 'zelfevaluatie als beleidsdaad'		
Doeltreffende communicatie	.17	.19
Innovatief vermogen	.54	.51
R^2	0.41	

Het innovatieve karakter van de zelfevaluatie speelt een zeer belangrijke rol (*B=0.51)*. De niet-gestandaardiseerde regressiecoëfficiënt bedraagt ruim 0.5. Dat betekent dat scholen per punt verschil op de schaal 'innovatief karakter van de zelfevaluatie' gemiddeld een half punt verschillen op de indicator 'resultaten van de zelfevaluatie' (beide lopen van 1 tot 5). Het communicatief karakter van de zelfevaluatie heeft een gestandaardiseerde regressiecoëfficiënt van 0.19. Het regressiemodel bevat dus slechts twee voorspellers. Deze vaststelling is mede te verklaren door de sterke correlaties die we hoger reeds rapporteerden tussen groepen van dragers van 'zelfevaluatie als beleidsdaad'. Dat maakt dat 'communicatie' in grote lijnen inwisselbaar is voor dragers als 'ondersteuning' en 'gezamenlijke doelgerichtheid' en 'geïntegreerde aanpak' voor 'reflectief vermogen tijdens de zelfevaluatie'. We concluderen daarom met de vaststelling dat de dragers van 'zelfevaluatie als beleidsdaad' in sterke mate samenhangen met de resultaten van de zelfevaluatie.

- *Impact van beleidsvoerend vermogen*



We ronden de beschrijving van de 'kwaliteit door de ogen van de schoolleiders' af door na te gaan of er een samenhang is met het beleidsvoerend vermogen van scholen. We vergelijken daartoe de scholen met respectievelijk een laag, gemiddeld en hoog beleidsvoerend vermogen. De variantie-analyse waarover we in Tabel 61 rapporteren, toont aan dat scholen met een hoog beleidsvoerend vermogen statistisch significant hoger scoren dan scholen die gemiddeld of laag scoren. Het effect van het beleidsvoerend vermogen is inhoudelijk noemenswaardig (eta² = .08) maar klein in vergelijking met het effect van de dragers van 'zelfevaluatie als beleidsdaad'.

Tabel 61: Variatie in de resultaten van de GOK-zelfevaluatie voor scholenclusters van BVV

	Scholenclusters BVV			ANOVA (post hoc)	Eta²
	Laag	Gemiddeld	Hoog		
De resultaten van de zelfevaluatie	3.74	3.80	4.23	F=14.5 - df(2,320) p< 0.001 (Laag = Gem < Hoog)	0.08

Aan de hand van een multiple regressieanalyse onderzochten we ook welke dragers van beleidsvoerend vermogen de variantie in de schaalscore kunnen verklaren. Tabel 62 vat de resultaten samen.

Tabel 62: Resultaten multiple regressieanalyse – Kwaliteit door de ogen van schoolleiders

De resultaten van de zelfevaluatie	Cte	DRAGERS VAN BELEIDSVOEREND VERMOGEN									R^2	Max CI
		Com	Ond Rel	Visie	Gez Doel	Ls	Resp	Inno	Int Bel	Re-flec		
	2.49		0.18				0.18				0.09	19
		p is telkens kleiner dan 0.01										

Op basis van de beoordeling van het eigen beleidsvoerend vermogen kan 9% van de variantie in de schaalscore met de zelfrapportage over de resultaten van de zelfevaluatie verklaard worden. In het regressiemodel blijken zowel 'ondersteunende relaties' als 'responsief vermogen' een rol te spelen. De gestandaardiseerde regressiecoëfficiënten bedraagt bij beide 0.18. Ook hier is de impact duidelijk kleiner dan bij de dragers van 'zelfevaluatie als beleidsdaad'.

b. KWALITEIT DOOR DE OGEN VAN DE GOK-INSPECTIE

De onderwijsinspectie hanteerde in haar analysekader vier indicatoren om de kwaliteit van de GOK-zelfevaluatie te beoordelen. Tabel 63 bevat de resultaten voor de vier indicatoren voor al de doorgelichte scholen. Daaruit blijkt dat 80 tot bijna 85% van de scholen 'voldoende tot goed' of zeer goed' scoort op de criteria. Slechts een zeer kleine minderheid van de scholen scoort 'onvoldoende'.

Tabel 63: Beoordeling van de GOK-zelfevaluatie door inspectie

Kwaliteitsindicatoren GOK-inspectie	Onvol-doende	Aanvaard-baar	Voldoende tot goed	Zeer goed
Indicator 1: Voorstellen tot verbetering en bijsturing	2.5%	17.6%	55.6%	24.4%
Indicator 2: Gegevensverzameling	1.1%	14.9%	54.0%	29.9%
Indicator 3: Gegevensanalyse	1.9%	16.5%	50.9%	30.7%
Indicator 4: Betrokkenheid	1.5%	14.3%	53.4%	30.7%

Deze resultaten wekken de indruk dat de GOK-zelfevaluatie behoorlijk kwaliteitsvol zijn verlopen. Uiteraard speelt de vraag naar de gehanteerde standaarden hier een cruciale rol (we komen hierop terug in de algemene conclusies bij dit deel). In het volgende zijn we evenwel in de eerste plaats geïnteresseerd in de variatie in de beoordelingen. We onderzoeken eerst of de score van GOK-scholen op deze indicatoren samenhangt met achtergrondkenmerken van de school. Daarna onderzoeken we ook hier de impact van de dragers van 'zelfevaluatie als beleidsdaad' en van het beleidsvoerend vermogen van scholen.

➢ *Impact van achtergrondkenmerken van de school*

Voor elk van de vier indicatoren kunnen we nagaan of de verhoudingen over de verschillende beoordelingsniveaus dezelfde zijn over verschillende groepen scholen (denk aan schoolgrootte, onderwijsniveau en onderwijsnet). Om de daarbij horende kruistabellen te analyseren doen we een beroep op chikwadraattoetsen. Daaruit blijkt dat het verschil tussen kleine, middelgrote en grote scholen niet statistisch significant is. Het onderwijsniveau en het onderwijsnet hebben wel een impact.

Tabel 64 geeft de kruistabel weer voor de indicator 'Voorstellen verbetering/bijsturing' en het onderwijsniveau waartoe de school behoort. De verdeling van scholen over de kwaliteitsniveaus blijkt in de verschillende onderwijsniveaus sterk te verschillen van de verdeling van deze van de totale groep. De chikwadraattoets wijst hier dan ook in de richting van statistisch significante verschillen tussen scholen (Pearson X^2 (6, n = 527) = 36.16, p = 0.001). Basisscholen krijgen duidelijk betere beoordelingen op deze kwaliteitsindicator. Slechts 16.7% van de basisscholen scoort 'onvoldoende' of 'aanvaardbaar'. In secundaire scholen ligt dit aantal op 34.5% voor de 1ste graad en 46,3% voor de 2de en 3de graad.

Tabel 64: Kruistabel onderwijsniveau en 'Voorstellen verbetering / bijsturing'

	Basis		1ste graad SO		2de & 3de graad SO		Totaal[26]	
	Aantal	%	Aantal	%	Aantal	%	Aantal	%
0nvoldoende	6	2.0%	7	3.9%	5	9.3%	18	3.4%
Aanvaardbaar	43	14.7%	55	30.6%	20	37.0%	118	22.4%
Voldoende tot goed	159	54.3%	78	43.3%	14	25.9%	251	47.6%
Zeer goed	85	29.0%	40	22.2%	15	27.8%	140	26.6%
totaal	293	100%	180	100%	54	100%	527	100%

Het effect van het onderwijsniveau blijkt zich ook bij de andere indicator te manifesteren (zie Tabel 65). Slechts 12,3% van de basisscholen scoort 'onvoldoende' of 'aanvaardbaar' op de beoordeling over de 'gegevensverzameling' bij de zelfevaluatie. In het secundair onderwijs ligt dit percentage op 34% voor de scholen met een 1ste graad en op 37% voor de scholen met een 2de en 3de graad (Pearson X^2 (6, n = 527) = 45.8, p < 0.001). Een zelfde trend zien we in de resultaten voor de indicator 'analyse van de gegevens'. Er zijn minder basisscholen dan secundaire scholen die 'onvoldoende' of slechts 'aanvaardbaar' scoren (Pearson X^2 (6, n = 527) = 13.6, p = 0.03). Het meest uitgesproken verschil tussen de onderwijsniveaus vinden we echter bij de indicator 'betrokkenheid bij de zelfevaluatie' (Pearson X^2 (6, n = 527) = 67.1, p < 0.001). Naar het oordeel van de onderwijsinspectie is de betrokkenheid in basisscholen slechts in minder dan 10% van de gevallen 'aan-

[26] Er wordt hier enkel gerapporteerd over de scores van scholen die deel uitmaakten van de steekproef bij het onderzoek naar de GOK-zelfevaluaties. Dat verklaart de licht verschillende totaalscores in Tabel 63 waarin al de door de inspectie doorgelichte scholen opgenomen zijn.

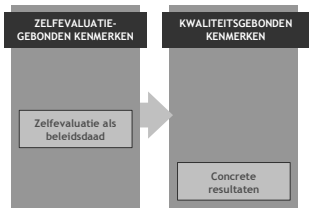
vaardbaar' of 'onvoldoende' geweest. In het secundair onderwijs loopt dit aantal op tot 32.2% voor de scholen met een 1ste graad en tot 37.1% voor de scholen met een 2de en 3de graad. Vooral het aantal scholen dat 'zeer goed' scoort ligt verhoudingsgewijs lager in het secundair onderwijs.

Tabel 65: Kruistabel onderwijsniveau en overige indicatoren

	Gegevensverzameling			Gegevensanalyse			Betrokkenheid		
	BaO	1ste GR SO	2de 3de GR SO	BaO	1ste GR SO	2de 3de GR SO	BaO	1ste GR SO	2de 3de GR SO
0nvoldoende	0.7%	3.3%	0.0%	1.7%	3.3%	3.7%	0.7%	2.8%	1.9%
Aanvaardbaar	11.6%	30.6%	37.0%	14.3%	25.6%	22.2%	8.9%	29.4%	35.2%
Voldoende tot goed	48.1%	42.8%	42.6%	48.1%	45.0%	38.9%	53.9%	55.6%	50.0%
Zeer goed	39.6%	23.3%	20.4%	35.8%	26.1%	35.2%	36.5%	12.2%	13.0%

Het onderwijsnet speelt bij één indicator een statistisch significante rol, met name bij 'de betrokkenheid van het team' (Pearson X^2 (6, n = 527) = 15.1, p = 0.02). In het vrij gesubsidieerd onderwijs (VGO) ligt het aantal scholen dat 'onvoldoende' (2.8%) of 'aanvaardbaar' (19%) scoort iets hoger dan in de andere netten (18% voor de beide categorieën). Deze verschillen zijn echter zeer beperkt. We concluderen dan ook dat van de bestudeerde achtergrondkenmerken voornamelijk het onderwijsniveau relevant is.

- *Impact van de dragers van 'zelfevaluatie als beleidsdaad'*



We kunnen de dragers van 'zelfevaluatie als beleidsdaad' ook koppelen aan de beoordeling van de GOK-inspectie. Omdat de beoordeling door inspectie op een ordinale schaal gemeten werd, vergt dit evenwel een andere methodologie. Wij hanteren daarom een discriminantanalyse. Discriminantanalyse wordt gebruikt om individuele scholen op basis van een aantal indicatoren in groepen in te delen. Het komt er in wezen op neer dat we op basis van de dragers van 'zelfevaluatie als beleidsdaad' de beoordeling door de inspectie proberen te voorspellen. De indicatoren zijn hier dus de dragers van 'zelfevaluatie als beleidsdaad'. We onderscheiden twee groepen in de kwaliteitsbeoordeling door de inspectie, met name wat de indicator 'Voorstellen tot verbetering en bijsturing' betreft. Deze groepen zijn enerzijds de scholen die 'onvoldoende' of 'aanvaardbaar' scoren (deze groep labellen we 'lage kwaliteit') en anderzijds de scholen die 'voldoende tot goed' of 'zeer goed' scoren (dit is de groep 'hoge kwaliteit').

We concluderen op basis van de discriminantanalyse dat er een zwak maar statistisch significant verband is tussen de perceptie van schoolleiders omtrent de dragers van 'zelfevaluatie als beleidsdaad' en de beoordeling op de indicator 'Voorstellen tot verbetering en bijsturing' door de GOK-inspectie. Wilks' Lambda (0.95) leunt sterk aan bij de waarde 1 maar is statistisch significant (p < 0.001). Er is met andere woorden een statistisch

significant maar inhoudelijk klein verschil tussen de tussen de groepen 'lage' en 'hoge' kwaliteit wat de indicatoren betreft. De canonische correlatiecoëfficiënt bedraagt slechts 0.22. Enkel de doeltreffendheid van de communicatie tijdens de zelfevaluatie blijkt een statistisch significante rol te spelen in de bijhorende discriminantfunctie. De voorspellende kracht is evenwel beperkt. De kappa coëfficiënt bedraagt slechts 0.02. Een kappawaarde van 1 betekent een perfecte voorspelling, een waarde van 0 betekent dat de voorspelling niet beter is dan op basis van toeval. De informatie van schoolleiders over de 'zelfevaluatie als beleidsdaad' voegt dus zeer weinig toe aan het toeval bij het voorspellen van de beoordeling in de GOK-inspectie.

Hierboven werd uitgegaan van slechts één indicator (Voorstellen tot verbetering en bijsturing). De kwaliteit van de GOK-zelfevaluaties werd door de onderwijsinspectie echter aan de hand van vier indicatoren in kaart gebracht. De gelijkenissen en verschillen op deze vier indicatoren maakten dat we hoger tot 3 clusters van scholen kwamen. Wij hebben op basis van de dragers van 'zelfevaluatie als beleidsdaad' eveneens geprobeerd om door middel van discriminantanalyse te voorspellen tot welke cluster een school behoort. Wilks' Lambda (0.93) ligt voor deze analyse iets lager ($p < 0.001$). De kappa-coëfficiënt ligt iets hoger maar bedraagt nog altijd maar 0.10. Als kwaliteit ruimer wordt ingevuld dan een goede beoordeling op 'Voorstellen tot verbetering en bijsturing', is de voorspellende kracht van de dragers van 'zelfevaluatie als beleidsdaad' dus groter. Dat is niet onlogisch omdat de andere indicatoren van de inspectie zich eerder op het proces dan op de resultaten van de zelfevaluatie richten. De enige drager die een statistisch significante rol speelt in deze discriminantfunctie is gezamenlijke doelgerichtheid.

Alvorens onze verkenning van het verband tussen de kwaliteitsbeoordeling door de GOK-inspectie en de dragers van 'zelfevaluatie als beleidsdaad' af te ronden, gebruiken we nog even een andere invalshoek. Aan de hand van variantieanalyses gaan we na of de clusters van scholen met betrekking tot de GOK-inspectie verschillen wat de dragers van 'zelfevaluatie als beleidsdaad' betreft. Tabel 66 bevat enkel de dragers waarbij ook bij post hoc analyses (Tukey) statistisch significante verschillen ($p < 0.04$) gevonden worden.

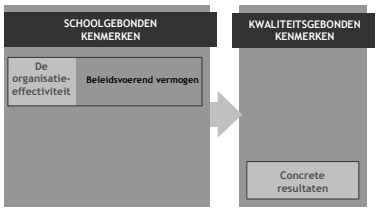
Tabel 66: Variatie in de dragers van 'zelfevaluatie als beleidsdaad' naar clusters van de GOK-inspectie

	Scholenclusters GOK-inspectie			ANOVA (post hoc)	Eta²
	Laag	Gemid-deld	Hoog		
Communicatief vermogen	3.65	3.83	3.98	Laag < Hoog	.04
Ondersteuning en samenwerking	3.49	3.65	3.71	Laag < Hoog	.02
Gezamenlijke doelgerichtheid	3.42	3.65	3.74	Laag < Gem = Hoog	.03
Reflectief vermogen	3.45	3.60	3.66	Laag < Hoog	.02

Een blik op de gemiddelde scores voor de verschillende clusters leert dat de verschillen – hoewel statistisch significant – eerder klein zijn. Eta kwadraat bedraagt maximaal 0.04. Opnieuw blijken de dragers van 'zelfevaluatie als beleidsdaad' wel samen te hangen met de kwaliteit van zelfevaluaties door de ogen van de GOK-inspectie. De effectgroottes zijn

doorgaans echter (zeer) beperkt. Dat kan zowel toe te schrijven zijn aan het werkelijk ontbreken van samenhang tussen de concepten als aan het ontbreken aan samenhang tussen de oordelen van schoolleiders en inspectieleden.

- *Impact van beleidsvoerend vermogen*



In lijn met de vorige analyses gaan we ook na of we op basis van de perceptie van schoolleiders over het beleidsvoerend vermogen van hun school kunnen voorspellen welke beoordeling de GOK-inspectie toekende aan de GOK-zelfevaluatie in de school. Opnieuw werd in eerste instantie een discriminantanalyse gebruikt om de beoordeling door de GOK-inspectie te voorspellen. De indicatoren zijn hier de dragers van beleidsvoerend vermogen. De groepen zijn opnieuw de scholen die 'onvoldoende of aanvaardbaar' scoren (lage kwaliteit) enerzijds en de scholen die 'voldoende tot goed of zeer goed' scoren (hoge kwaliteit) anderzijds.

De discriminantanalyse leert dat het niet mogelijk is de verwachte voorspelling hard te maken. Wilks' Lambda (0.96) leunt aan bij de waarde 1, wat impliceert dat er tussen de twee groepen geen statistisch significante verschillen zijn wat de indicatoren betreft. De canonische correlatiecoëfficiënt bedraagt 0.19. Door deze te kwadrateren bekom je het percentage variantie in de eerste discriminantfunctie dat kan verklaard worden door het verschil tussen de twee groepen. Dit is dus slechts 3.6%. We concluderen dan ook dat er geen sterk verband is tussen de perceptie van schoolleiders van hun eigen beleidsvoerend vermogen en de beoordeling door de GOK-inspectie van de kwaliteit van de zelfevaluatie die in de school plaatsvond.

Om af te ronden hanteren we ook hier de invalshoek van de clusters. Op basis van de beschikbare data kwamen we tot groepen scholen met een laag, gemiddeld en hoog beleidsvoerend vermogen enerzijds en een zwakke, gemiddelde en sterke GOK-beoordeling anderzijds (zie deel 5). Als er een sterk verband is tussen de bestudeerde concepten verwachten we dat scholen in beide clusters in overeenstemmende categorieën zullen zitten. Tabel 67 bevat in dat opzicht een aantal verhelderende percentages. We leren daaruit bijvoorbeeld dat er van de 100 scholen met een laag beleidsvoerend vermogen (laag BVV) 35% de beoordeling van verhoudingsgewijs zwak krijgen van de GOK-inspectie, 38% de beoordeling van verhoudingsgewijs gemiddeld en 26% verhoudingsgewijs sterk. Scholen met een hoog beleidsvoerend vermogen hebben evenveel kans om de beoordeling zwak als sterk te krijgen (32.7%). Ook deze resultaten geven dus aan dat er geen sprake is van een verband tussen beide elementen ($p = 0.85$ bij de chikwadraattoets).

Tabel 67: Vergelijking van scholen over clusters van BVV en GOK-inspectie

	GOK-inspectie			
Dragers BVV	zwakke	Gemiddeld	sterk	Totaal
Laag BVV	35.4%	38.3%	26.3%	100%
Gemiddeld BVV	35.1%	33.2%	31.7%	100%
Hoog BVV	32.7%	34.5%	32.8%	100%
Totaal	34.9%	34.4%	30.7%	100%

De perceptie van het beleidsvoerend vermogen van scholen door schoolleiders blijkt dus niet samen te hangen met de beoordeling die de inspectie aan hun school toekent op de indicatoren van kwaliteitsvolle zelfevaluatie. Ook hier moeten we ons ervan bewust zijn dat dit ook kan toe te schrijven zijn aan de verschillende standpunten in het beoordelen van beide concepten. Bij beleidsvoerend vermogen gaat het immers om schoolleiders, bij de kwaliteitsindicatoren gaat het om inspecteurs. In de volgende paragraaf gaan we hier dieper op in.

c. *KWALITEIT : HET BELANG VAN DE OGEN*

Voor ruim 400 GOK-scholen zijn wij in staat de uitspraken van hun schoolleiders over het verloop en de kwaliteit van de zelfevaluatie te vergelijken met deze van externe beoordelaars (c.q. de inspectie). Hoewel de gegevensverzameling dit niet als doel had, kunnen deze analyses een licht werpen op de vraag of zelfrapportage door schoolleiders al dan niet een betrouwbaar beeld geeft over de bestudeerde fenomenen.

In eerste instantie kijken we naar de correlaties tussen de verschillende beoordelingen (zie Tabel 68). Al de correlaties linksboven en rechtsonder zijn statistisch significant ($p < 0.01$). Dit zijn de correlaties tussen de beoordelingen van de inspectie enerzijds en van schoolleiders anderzijds. Over de beoordelaars heen blijkt geen enkel verband statistische significant. Er is – volgens deze benadering - dus geen statistisch significant verband tussen de beoordeling van de onderwijsinspectie en van schoolleiders.

Tabel 68: Correlaties tussen de verschillende indicatoren

		Oordeel Inspectie				Oordeel schoolleiders			
		Ind 1	Ind 2	Ind 3	Ind 4	Ind 1	Ind 2	Ind 3	Ind 4
Oordeel inspectie	Ind 1	1							
	Ind 2	0.58	1						
	Ind 3	0.70	0.70	1					
	Ind 4	0.57	0.59	0.58	1				
Oordeel schoolleiders	Ind 1	0.08	0.02	0.02	0.14	1			
	Ind 2	0.13	0.08	0.11	0.14	0.31	1		
	Ind 3	0.09	0.09	0.04	0.13	0.32	0.50	1	
	Ind 4	0.12	0.10	0.09	0.14	0.34	0.44	0.41	1

We analyseren de vier indicatoren ook meer diepgaand door telkens te kijken of verschillen in de beoordeling door de inspectie ook samenhangen met verschillen in beoordeling door de respondenten. Dat doen we telkens aan de hand van variatieanalyses.

Wanneer we kijken naar informatie over 'de resultaten van de zelfevaluatie' (zie Tabel 69) dan blijkt dat er wel statistisch significante verschillen zijn. Deze zijn evenwel beperkt. Enkel schoolleiders van scholen die volgens de onderwijsinspectie 'onvoldoende' scoren op deze indicator scoren verschillend van de andere schoolleiders. Zij geven meer dan de andere schoolleiders aan dat hun resultaten niet/minder goed zijn. Andere verschillen zijn er niet. Schoolleiders van scholen waarvan de inspectie oordeelde dat het resultaat van hun zelfevaluatie 'aanvaardbaar was' scoren dus in de zelfrapportage niet (statistisch) significant lager dan schoolleiders die de beoordeling 'zeer goed' ontvingen.

Tabel 69: Zelfrapportage of externe beoordeling? Resultaten van de zelfevaluatie

Indicator: kwaliteit van de zelfevaluatie		Aantal scholen	Zelfrapportage Scholen	ANOVA
Beoordeling inspectie	0nvoldoende	13	3.35	(F= 3.26; df(3,422) p= 0.02) Eta² = .02 **Post-hoc (Tukey):** Onvoldoende < rest
	Aanvaardbaar	86	3.84	
	Voldoende tot goed	212	3.85	
	Zeer goed	115	3.87	
	Totaal	426	3.84	

Dezelfde analyse werd uitgevoerd voor de indicatoren omtrent gegevensverzameling en -analyse. Voor deze indicatoren worden geen statistisch significante verschillen gevonden. Ook deze indicatoren wijzen er dus op dat er geen één op één verband bestaat tussen de beoordeling door externen en zelfrapportage door schoolleiders. Voor de gegevensverzameling geldt zelfs dat schoolleiders die een 'onvoldoende' krijgen van de onderwijsinspectie zich het meest positief uitspreken. Door het kleine aantal (n=5) is het verschil echter niet statistisch significant. Wat de gegevensanalyse betreft zien we wel de verwachte trend maar de verschillen zijn te klein om statistisch significant te zijn. Tenslotte kijken we ook naar de beoordeling van de betrokkenheid tijdens de zelfevaluatie. Het gaat hier om één van de dragers van de 'zelfevaluatie als beleidsdaad' (namelijk gedeeld leiderschap), beoordeeld door de schoolleiders enerzijds en het oordeel van de inspectie anderzijds. De vaststellingen in Tabel 70 zijn gelijkaardig aan deze in Tabel 69. In het kleine aantal scholen dat de beoordeling onvoldoende kreeg, geeft ook de schoolleider meer dan in de andere scholen aan dat de betrokkenheid bij de zelfevaluatie beperkt was. Opnieuw zijn er geen statistisch significante verschillen tussen de categorieën 'aanvaardbaar', 'voldoende tot goed' en 'zeer goed'.

Tabel 70: Zelfrapportage of externe beoordeling? Betrokkenheid bij de zelfevaluatie

Indicator: kwaliteit van de zelfevaluatie		Aantal scho-len	Zelfrapportage Scholen	ANOVA
Beoordeling inspectie	0nvoldoende	5	3,02	(F= 3.9; df(3,402) p= 0.01) Eta² = .03 **Post-hoc (Tukey):** Onvoldoende < rest
	Aanvaardbaar	68	3,71	
	Voldoende tot goed	220	3,80	
	Zeer goed	113	3,94	
	Totaal	406	3,82	

Op basis van de gepresenteerde gegevens moeten we vaststellen dat de beoordeling van een schoolleider en van een extern beoordelaar niet (sterk) samenhangen. Het verschil tussen het inspectieoordeel 'aanvaardbaar' en 'zeer goed' vertaalt zich nergens in een verschillende beoordeling door de schoolleiders zelf. We moeten hierbij weliswaar goed in het achterhoofd houden dat de gegevens betrekking hebben op dezelfde thema's, maar op een andere manier verzameld werden. Het gebrek aan methodologische consistentie kan dus een rol spelen. Daardoor kunnen we ook niet uitmaken of de perceptie van schoolleiders al dan niet positiever is dan deze van de onderwijsinspectie. Daarvoor zouden beide dezelfde instrumenten moeten gebruiken. We leren wel dat voorzichtigheid aan de orde is. De door schoolleiders gepercipieerde kwaliteit van een (het verloop van) zelfevaluatie geeft dus niet noodzakelijk een betrouwbaar beeld van de feitelijke kwaliteit (als we het oordeel van de inspectie zo zouden omschrijven). Dat noodzaakt tot een grote voorzichtigheid om de validiteit van interpretaties te bewaken en zet ons er toe aan om waar mogelijk niet alleen op de beoordeling van schoolleiders te berusten.

1.5.2 Kwaliteit in het kader van 'zelfevaluatie binnenstebuiten'

De kwaliteit van zelfevaluaties werd in de bevraging 'zelfevaluatie binnenstebuiten' op drie verschillende manieren benaderd. De schooleigen benadering is gebaseerd op de vraag of zelfevaluaties aan de verwachtingen die de school/respondent zelf vooropstelt tegemoet komen. De schooloverstijgende benadering brengt in kaart in welke mate zelfevaluaties resulteren in actiepunten enerzijds en in een aantal concrete resultaten anderzijds. Zowel in de beschrijvende resultaten als in de meerniveau-analyses komen de resultaten van elke invalshoek aan bod.

a. BESCHRIJVENDE RESULTATEN

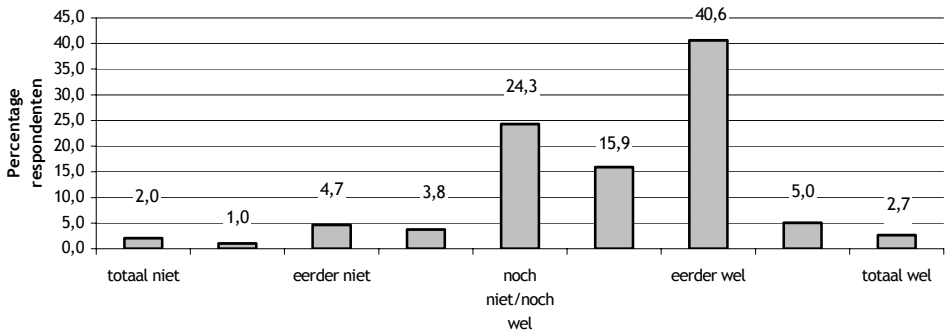
De beschrijvende resultaten richten zich achtereenvolgens op de schooleigen benadering, op het resulteren in passende actiepunten en op het resulteren in concrete resultaten.

- *De schooleigen benadering van kwaliteit*

In het theoretisch kader omschreven wij kwaliteit als het voldoen aan verwachtingen. In de vragenlijst gaven respondenten aan in welke mate de zelfevaluatie in hun school slaagde in het bereiken van de twee verwachtingen die zijzelf als het meest belangrijk

naar voren schoven. Het gemiddelde voor beide is de schooleigen indicator voor de kwaliteit van de zelfevaluatie. Figuur 24 geeft de percentages respondenten (n=2524) weer die een bepaalde score hebben op deze indicator. De beoordelingsschaal liep van totaal niet (1) tot totaal wel (5).

Figuur 24: Scores van de respondenten op de schooleigen kwaliteitsindicator



Ruim 48 % van de respondenten geeft aan dat de zelfevaluatie in hun school er 'eerder wel' tot 'totaal wel' in slaagde om de twee belangrijkste verwachtingen waar te maken. 16% is beperkt positief. Een kwart geeft aan dat dit 'noch niet/nog wel' het geval was. Het aantal respondenten dat 'eerder niet' of 'totaal niet' scoorde is dan ook eerder beperkt. 11,3% scoort onder het middenpunt van de schaal. De indicator heeft een gemiddelde score van 3.49 en de standaarddeviatie bedraagt 0.76. De respondenten spreken zich dus gematigd positief uit over de resultaten van de zelfevaluatie in termen van het realiseren van de eigen verwachtingen.

- *De schooloverstijgende benadering: resulteren in passende actiepunten*

Resulteren in passende actiepunten omschreven we in het theoretisch kader als het komen tot actiepunten die specifiek, meetbaar, aanvaardbaar, realistisch en tijdsgebonden zijn. Tabel 71 bevat de beschrijvende resultaten voor de schaal 'resulteren in actiepunten' en voor de verschillende items van deze schaal.

Tabel 71: Frequentieverdeling (%), gemiddelde score en standaarddeviatie voor de items m.b.t. 'Resulteren in actiepunten'

De zelfevaluatie resulteerde:	1	2	3	4	5	Gem.	SD
... in duidelijk omschreven actiepunten.	4	12	23	50	11	3.52	0.97
... in actiepunten waarvan de realisatie achteraf na te gaan is.	5	15	30	44	7	3.33	0.97
... in actiepunten waarvoor een groot draagvlak bestond.	5	14	34	42	5	3.30	0.94
... in realistische actiepunten.	3	11	26	52	8	3.51	0.90
... in actiepunten met een duidelijke tijdsplanning.	5	18	33	37	7	3.21	1.00
Schaalscore 'Resulteren in actiepunten'						3.37	.85

De gemiddelde schaalscore bedraagt 3.37. Gemiddeld spreekt de groep respondenten (n = 2145) zich dus gematigd positief uit over de mate waarin de zelfevaluatie in de eigen school tot gepaste actiepunten heeft geleid. Dat komt onder meer omdat het aantal respondenten dat het 'volledig eens' is met de verschillende stellingen eerder beperkt is. 70% tot 80% van de respondenten scoort 'noch niet/nog wel' of 'eerder wel'. Een noemenswaardige groep respondenten is het verder (eerder) niet eens met de stellingen. De variatie in antwoorden is dan ook groot, zowel voor de schaalscore als voor de afzonderlijk items.

- *De schooloverstijgende benadering: resulteren in concrete resultaten*

Van zelfevaluaties wordt niet alleen verwacht dat ze resulteren in passende actiepunten. De schooloverstijgende benadering van de kwaliteit van zelfevaluatie houdt immers ook in dat zelfevaluaties tot een aantal wenselijk geachte resultaten moeten leiden. In Tabel 72 wordt gerapporteerd in welke mate dat volgens de bevraagde respondenten het geval was.

Tabel 72: Frequentieverdeling (%), gemiddelde score en standaarddeviatie voor de items m.b.t. 'Concrete resultaten'

De zelfevaluatie:	1	2	3	4	5	Gem.	SD
... stimuleerde het bespreken van opvattingen over goed onderwijs.	5	15	26	46	8	3.38	0.99
... bevorderde de wil om kritischer tegen de zaken aan te kijken.	4	11	26	51	8	3.50	0.92
... resulteerde in extra motivatie om kwaliteit te leveren.	4	10	27	49	10	3.49	0.95
... versterkte de betrokkenheid van teamleden bij het schoolbeleid.	7	16	31	39	7	3.22	1.02
... heeft bijgedragen tot meer samenwerking in onze school.	7	15	36	35	7	3.18	1.01
... heeft ertoe bijgedragen dat men in de school in sterkere mate dezelfde doelen is gaan nastreven.	5	12	33	43	7	3.33	0.96
... versterkte de samenhang in de school.	8	16	38	32	6	3.14	1.02
Schaalscore 'Resulteren in actiepunten'						3.32	.86

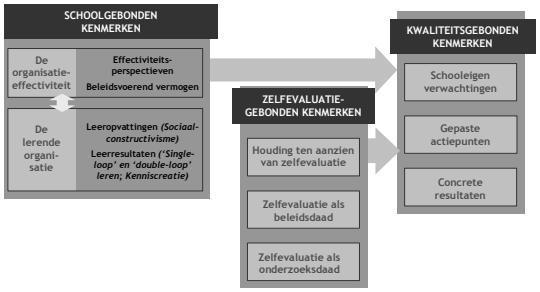
Ook op deze indicator van de kwaliteit van zelfevaluaties scoren de respondenten gemiddeld gematigd positief. Opnieuw is enkel een minderheid er heel sterk van overtuigd dat de zelfevaluatie tot de genoemde resultaten leidde. Het grootste deel van de respondenten scoort 'eerder wel' of 'noch niet/nog wel'. Opnieuw blijkt uit de frequentietabel dat de variatie in de antwoorden groot is. In de meerniveau-analyses kunnen we nagaan in welke mate deze variatie toe te schrijven is aan verschillen tussen scholen en respondenten.

b. MEERNIVEAU-ANALYSES

Onze onderzoekshypothese luidt dat scholen hoger scoren op de verschillende kwaliteitsindicatoren naarmate ze doorheen het zelfevaluatieproces hoger scoren op de zelfevaluatiegebonden kenmerken. We onderzoeken hier in de eerste plaats of deze onafhankelijke variabelen een impact hebben op de kwaliteit van zelfevaluatie. De gepostuleerde voorspellers zijn naast de klassieke school- en respondentenkenmerken dus de houding ten

aanzien van zelfevaluatie en de scores van scholen op de dragers van 'zelfevaluatie als beleidsdaad' en van 'zelfevaluatie als onderzoeksdaad'. Later in paragraaf 2 van dit deel gaan we op zoek naar een spaarzaam model waarin we de kwaliteit op basis van zo eenvoudig mogelijk meetbare voorspellers trachten te voorspellen. In de rapportering over de meerniveau-analyses komen eerst de verschillende benaderingen van kwaliteit van zelfevaluaties aan bod. Daarna onderzoeken we de samenhang op school- en respondentniveau tussen de drie kwaliteitsindicatoren.

➢ *Focus op het voorspellen van de kwaliteit van zelfevaluaties*



De basismodellen onderaan in Tabel 73 maken duidelijk dat de variantie in de perceptie van respondenten omtrent de kwaliteitindicatoren van zelfevaluatie zowel toe te schrijven is aan verschillen tussen scholen als aan verschillen tussen respondenten. Het aandeel van het schoolniveau bedraagt 10% in de schooleigen benadering en respectievelijk 16% à 17% in de schooloverstijgende benaderingen. Opnieuw stellen we dus een grote variantie vast tussen respondenten.

De drie brutomodellen tonen aan dat de concepten die in het theoretisch kader onderscheiden worden allemaal statistisch significant samenhangen met de kwaliteitsindicatoren van zelfevaluaties. Het gaat in de eerste plaats om 'zelfevaluatie als beleidsdaad'. Het verschil tussen respondenten die in dit opzicht tot de cluster 'laag' behoren en respondenten die tot de cluster 'hoog' behoren, bedraagt voor de schooloverstijgende kwaliteitsindicatoren meer dan 2 standaarddeviaties. Als schoolkenmerk bedraagt het effect van deze clusters ongeveer 1.3 maal de standaarddeviatie. Voor de 'zelfevaluatie als onderzoeksdaad' gaat het op respondent- en schoolniveau respectievelijk om ongeveer 1.2 en 1.3 maal de standaarddeviatie. Ook voor de 'houding ten aanzien van zelfevaluatie' is er een grote effectgrootte, weliswaar voornamelijk van de score op schoolniveau. Respondenten geven verder meer aan dat de zelfevaluatie in hun school kwaliteitsvol is wanneer ze van mening zijn dat hun school hoger scoort op de vier effectiviteitsperspectieven (human relations, interne proces, open systeem én rationele doel), op diepgaand leren en op kenniscreatie en naarmate ze zelf hoger scoren op de sociaalconstructivistische visie op leren. De schoolcultuur is bovendien ook als schoolkenmerk van belang. De regressiecoëfficiënten van de schoolgebonden kenmerken liggen wel lager dan bij het voorspellen van de proceskenmerken van de zelfevaluatie (zie 1.2). Verder blijken ook het als respondent genieten van ondersteuning en het waarderen van de genoten ondersteuning positief samen te hangen met de kwaliteit van zelfevaluaties.

Tabel 73: Basis- en brutomodellen van de meerniveau-analyses voor drie kwaliteits-benaderingen van zelfevaluatie (schattingen en standaardfouten (SF))

	Schooleigen benadering	Resulteren in actiepunten	Resulteren in concrete resultaten
VAST	Schatting (SF)	Schatting (SF)	Schatting (SF)
Voorspellers: info over de school			
Basisonderwijs (vs. secundair onderwijs)	**.258 (.072)**	**.425 (.25)**	**.447 (.080)**
Vrij gesubsidieerd onderwijs (vs. GO)	.133 (.078)	.046 (.097)	.031 (.094)
Officieel gesubsidieerd onderwijs (vs. GO)	.015 (.124)	.085 (.151)	.011 (.148)
Geen ervaring met zelfevaluatie (vs. wel ervaring)	-.060 (.079)	-.004 (.098)	.049 (.095)
Geen ondersteuning (vs. wel ondersteuning)	-.121 (.081)	-.090 (.100)	-.179 (.096)
Schoolgrootte (aantal leerkrachten)	**-.129 (.041)**	**-.170 (.052)**	**.157 (.050)**
Schoolcultuur cluster 1 (vs. cluster2)	.130 (.145)	.014 (.177)	-.010 (0.171)
Schoolcultuur cluster 3 (vs. cluster2)	**.359 (.071)**	**.461 (.085)**	**.525 (.077)**
Houding ten aanzien van zelfevaluatie	**.546 (.100)**	**.728 (.118)**	**.839 (.105)**
Zelfevaluatie als beleidsdaad 'hoog' (vs. gem.)	**.541 (.106)**	**.699 (.125)**	**.620 (.123)**
Zelfevaluatie als beleidsdaad 'laag' (vs. gem.)	**-.390 (.072)**	**-.647 (.081)**	**-.631 (.076)**
Zelfevaluatie als onderzoeksdaad 'hoog' (vs. gem.)	**.412 (.062)**	**.613 (.071)**	**.507 (.075)**
Zelfevaluatie als onderzoeksdaad 'laag' (vs. gem.)	**-.433 (.161)**	**-.708 (.200)**	**-.700 (.190)**
Voorspellers: info over de (percepties van de) respondent			
Schoolleiding (vs. teamleden)	**.468 (.072)**	**.341 (.071)**	**.309 (.071)**
Wel ondersteuning (vs. geen ondersteuning)	**.199 (.040)**	**.297 (.041)**	**.278 (.041)**
Perceptie kwaliteit proces ondersteuning	**.258 (.031)**	**.363 (.030)**	**.402 (.030)**
Perceptie kwaliteit resultaat ondersteuning	**.285 (.030)**	**.451 (.028)**	**.536 (.027)**
Leeftijd	.033 (.020)	-.016 (.021)	.008 (.021)
Perceptie van 'human relations'	**.170 (.020)**	**.223 (.020)**	**.292 (0.19)**
Perceptie van 'interne proces'	**.103 (.020)**	**.159 (.021)**	**.182 (.020)**
Perceptie van 'open systeem'	**.135 (.020)**	**.210 (.020)**	**.310 (.019)**
Perceptie van 'rationele doel'	**.122 (.020)**	**.185 (.021)**	**.222 (.020)**
Perceptie van het eigen leren	**.075 (.021)**	**.120 (.021)**	**.123 (.021)**
Perceptie van diepgaand leren in de school	**.301 (.022)**	**.366 (.022)**	**.458 (.022)**
Perceptie van mate van kenniscreatie in de school	**.262 (.019)**	**.357 (.020)**	**.439 (.019)**
Houding ten aanzien van zelfevaluatie	**.237 (.019)**	**.301 (.020)**	**.358 (.019)**
Zelfevaluatie als beleidsdaad 'hoog' (vs. gem.)	**.667 (.044)**	**.901 (.041)**	**.993 (.040)**
Zelfevaluatie als beleidsdaad 'laag' (vs. gem.)	**-.803 (.056)**	**-1.184 (.052)**	**-1.242 (.052)**
Zelfevaluatie als onderzoeksdaad 'hoog' (vs. gem.)	**.467 (.056)**	**.540 (.055)**	**.529 (.056)**
Zelfevaluatie als onderzoeksdaad 'laag' (vs. gem.)	**-.507 (.074)**	**-.709 (.071)**	**-764 (.072)**
VARIABEL			
Basismodel: Variantie schoolniveau	.102 (.021)	.174 (.032)	.162 (.030)
Basismodel: Variantie respondentenniveau	.906 (.026)	.827 (.026)	.853 (.026)

Zoals in de vorige bruto-analyses stellen we bovendien ook hier een statistisch significant effect vast van de functie die de respondent in de school waarneemt. Schoolleiders spre-

ken zich positiever uit over de kwaliteitsindicatoren dan andere teamleden. Tenslotte zijn er ook een aantal achtergrondkenmerken die een statistisch significant effect hebben. Op schoolniveau gaat het om het onderwijsniveau en schoolgrootte. Respondenten uit secundaire scholen en uit grote scholen geven aan dat de zelfevaluatie in hun school minder kwaliteitsvol verliep dan respondenten uit basisscholen en uit kleine scholen. Maar, vermits het hier om bruto-effecten gaat is het niet ondenkbaar dat er sprake is van schijnverbanden. Basisscholen zijn immers klein in vergelijking met secundaire scholen. Bovendien is het niet ondenkbaar dat kleine scholen anders (met name beter) scoren op de schoolgebonden kenmerken. Vandaar het belang om ook naar de nettomodellen te kijken. De nettomodellen worden gepresenteerd in Tabel 74.

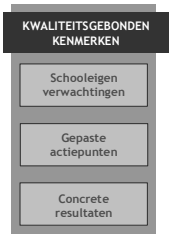
Tabel 74: Nettomodellen van de meerniveau-analyses voor drie kwaliteitsbenaderingen van zelfevaluatie

	Schooleigen benadering	**Resulteren in actiepunten**	**Resulteren in concrete resultaten**
VAST	Schatting (SF)	Schatting (SF)	Schatting (SF)
Intercept	.109 (.049)	-.016 (.065)	.101 (.049)
Voorspellers: info over de school			
Zelfevaluatie als onderzoeksdaad 'hoog' (vs. gem.)		**.225 (.070)**	
Voorspellers: info over de (percepties van de) respondent			
Perceptie van 'human relations'	**.094 (.031)**	**.103 (.031)**	**.165 (.031)**
Perceptie van diepgaand leren in de school		**.075 (.035)**	**.149 (.035)**
Perceptie van mate van kenniscreatie in de school			**.122 (.034)**
Houding ten aanzien van zelfevaluatie	**.135 (.034)**		
Zelfevaluatie als beleidsdaad 'hoog' (vs. gem.)	**.274 (.063)**	**.391 (.062)**	**.416 (.061)**
Zelfevaluatie als beleidsdaad 'laag' (vs. gem.)	**-.313 (.109)**	**-.609 (.102)**	**-.614 (.100)**
Zelfevaluatie als onderzoeksdaad 'hoog' (vs. gem.)	**.239 (0.060)**	**.197 (.058)**	
Zelfevaluatie als onderzoeksdaad 'laag' (vs. gem.)		**-.189 (.078)**	**-.278 (.074)**
VARIABEL			
Nettomodel:Variantie schoolniveau	.027 (.014)	.044 (.015)	.026 (.012)
Nettomodel:Variantie respondentenniveau	.553 (.029)	.418 (.024)	.411 (.023)
R^2 - Reductie van variantie op schoolniveau	39%	50%	52%
R^2 - Reductie van variantie op respondentenniveau	42%	54%	57%

De nettomodellen tonen meteen aan dat heel wat voorspellers die in de brutomodellen een statistisch significante invloed hadden verdwijnen van zodra voor zelfevaluatiegebonden voorspellers gecontroleerd wordt. Voornamelijk de verschillen tussen respondenten en scholen inzake 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad' blijken belangrijke voorspellers van de kwaliteit van zelfevaluatie. De effectgroottes zijn redelijk tot groot. Afhankelijk van de kwaliteitsindicator die gehanteerd wordt, hebben nog drie andere kenmerken van het schoolfunctioneren een statistisch significant effect. Deze zijn de perceptie van de respondent over het effectiviteitperspectief 'human relations' enerzijds en over de mate van diepgaand leren en kenniscreatie in de

school anderzijds. De regressiecoëfficiënten van de schoolgebonden kenmerken zijn wel kleiner dan deze van de zelfevaluatie gebonden kenmerken. In de schooleigen benadering van kwaliteit blijkt tenslotte ook de houding van de respondent nog een statistisch significante rol te spelen. De gepresenteerde modellen zijn sterke modellen. Voor de schooloverstijgende benaderingen van de kwaliteit van zelfevaluaties wordt de fout in de voorspelling met meer dan 50% gereduceerd in vergelijking met het nulmodel (zowel op school- als op respondentniveau).

- *Focus op de samenhang van de drie kwaliteitsindicatoren*



We hebben ook onderzocht hoe de drie kwaliteitsindicatoren samenhangen op school- en respondentniveau. Deze resultaten hiervan worden samengebracht in Tabel 75. Op schoolniveau blijken de drie benaderingen (zeer) sterk samen te hangen. De correlaties variëren van .85 tot .95. De score op 'resulteren in actiepunten' hangt bijna perfect samen met de score op de schooleigen benadering. Ook op respondentniveau is er een aanzienlijke samenhang. Scholen en respondenten die goed scoren op één indicator doen het doorgaans dus ook goed op de andere indicatoren. In paragraaf 2 van dit deel houden we in de analyses rekening met de onderlinge verbanden tussen de kwaliteitsindicatoren.

Tabel 75: Correlaties tussen kwaliteitsindicatoren

	Schooleigen benadering	Resulteren in actiepunten	Resulteren in concrete resultaten	
Schooleigen benadering		0.49	0.46	Respondentniveau
Resulteren in actiepunten	0.95		0.67	
Resulteren in concrete resultaten	0.85	0.89		
	Schoolniveau			

1.5.3 Conclusies bij de vooropgestelde onderzoekshypothesen

Het oordeel over de verschillende kwaliteitsindicatoren van zelfevaluatie is sterk afhankelijk van de persoon die het oordeel uitspreekt. Schoolleiders spreken zich doorgaans positief uit over de zelfevaluaties die in hun school plaatsvonden. In ieder geval schatten zij de kwaliteit van zelfevaluaties hoger in dan andere teamleden in hun school. Dat is conform de verwachtingen uit eerder onderzoek (cf. hypothese 5c). De oordelen van schoolleiders blijken verder niet mooi samen te hangen met deze van externe beoordelaars. We stellen vast dat we op basis van de kwaliteitsbeoordeling door een schoolleider over de GOK-zelfevaluatie die in zijn school plaatsvond geen betrouwbare voorspellingen kunnen doen over de beoordeling van diezelfde zelfevaluatie door de onderwijsinspectie. Dat strookt niet met wat wij in onze hypothese hieromtrent naar voren schoven, namelijk dat beide duidelijk positief zouden samenhangen. In het beoordelen van zelfevaluaties blijkt het perspectief van betrokkenen een belangrijke impact te hebben.

In 'zelfevaluatie binnenstebuiten' is het gemiddelde waardeoordeel van respondenten over de kwaliteitsindicatoren gematigd positief (cf. hypothese 5a). We kunnen dus niet concluderen dat de onderzochte zelfevaluatie in de ogen van alle respondenten zonder meer kwaliteitsvol waren. Een groot deel van de respondenten is weliswaar telkens neutraal of eerder positief. We stellen echter vast dat het gemiddelde beeld een aanzienlijke variatie tussen de bevraagde scholen en tussen respondenten in zich draagt. Er zijn dus wel degelijk verschillen tussen scholen. Zelfevaluaties verlopen in sommige scholen kwaliteitsvoller dan in andere. We hebben echter ook vastgesteld dat er grote verschillen zijn tussen respondenten binnen scholen. Een zelfevaluatie is dus niet in dezelfde mate kwaliteitsvol in de ogen van al de leden van hetzelfde schoolteam. Onze hypothese over de verwachte verschillen in de kwaliteitsindicatoren wordt dan ook bevestigd.

Ook de verwachting (cf. hypothese 5b) dat zelfevaluaties die hoog scoren op één kwaliteitsindicator eveneens hoog scoren op de andere kwaliteitsindicatoren werd door de empirie bevestigd.

Een constante over de gegevensverzamelingen is de bevinding dat verschillen in de kwaliteit van zelfevaluaties in de eerste plaats verklaard worden door kenmerken die betrekking hebben op het verloop van de zelfevaluatie (cf. hypothese 6). In de analyses omtrent de kwaliteit van de GOK-zelfevaluatie bleek bijvoorbeeld dat het beleidsvoerend vermogen van een school nauwelijks samenhangt met de gepercipieerde kwaliteit van de zelfevaluatie. Informatie over de dragers van 'zelfevaluatie als beleidsdaad' draagt daarentegen wel bij tot het voorspellen van de kwaliteitsindicatoren. Ook de meer-niveauanalyses in het kader van 'zelfevaluatie binnenstebuiten' tonen aan dat de grootste netto-effecten toe te schrijven zijn aan zelfevaluatieprocesgebonden kenmerken. De kwaliteit van een zelfevaluatie is met andere woorden afhankelijk van het verloop van die zelfevaluatie. Cruciaal blijkt het tegemoet komen aan de concepten 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad'. In mindere mate speelt ook de 'houding ten aanzien van zelfevaluatie' een rol.

We stelden hoger reeds vast dat het verloop van een zelfevaluatie samenhangt met schoolgebonden kenmerken. Het effect van bijvoorbeeld de schoolcultuur en het lerende karakter op de kwaliteit van een zelfevaluatie loopt dus wellicht over het proces van de zelfevaluatie. Dat roept de vraag op of kenmerken van het functioneren van scholen een rechtstreeks effect hebben op de kwaliteit van zelfevaluaties dan wel een onrechtstreeks effect dat loopt over de proceskenmerken van de zelfevaluatie. Of gaat het om een combinatie van beide? In paragraaf 2 verdiepen we ons in deze vraag. Uit de bovenstaande analyses leren we alvast wel dat zelfevaluaties minder kwaliteitsvol verlopen in scholen waar minder sprake is van bijvoorbeeld diepgaand leren of vormen van kenniscreatie. De onderzoeksresultaten zijn dus een bevestiging van de relaties in het theoretisch model en logischerwijze ook van de hypothesen die we omtrent het verklaren van de kwaliteit van zelfevaluaties vooropstelden.

1.6 De beoordeling en de impact van ondersteuning

Op basis van het theoretisch kader werden een aantal hypothesen geformuleerd omtrent de impact van ondersteuning op zowel de processen als de resultaten van zelfevaluaties. In deze paragraaf toetsen we deze hypothesen. We staan evenwel eerst stil bij een aantal beschrijvende resultaten omtrent de ondersteuning.

1.6.1 Beschrijvende resultaten

De aard van de ondersteuning waarop teamleden in scholen een beroep doen werd reeds in deel 6 beschreven (zie paragraaf 4). We focussen in deze beschrijvende resultaten daarom op de beoordeling van het proces en van de resultaten van de ondersteuning. Tabel 76 bevat de resultaten van de items en schalen die dit in kaart brachten.

Tabel 76: Frequentieverdeling (%), gemiddelde score en standaarddeviatie voor de items en schalen m.b.t. ondersteuning

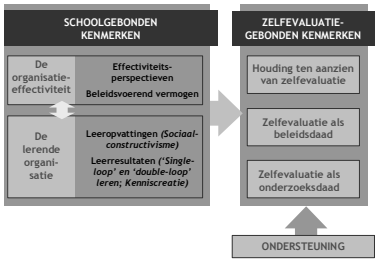
Proces: De ondersteuning waarin ik deelnam...	1	2	3	4	5	Gem	SD
...speelde voldoende in op mijn behoeftes.	4	13	28	48	7	3.42	.93
...hield voldoende rekening met de eigenheid van onze school.	4	14	28	46	8	3.41	.96
Schaalscore 'Proces ondersteuning'						*3.41*	*.84*
Resultaat: De ondersteuning waarin ik deelnam...							
...slaagde erin mij te overtuigen van het belang van het uitvoeren van zelfevaluaties.	4	10	25	46	15	3.58	.98
...slaagde erin mijn kennis over (het verloop van) zelfevaluaties uit te breiden.	3	11	28	47	11	3.52	.92
...slaagde erin de mate waarin ik verbanden leg en samenhangen begrijp tussen aspecten van zelfevaluatie te vergroten.	3	10	34	47	6	3.45	.86
...stelt mij in staat om in de praktijk een kwaliteitsvolle zelfevaluatie uit te voeren.	4	13	32	44	7	3.38	.92
...slaagde erin mij er toe te brengen om zelfevaluaties spontaan te gaan toepassen.	5	14	34	41	6	3.31	.95
Schaalscore 'Resultaat ondersteuning'						*3.44*	*.79*

Respondenten spreken zich gematigd positief uit over beide aspecten van de ondersteuning. De gemiddelde scores liggen telkens een eind boven het middelpunt van de schaal. Grofweg de helft van de respondenten die ondersteuning genoten geeft aan dat deze ondersteuning (eerder wel) voldeed aan de twee proceskenmerken en in de bevraagde resultaten resulteerde. Ongeveer een derde is neutraal en 15% (eerder) negatief.

We zien overigens dat de theoretisch verwachtte verschillen in de scores op de items van de schaal 'resultaat ondersteuning' zich enkel in beperkte mate manifesteren. Er zit een dalende lijn in de gemiddelde scores als men de items van boven naar onder leest. De voorbereidende analyses hebben echter aangetoond dat hoge scores op het ene item samen hangen met hoge scores op de andere items.

Een meerniveau-analyse leert dat de twee schalen onderling sterk samen hangen. Op schoolniveau bedraagt de correlatie .98, op respondentniveau .62.

1.6.2 Meerniveau-analyses



In de reeds uitgevoerde (bruto-)analyses bleek ondersteuning sterk samen te hangen met de bestudeerde zelfevaluatiegebonden kenmerken en met de kwaliteitsindicatoren van zelfevaluatie. Omdat het aantal respondenten dat zich uitsprak over deze aspecten van de genoten ondersteuning maar een derde van het totale aantal bedraagt, hebben we de schalen in Tabel 76 voorheen niet in de nettomodellen opgenomen. Het zou de analyses niet alleen minder 'krachtig' gemaakt hebben, het zou ze ook beperkt hebben tot de respondenten die ondersteuning genoten. In de vorige netto-analyses werd enkel de vraag naar de impact van ondersteuning gesteld, los van de kwaliteit van het proces en het resultaat van die ondersteuning. In deze paragraaf richten we ons specifiek op de impact van de beoordeling van ondersteuning op de nettomodellen inzake zelfevaluatiegebonden kenmerken. De volgende analyses hebben dus noodgedwongen enkel betrekking op de groep van respondenten die een vorm van ondersteuning genoten heeft én zich in de vragenlijst heeft uitgesproken over het proces en het resultaat van deze ondersteuning[27]. We richten ons in Tabel 77 op de drie zelfevaluatiegebonden concepten. De tabel is anders opgebouwd dat voorheen. Tussen de haakjes staan deze keer immers de parameterschattingen voor het nettomodel zonder de indicatoren van ondersteuning (zoals hoger gepresenteerd).

We leren uit het variabele luik van de drie modellen meteen dat de reductie van variantie op school- en op respondentniveau (R^2) aanzienlijk groter is wanneer de indicatoren omtrent de ondersteuning in het model opgenomen worden. Dat is meteen een eerste indicator van het belang van de beoordeling van de ondersteuning in het verklaren van verschillen in de zelfevaluatiegebonden kenmerken. In de gepresenteerde nettomodellen is de perceptie van het resultaat van de ondersteuning telkens de meest gewichtige voorspeller. Bij het concept 'zelfevaluatie als beleidsdaad' speelt ook het proces een statistisch significante rol. Het sterke bruto-effect dat we hoger vonden, blijft dus aanwezig in de nettomodellen. De hypothese dat kwaliteitsvolle ondersteuning samenhangt met hoge scores op de zelfevaluatiegebonden kenmerken wordt hiermee bevestigd. We merken ook dat een aantal voorspellers niet meer statistisch significant (NS) is[28]. Een opmerkelijke vaststelling is het verdwijnen van het functie-effect bij het concept 'zelfevaluatie als beleidsdaad'. Schoolleiders scoren niet hoger dan andere teamleden van zodra het effect van ondersteuning in rekening wordt gebracht. Ook bij de andere zelfevaluatiegebonden

[27] Dit verklaart waarom de indicator 'Wel ondersteuning (vs. geen ondersteuning)' niet in de analyses werd opgenomen (cf. NO).

[28] Merk op dat dit zowel het gevolg kan zijn van het feit dat er gecontroleerd wordt voor de kenmerken van de ondersteuning als van het feit dat de geschatte standaardfouten groter zijn omdat de analyses gebaseerd zijn op een beperkter aantal respondenten. In deze analyses worden kleine effecten immers minder snel gedetecteerd.

concepten verkleint het effect van de functie. Dat kan betekenen dat de teamleden die ondersteuning genieten sterker op schoolleiders gelijken dan teamleden die niet deelnemen aan ondersteuning. Het kan echter ook betekenen dat de verschillen verdwijnen, precies als gevolg van de ondersteuning. Of nog: vermits we weten dat de schoolleiding vaker aan ondersteuning deelneemt dan andere teamleden (zie deel 6), kan het effect dat we voorheen vonden ook moeten toegeschreven worden aan het feit dat al dan niet deelnemen aan ondersteuning de verschillen tussen schoolleiding en de grote groep teamleden in de school nog groter maakt. Om dit verder te exploreren lijkt het zinvol een onderscheid te maken tussen teamleden die aan ondersteuning deelnemen en teamleden die niet aan ondersteuning deelnemen. In onze zoektocht naar een volledig verklaringsmodel (zie paragraaf 2) wordt dit meegenomen.

Tabel 77: Nettomodellen met indicatoren voor de kwaliteit van de ondersteuning

	Houding t.a.v. ZEV	ZEV als beleidsdaad	ZEV als onderzoeksdaad
VAST	Schattingen	Schattingen	Schattingen
Intercept	.060 (.025)	.141 (-.135)	.014 (-.009)
Voorspellers: info over de school			
Schoolcultuur cluster 3 (vs. cluster2)	.136 (.131)	.131 (.201)	
Voorspellers: info over de (percepties van de) respondent			
Schoolleiding (vs. teamleden)	.230 (.453)	NS (.259)	.271 (.457)
Wel ondersteuning (vs. geen ondersteuning)	NO (.108)	NO (.203)	
Perceptie kwaliteit 'proces ondersteuning'	NS (NO)	.105 (NO)	NS (NO)
Perceptie kwaliteit 'resultaat ondersteuning'	.313 (NO)	.300 (NO)	.347 (NO)
Perceptie van 'human relations'	-.073 (-.061)		
Perceptie van 'interne proces'			.089 (.068)
Perceptie van 'open systeem'	NS (.057)	NS (.074)	
Perceptie van 'rationele doel'	NS (.064)	NS (.059)	NS (.093)
Perceptie van het eigen leren	.098 (.154)		.125 (.117)
Perceptie van diepgaand leren in de school	.172 (.210)	.272 (.297)	.188 (.132)
Perceptie van mate van kenniscreatie in de school	.158 (.170)	.210 (.267)	NS (.110)
VARIABEL			
Variantie op schoolniveau	.030 (.031)	.043 (.062)	.006 (.051)
Variantie op respondentenniveau	.518 (.655)	.328 (.483)	.579 (.702)
R^2 - Reductie van variantie op schoolniveau	45% (30%)	58% (38%)	31% (16%)
R^2 - Reductie van variantie op respondentenniveau	46% (32%)	62% (44%)	36% (17%)

1.6.3 Conclusies bij de vooropgestelde onderzoekshypothesen

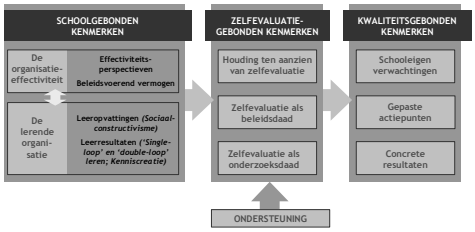
Respondenten die een vorm van ondersteuning genoten, spreken zich gemiddeld gematigd positief uit over het proces en het resultaat hiervan. Dat verklaart waarom we in eerdere analyses een positief effect vonden van het al dan niet genieten van ondersteuning. De meerniveau-analyses waarover we zonet rapporteerden geven immers aan dat het effect van ondersteuning sterk afhankelijk is van de gepercipieerde kwaliteit van die ondersteu-

ning. De perceptie van de drie zelfevaluatiegebonden kenmerken is duidelijk hoger naarmate de ondersteuning als meer kwaliteitsvol beoordeeld wordt. Daardoor verwachten we ook een positief effect van kwaliteitsvolle ondersteuning op de kwaliteit van zelfevaluaties. Of het inderdaad correct is te veronderstellen dat ondersteuning enkel een indirect effect heeft op de kwaliteit van zelfevaluaties onderzoeken we in de volgende paragraaf.

2 EEN PASSEND VERKLARINGSMODEL VOOR DE KWALITEIT VAN ZELFEVALUATIES

In deze paragraaf komen we tot een aantal verklaringsmodellen. Dat er sprake is van een aantal modellen is het gevolg van de twee invalshoeken waarmee we de analyses benaderen. Het gaat om de invalshoek van het 'volledige verklaringsmodel' enerzijds en van het 'meest spaarzame model' anderzijds. In eerste instantie willen we al de gegevens waarover we beschikken, aanwenden om een zo volledig mogelijk beeld te geven van de verbanden tussen de bestudeerde concepten. In tweede instantie willen we een gepaste balans vinden tussen 'verklaringskracht' en de hoeveelheid informatie die in het model moet opgenomen worden. Hieronder worden eerst de resultaten van beide invalshoeken gepresenteerd. Daarna komen we tot een aantal conclusies.

2.1 Het volledige verklaringsmodel



In het theoretisch kader werden een aantal verwachtingen naar voren geschoven omtrent (het pad van de) de relaties tussen de school-, zelfevaluatieproces- en zelfevaluatiekwaliteitgebonden concepten. Structural equation modelling (SEM) maakt het mogelijk te toetsen of het veronderstelde pad aansluit bij de verzamelde gegevens. Zulke padanalyses kunnen tegenwoordig ook toegepast worden op meerniveau data. Op die manier kunnen we in de verschillende concepten komen tot een decompositie van de antwoorden van respondenten in een individuele component en een schoolcomponent (De Maeyer & Rymenans, 2004). De volgende analyses werden uitgevoerd aan de hand van Mplus, een softwarepakket voor (onder meer) meerniveau SEM dat voortbouwt op de benadering van Muthén (2003).

In de meerniveau analyses werden – naast de concepten uit het theoretisch kader - ook de kwaliteit van de ondersteuning en de functie van respondenten gemodelleerd. De eerdere analyses hebben immers het belang van deze variabelen aangetoond. Ondersteuning wordt aan de hand van twee dummies geanalyseerd[29]. De dummy 'ZWAK vs. GEEN

[29] Enkel respondenten die ondersteuning genoten, hebben zich over de kwaliteit van de ondersteuning uitgesproken. Door de indicator 'resultaat van de ondersteuning' in de analyses op te nemen

OND' zet de respondenten die geen ondersteuning genoten af tegen de respondenten die wel, maar verhoudingsgewijs volgens henzelf minder kwaliteitsvolle ondersteuning genoten. De dummy 'STERK vs. GEEN OND' zet de respondenten die geen ondersteuning genoten af tegen de respondenten die wel, en verhoudingsgewijs kwaliteitsvolle ondersteuning genoten. Het onderscheid tussen beide is gebaseerd op de gestandaardiseerde score op de schaal 'resultaten van de ondersteuning'. 'Zwakke (sterke) ondersteuning' is van toepassing op de respondenten die een negatieve (positieve) gestandaardiseerde score hebben op deze variabele. De andere dummy 'functie' richt zich op het onderscheid tussen schoolleiders en teamleden. Deze dummy werd enkel op respondentniveau in de analyses opgenomen[30].

Het volledige verklaringsmodel bevatte bij aanvang alle concepten uit het theoretisch kader. Niet statistisch significante regressielijnen en covarianties werden vervolgens uit het model verwijderd. De concepten die niet in het finale model opgenomen zijn, dragen dus – na controle voor de wel aanwezige concepten - niet noemenswaardig bij aan het verklaren van variantie. Informatie over de verklaarde variantie, de model fit en de intraklasse correlaties voor het volledige verklaringsmodel is opgenomen in Tabel 78. De verschillende fit indices geven aan dat het model goed bij de data aansluit: zowel CFI als TLI zijn groter dan .95, de RMSEA is kleiner dan .06 en de beide SRMR-waarden zijn kleiner dan .08 (Hu & Bentler, 1999). Het model past dus zowel op respondent als op schoolniveau. Voornamelijk op schoolniveau blijken de opgenomen concepten goede voorspellers. De verklaarde variantie ligt voor de drie kwaliteitsindicatoren tussen .81 en .94. Voor de zelfevaluatiegebonden kenmerken liggen deze tussen .47 en .73. Op respondentniveau ligt R^2 enigszins lager. Het model is hier het meest sterk in het verklaren van de variantie in de kwaliteitsindicatoren 'resulteren in actiepunten' en 'concrete resultaten' (respectievelijk 37% en 45%). Ook voor 'zelfevaluatie als beleidsdaad' kan een behoorlijk deel van de variantie verklaard worden (met name 36%). De andere concepten blijken in sterkere mate verklaard te worden door variabelen die niet in ons model opgenomen zijn. De intraklasse correlaties geven overigens aan dat werken met meerniveau-analyses inderdaad aangewezen is. Tussen 9% en 20% van de variantie in de scores van respondenten is toe te schrijven aan verschillen tussen scholen.

zou het aantal respondenten zakken naar (n=629). Het gemiddelde aantal respondenten per school zou dan zakken naar iets meer dan 6. Vier scholen zouden uit de analyses verdwijnen omdat er geen enkele respondent ondersteuning genoot. Verder zouden er in een kwart van de scholen slechts 3 of minder respondenten zijn die zich uitspraken over de kwaliteit van de ondersteuning. We hebben daarom gebruik gemaakt van een alternatieve benadering aan de hand van twee dummies. Op die manier kan de informatie van alle respondenten in de analyse opgenomen worden.

[30] Vermits niet alle teamleden van een school aan de bevraging deelnamen bekomen we op basis van de dummy functie immers geen valide beeld van de verhouding schoolleiders en teamleden in scholen.

Tabel 78: Verklaarde variantie en model fit voor het 'best verklarende model'

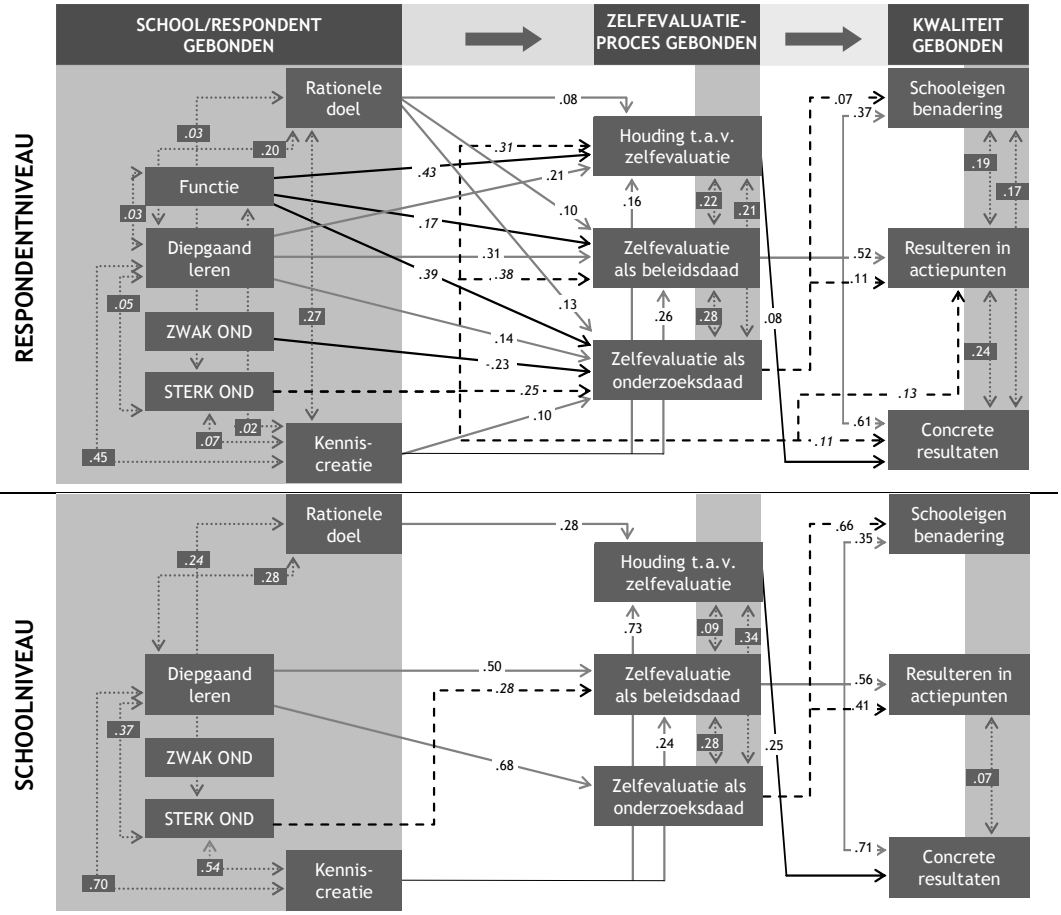
	Verklaarde variantie – R^2		Intra-klasse correlaties
	Respondent-niveau	School-niveau	
Kwaliteitsindicator 1: Schooleigen benadering	.17	.94	.09
Kwaliteitsindicator 2: Resulteren in actiepunten	.37	.84	.17
Kwaliteitsindicator 3: Concrete resultaten	.45	.81	.16
Zelfevaluatie als beleidsdaad	.36	.73	.20
Houding ten aanzien van zelfevaluatie	.21	.61	.09
Zelfevaluatie als onderzoeksdaad	.15	.47	.09
Model fit			
Chi^2 (df)	203.31 (71) ($p < 0.001$)		
CFI	.97		
TLI	.96		
RMSEA	.036		
SRMR within (respondentniveau)	.073		
SRMR between (schoolniveau)	.037		

De regressiecoëfficiënten en correlaties in het verklaringsmodel worden gerepresenteerd in Figuur 25. Enkel de statistisch significante parameters werden in het model opgenomen. De parameters die binnen de grijze vakken gerapporteerd worden zijn correlaties. De parameters tegen de witte achtergrond zijn de regressiecoëfficiënten. Deze zijn, uitgezonderd voor de dummy variabelen, telkens gestandaardiseerd.

Een eerste vaststelling in het padmodel is dat er slechts drie schoolgebonden kenmerken een statistisch significant effect hebben op de zelfevaluatiegebonden kenmerken wanneer de correlaties tussen de zelfevaluatieprocesgebonden concepten in het model opgenomen worden. Dat geldt zowel voor het school- als voor het respondentniveau. Het gaat om de score op het effectiviteitperspectief 'rationele doel', de score op 'diepgaand leren' en de score op 'kenniscreatie'. Op *respondentniveau* heeft elk van deze drie voorspellers een statistisch significante impact op de drie zelfevaluatieprocesgebonden kenmerken. 'Kenniscreatie' en 'diepgaand leren' hebben de hoogste regressiecoëfficiënten. Respondenten die één SD hoger scoren dan gemiddeld op deze voorspellers scoren respectievelijk .26 en .31 hoger op 'zelfevaluatie als beleidsdaad'. Het effect van de schoolgebonden kenmerken op de kwaliteit van zelfevaluatie is enkel indirect. Ook de impact van de functie en van ondersteuning op de zelfevaluatiegebonden kenmerken wordt bevestigd. Schoolleiders scoren bijvoorbeeld respectievelijk .43 SD, .17 SD en .39 SD hoger dan andere teamleden op de 'houding ten aanzien van zelfevaluatie', 'zelfevaluatie als onderzoeksdaad' en 'zelfevaluatie als beleidsdaad'. Figuur 25 geeft verder ook zeer duidelijk de impact aan van de kwaliteit van de ondersteuning waarop respondenten een beroep deden. De ondersteuning heeft zowel een direct als een indirect effect op de kwaliteit van de bestudeerde zelfevaluatie. Respondenten die in sterke mate aangeven dat de ondersteuning volgens hen tot een aantal wenselijke resultaten leidde, geven ook aan dat

de zelfevaluatie hoger scoorde op de proceskenmerken (we noteren de waarden .31, .38 en .25). Precies door hoger te scoren op deze kenmerken blijkt de ondersteuning ook tot betere resultaten te leiden. Daarnaast is er het directe effect. Respondenten die ondersteuning genoten die bovengemiddeld kwaliteitsvol was, scoren hoger op 'resulteren in actiepunten' (.13 SD) en 'concrete resultaten' (.11 SD) dan zij die aangeven dat de ondersteuning niet kwaliteitsvol was. Doorgaans scoren respondenten die geen ondersteuning genoten niet slechter dan zij die ondergemiddeld kwaliteitsvolle ondersteuning genoten. Enkel bij 'zelfevaluatie als onderzoeksdaad' is er een statistisch significant effect van zwakke ondersteuning (ZWAK OND). Respondenten die aangeven dat de ondersteuning die ze genoten niet kwaliteitsvol was, geven aan dat hun school minder hoog scoorde op de 'zelfevaluatie als onderzoeksdaad' dan respondenten die geen ondersteuning genoten (-.23 SD).

Figuur 25: Resultaten meerniveau padanalyse



Op *schoolniveau* is het aantal statistisch significante regressielijnen beperkter. De score van scholen op 'houding ten aanzien van zelfevaluatie' laat zich voorspellen door de mate van 'kenniscreatie' (.73) en het 'rationele doel' (.28); 'zelfevaluatie als beleidsdaad' door de mate van 'diepgaand leren' (.50) en 'kenniscreatie' (.24) in de school. 'Zelfevaluatie als onderzoeksdaad' laadt enkel op de mate van 'diepgaand leren' (.68). Ondersteuning heeft op schoolniveau enkel een statistisch significante impact op 'zelfevaluatie als beleidsdaad'. Scholen waar verhoudingsgewijs meer respondenten aangeven dat de ondersteuning die ze genoten kwaliteitsvol was, scoren hoger op 'zelfevaluatie als beleidsdaad' (.28). Er is geen negatief effect van ondergemiddeld kwaliteitsvolle ondersteuning in vergelijking met geen ondersteuning.

In het voorspellen van de kwaliteitsindicatoren blijken de effecten op respondentniveau doorgaans ook op te gaan voor het schoolniveau. De score op 'zelfevaluatie als beleidsdaad' is hier de beste voorspeller (de parameters zijn .35, .56 en .71). De score op 'zelfevaluatie als onderzoeksdaad' heeft ook op schoolniveau een statistisch significante impact op de 'schooleigen benadering' en 'resulteren in actiepunten' (de parameters zijn respectievelijk .66 en .41). De 'houding ten aanzien van zelfevaluatie' hangt enkel samen met 'concrete resultaten' (.25). Het directe effect van kwaliteitsvolle ondersteuning op de kwaliteitsindicatoren blijkt zich op schoolniveau niet te manifesteren.

De correlaties tussen de drie zelfevaluatieprocesgebonden kenmerken zijn op beide niveaus statistisch significant. We stellen echter vast dat de correlaties hier lager liggen dan bij eerdere analyses (zie Tabel 57). De correlaties in Figuur 25 zijn dan ook gecontroleerd voor de schoolgebonden kenmerken. Een gelijkaardige vaststelling – meer uitgesproken zelfs - geldt overigens voor de correlaties tussen de kwaliteitsgebonden kenmerken (in vergelijking met Tabel 75). We leren uit deze vaststelling dat een school het niet alleen goed doet op één indicator wanneer ze het ook goed doet op een andere indicator. Ze doet het op de verschillende indicatoren goed omdat bijvoorbeeld haar score op bijvoorbeeld 'zelfevaluatie als beleidsdaad' samenhangt met goede prestaties op de drie verschillende kwaliteitsindicatoren. Op schoolniveau is er immers enkel een beperkte maar statistisch significante positieve samenhang tussen 'resulteren in actiepunten' en 'concrete resultaten'.

We onthouden tenslotte ook dat het niet nodig is rechtstreekse effecten te modelleren tussen de schoolgebonden kenmerken en de zelfevaluatiekwaliteitgebonden kenmerken. Het effect van de schoolgebonden kenmerken loopt op beide analyseniveaus over het proces van de zelfevaluatie. Dat maakt dat de schoolgebonden kenmerken in het voorspellen van de kwaliteit van zelfevaluatie niets toevoegen aan de zelfevaluatiegebonden kenmerken.

2.2 Het meest spaarzame model

In het volledige verklaringsmodel ging het ons in de eerste plaats om de onderlinge relaties tussen de concepten in het model. Het voorspellen van de kwaliteit van zelfevaluaties op basis van het volledige verklaringsmodel vergt veel informatie. Het model bevat

bijvoorbeeld alle zelfevaluatiegebonden concepten en bovendien dragen twee van de drie concepten nog verschillende deelconcepten in zich. De score op het concept 'zelfevaluatie als beleidsdaad' is bijvoorbeeld gebaseerd op de score op 8 dragers van 'zelfevaluatie als beleidsdaad'. Uitspraken doen over het concept 'zelfevaluatie als beleidsdaad' betekent dus dat respondenten alle items van de verschillende dragers behoren in te vullen. Alleen al voor 'zelfevaluatie als beleidsdaad' zijn er dit 40. In deze paragraaf onderzoeken we of het mogelijk is de hoeveelheid informatie die verzameld moet worden om de kwaliteit van zelfevaluatie te voorspellen, te beperken. We willen de variantie in de kwaliteitsindicatoren van zelfevaluatie verklaren en dit op een zo spaarzame manier. Het achterliggende doel is met zo weinig mogelijk informatie een zo goed mogelijk voorspelling doen over de kwaliteit van zelfevaluaties.

Uit het volledige verklaringsmodel bleek dat de school/respondentgebonden kenmerken, uitgezonderd de 'kwaliteit van ondersteuning', enkel een indirect effect hebben op de kwaliteitsindicatoren van zelfevaluatie. Als we de score van respondenten en scholen op de kwaliteitsindicatoren willen voorspellen, volstaat het dus om informatie te verzamelen over de drie zelfevaluatieprocesgebonden kenmerken en de ondersteuning. In onze zoektocht naar het meest spaarzame model kunnen we ons dus beperken tot het meest rechtse deel van het theoretisch kader (en van Figuur 25).

Om te komen tot een instrument dat zo gebruiksvriendelijk/kort als mogelijk is, en toch voldoende sterk is om verschillen in de kwaliteitsindicatoren te voorspellen, vergeleken we de resultaten van de volgende zeven verklaringsmodellen. We geven nog mee dat in de modellen telkens alleen statistisch significante regressiecoëfficiënten opgenomen zijn. Uit Tabel 79 blijkt overigens dat de model fit voor alle modellen prima is.

1. Model 1: Het 'Volledige verklaringsmodel' (rechtse deel van Figuur 25). Dit model bevat informatie over de 'houding ten aanzien van zelfevaluatie', de concepten 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad', en de kwaliteit van de ondersteuning.

2. Model 2: Het 'Volledige verklaringsmodel met verkorte vragenlijsten'. Dit model bevat dezelfde voorspellers als model 1. Verschillend is dat in dit model de verkorte instrumenten van 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad' als voorspellers zijn opgenomen (zoals in de voorbereidende analyses beschreven). Daardoor bevat het model de helft minder items, met name 39 in plaats van 78. Het aantal opgenomen concepten blijft op 4.

3. Model 3: Het 'Afzonderlijke dragers model'. Het verschil tussen dit model en model 2 is dat het concept 'zelfevaluatie als beleidsdaad' opgenomen wordt aan de hand van de twee dragers die het meest informatie geven over het geheel van de acht dragers (zie eerdere analyses). Deze zijn 'doeltreffende communicatie' en 'reflectief vermogen'. Het model bevat 38 items die verdeeld zijn over 5 concepten.

   De model fit en verklaarde variantie van deze drie modellen is gelijkaardig. Model 2 is echter het meest spaarzame. Het onderscheid tussen de twee dragers in model 3

draagt onvoldoende bij om dit verder mee te nemen. In het volgende model proberen we daarom Model 2 nog verder te vereenvoudigen.

4. Model 4: Het 'Zelfevaluatiegebonden voorspellers model'. In dit model worden enkel de drie zelfevaluatiegebonden kenmerken als voorspellers voor de kwaliteitsindicatoren opgenomen. In vergelijking met model 2 verdwijnt dus het effect van ondersteuning.

   Model 4 blijkt, gegeven de inperking van 38 naar 28 items, niet te moeten onderdoen voor model 3. Enkel bij de indicator 'concrete resultaten' daalt de verklaarde variantie enigszins. Gezien de winst inzake spaarzaamheid is het verlies aan voorspellingskracht evenwel aanvaardbaar.

5. Model 5: Het 'Twee voorspellers model'. In dit model worden enkel 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad' als voorspellers opgenomen. De 'houding ten aanzien van zelfevaluatie' wordt niet gemodelleerd omdat dit concept in model 4 globaal genomen het minst sterk laadt op de drie kwaliteitsindicatoren. Het 'Twee voorspellers model' bevat twee concepten die aan de hand van 16 items gemeten kunnen worden.

   Model 5 verklaart gemiddeld 1.3% minder variantie dan model 4. Het aantal te bevragen items daalt echter van 28 naar 16. In de pragmatische benadering van het zoeken naar het meest spaarzame model geniet model 5 dan ook de voorkeur. Als we dit model nog verder willen vereenvoudigen komen we uit bij de volgende twee 'Eén voorspeller modellen'.

6. Model 6: Het 'Zelfevaluatie als beleidsdaad model'. Dit model bevat enkel de verkorte meting van het concept 'zelfevaluatie als beleidsdaad' als voorspeller.

7. Model 7: het 'Zelfevaluatie als onderzoeksdaad model'. Dit model bevat enkel de verkorte meting van het concept 'zelfevaluatie als onderzoeksdaad' als voorspeller.

Model 7 verklaart noemenswaardig minder variantie in de kwaliteitsindicatoren dan modellen 5 en 6. De finale keuze moet dan ook gemaakt worden tussen het 'Twee voorspellers model' (model 5) en het 'Zelfevaluatie als beleidsdaad model' (model 6). Hoewel model 6 gemiddeld 1.7% minder variantie verklaart selecteren we dit model toch als het meest spaarzame model. In de eerste plaats omdat het spaarzamer is in termen van concepten en items. Daarnaast omdat kiezen voor 'zelfevaluatie als beleidsdaad' impliceert dat men zich beperkt tot items die men aan een grotere groep respondenten kan voorleggen. Respondenten in scholen die niet nauw bij de zelfevaluatie betrokken werden kunnen er toch uitspraken over doen. Dat geldt niet voor 'zelfevaluatie als onderzoeksdaad'. Model 6 is dus breder inzetbaar. Dat blijkt ook uit het aantal respondenten in Tabel 79.

Tabel 79: Verklaarde variantie, model fit en aantal items voor de getoetste modellen

Verklaarde variantie- R^2							Gem. %[31]
	Schooleigen benadering		Resulteren in actiepunten		Concrete resultaten		
	Respondent	School	Respondent	School	Respondent	School	
Model 1:	.18	.88	.37	.87	.45	.83	59,7
Model 2:	.17	.93	.33	.90	.36	.81	58,3
Model 3:	.18	.92	.34	.90	.41	.89	60,7
Model 4:	.17	.93	.33	.90	.36	.81	58,3
Model 5:	.17	.88	.33	.91	.35	.78	57,0
Model 6:	.17	.85	.32	.83	.35	.79	55,2
Model 7:	.10	.87	.18	.86	.17	.59	46,2

Aantal items en model fit								
	Aantal respondenten	Aantal items (concepten)	Chi^2 (df) – (p)	CFI	TLI	RMSEA	SRMR Respondent	SRMR School
Model 1:	1786	78 (4)	36.6 (42) (< .001)	1.0	.99	.03	.01	.08
Model 2:	1554	39 (4)	8.2 (11) (.70)	1.0	1.0	.00	.01	.01
Model 3:	1534	38 (5)	20.4 (14) (.12)	1.0	.99	.02	.01	.03
Model 4:	1554	28 (3)	5.7 (6) (.45)	1.0	1.0	.00	.01	.02
Model 5:	1555	16 (2)	3.5 (3) (.69)	1.0	1.0	.01	.00	.01
Model 6:	2063	10 (1)	3.4 (1) (.07)	1.0	1.0	.03	.00	.02
Model 7:	1562	6 (1)	1.0 (1) (.33)	1.0	1.0	.00	.00	.01

Uit het bovenstaande blijkt dat we door een reductie van 78 naar 10 items slechts 4.5% van de gemiddelde verklaarde variantie in de kwaliteitsindicatoren op respondent- en schoolniveau moeten inboeten. Het effect van de score op het verkorte meetinstrument van 'zelfevaluatie als beleidsdaad' laadt dan ook sterk op de kwaliteitsindicatoren van zelfevaluatie (zie Tabel 80). Het effect is in grote mate een schijneffect, maar dat doet er in de zoektocht naar het meest spaarzame model niet toe.

Tabel 80: Gestandaardiseerde regressiecoëfficiënten in het meest spaarzame model

	Schooleigen benadering	Resulteren in actiepunten	Resulteren in concrete resultaten
Respondentniveau: Zelfevaluatie als beleidsdaad	.41	.56	.59
Schoolniveau: Zelfevaluatie als beleidsdaad	.92	.91	.89

De perceptie van respondenten omtrent 10 vragen over het proces van de zelfevaluatie in hun school (zie Tabel 81), stelt ons in staat verwachtingen uit te spreken over de resultaten van die zelfevaluatie. Indien we de schoolscore benaderen als indicator van de kwaliteit van zelfevaluatie wordt op die manier immers grofweg 80% van de variantie in de

[31] Het doel van deze indicator is enkel de verschillende modellen met elkaar te vergelijken. Inhoudelijk heeft het gemiddelde percentage verklaarde variantie verder geen betekenis.

drie kwaliteitsindicatoren tussen scholen verklaard. Respondenten binnen scholen verschillen echter ook van mening. Van deze verschillen kan bijkomend één derde verklaard worden wat het 'resulteren in actiepunten' en het 'resulteren in concrete resultaten' betreft. Wat de 'schooleigen benadering' betreft bedraagt de verklaarde variantie op respondentniveau slechts 17%.

Tabel 81: Items in het 'Zelfevaluatie als beleidsdaad model'

	Item
1.	Al de nuttige expertise werd gebruikt om de zelfevaluatie tot een goed einde te brengen.
2.	Er was een goede afstemming tussen de leerkrachten en de directie wat hun taken m.b.t. de zelfevaluatie betreft.
3.	Wij werkten als een hecht team samen om de zelfevaluatie tot een goed einde te brengen.
4.	Er werd intern goed over de zelfevaluatie gecommuniceerd.
5.	Er werden tijdens het uitvoeren van de zelfevaluatie goede professionele afspraken gemaakt.
6.	De verantwoordelijken voor de zelfevaluatie vonden volgens mij steun bij collega's.
7.	De draagkracht van ons team was voldoende groot om deze zelfevaluatie uit te voeren.
8.	De zelfevaluatie vond plaats in een sfeer van vertrouwen tussen de directie en leerkrachten.
9.	Er waren voor de zelfevaluatie overzichtelijke overlegprocedures.
10.	Iedereen was goed geïnformeerd over de doelen van de zelfevaluatie.

Naarmate respondenten in sterkere mate aangeven dat de zelfevaluatie in hun school aan deze criteria voldeed, blijken zij ook in sterkere mate van mening dat de zelfevaluatie kwaliteitsvol was.

2.3 Conclusies bij de vooropgestelde onderzoekshypothesen

Het volledige verklaringsmodel biedt in verschillende opzichten een bevestiging van de conclusies die we na de verschillende deelanalyses presenteerden. Deze conclusies gaan we hier niet herhalen[32]. We focussen dan ook op wat we bijkomend kunnen leren uit de zoektocht naar een passend verklaringsmodel.

De padanalyses bleken in de eerste plaats verhelderend omdat ze duidelijkheid verschaffen over de aard van de effecten in het model. Zoals verwacht (zie Hypothese 6) is er enkel een indirect effect van de schoolgebonden kenmerken op de kwaliteit van zelfevaluaties. Het volledige verklaringsmodel maakt het echter mogelijk om meer specifieke conclusies te trekken over het verband tussen de schoolgebonden en zelfevaluatie gebonden kenmerken. Hoewel onze verwachtingen (cf. hypothese 3b) reeds eerder bevestigd

[32] We verwijzen daarvoor naar de eerdere stukken met 'conclusies bij de vooropgestelde onderzoekshypothesen' aan het eind van de verschillende onderdelen van paragraaf 1 van dit deel. Bovendien worden de eerdere conclusies in de 'Algemene conclusies en discussie' samengevat (zie de volgende paragraaf 3).

werden, weten we nu dat slechts drie schoolgebonden kenmerken een statistisch significante rol spelen. Deze zijn de score op het effectiviteitperspectief 'rationele doel', de score op 'diepgaand leren' en de score op 'kenniscreatie'.

Het effect van ondersteuning is gecompliceerder dan wij in onze hypothese (7b) vooropstelden. Op respondentniveau is het effect van kwaliteitsvolle ondersteuning op de kwaliteit van zelfevaluatie zowel direct als indirect (over de drie zelfevaluatiegebonden kenmerken). Op schoolniveau heeft kwaliteitsvolle ondersteuning enkel een direct effect op 'zelfevaluatie als beleidsdaad'.

Om de variantie in de kwaliteit van zelfevaluaties te voorspellen is informatie over de schoolgebonden kenmerken en de kwaliteit van de genoten ondersteuning echter overbodig. Uit het meest spaarzame model blijkt dat het gebruik van de verkorte versie van het instrument om 'zelfevaluatie als beleidsdaad' te meten, quasi evenveel variantie kan verklaren in de kwaliteitsindicatoren dan meer complexe modellen. Daarmee bleken ook onze verwachtingen hieromtrent gegrond (uitgedrukt in hypothese 8).

3 ALGEMENE CONCLUSIES EN DISCUSSIE

De verschillende analyses werden telkens afgerond met een conclusie die gericht was op het bevestigen dan wel bijsturen van de onderzoekshypothesen die we aan het einde van deel 4 vooropstelden. In deze paragraaf vatten we de grote lijnen samen en situeren we dit alles in het ruimere kader van deze studie. Het geheel is gestructureerd op basis van de centrale onderzoeksvragen van deze studie. We onderscheiden dan ook vier thema's: (1) Aan de hand van welke indicatoren kan men de kwaliteit van een zelfevaluatie operationaliseren? (2) Hoe kwaliteitsvol zijn zelfevaluaties in de bestudeerde scholen? (3) Hoe kunnen de verschillen in het verloop van zelfevaluaties verklaard worden? en (4) Hoe kunnen de verschillen in de resultaten van zelfevaluaties verklaard worden?

3.1 Aan de hand van welke indicatoren kan men de kwaliteit van een zelfevaluatie operationaliseren?

Het antwoord op deze onderzoeksvraag bestaat uit twee luiken: een conceptueel en een operationeel. Het eerste luik focust als het ware op wat we moeten weten/meten om de kwaliteit van zelfevaluatie te beschrijven en beoordelen. Het tweede luik focust op hoe we het kunnen meten.

3.1.1 Een conceptueel antwoord

In de kwantitatieve analyses werden drie indicatoren voor kwaliteitsvolle zelfevaluatie gehanteerd. De benaming van deze benaderingen vindt haar oorsprong in de aard van de verwachtingen die aan de benadering ten grondslag liggen. Deze verwachtingen leren dat een zelfevaluatie kwaliteitsvol is wanneer ze (a) voldoet aan de schooleigen verwachtingen, (b) resulteert in gepaste actiepunten en (c) bijdraagt tot een aantal wenselijk ge-

achte resultaten. We spreken daarom van 3 kwaliteitsindicatoren van een zelfevaluatie. De kwantitatieve analyses toonden aan dat deze drie kwaliteitsindicatoren alle drie met succes werden geoperationaliseerd en op een betrouwbare manier kunnen worden gemeten.

Eén indicator vergt nog enige conceptuele verduidelijking. We zullen in het vervolg immers spreken van een respondenteigen benadering in plaats van een schooleigen benadering. Het uitgangspunt bij de schooleigen benadering was immers dat scholen kunnen verschillen in de functies en verwachtingen die ze bij het initiëren van een zelfevaluatie vooropstellen. Dat bleek ook het geval te zijn. We hebben echter vastgesteld dat respondenten ook binnen scholen andere verwachtingen vooropstellen. Eerder dan van een schooleigen benadering is er dan ook sprake van een respondenteigen benadering. Ook binnen scholen worden verschillende verwachtingen vooropgesteld. Die individuele verschillen maken dat er binnen scholen geen sprake is van een soort gezamenlijk beleefde en onbetwiste 'kwaliteit van de zelfevaluatie'. Wat voor het ene teamlid een waardevol resultaat is, is dat niet noodzakelijk ook voor een ander teamlid. De 'respondenteigen' benadering van kwaliteit speelt hier op in omdat ze nagaat in welke mate de zelfevaluatie bijdroeg tot de verwachtingen die respondenten zelf als het meest belangrijk naar voren schuiven.

In wezen kunnen de drie genoemde concepten aangewend worden als drie complementaire kwaliteitsindicatoren. Deze kwaliteitsindicatoren zijn meteen ook gepaste evaluatiecriteria in het kader van meta-evaluatie. Een zelfevaluatie is dan kwaliteitsvol als:

1. ...*deze de verwachtingen die de verschillende respondenten bij de start vooropstellen realiseert*. Deze verwachtingen kunnen niet a priori door externen vastgelegd worden. Ze zijn school- en zoals bleek zelfs respondentspecifiek. Het gaat hem in deze indicator echter niet om de inhoudelijke doelen als dusdanig, wel om de vraag of deze doelen volgens respondenten gerealiseerd werden.
2. ...*deze resulteert in gepaste actiepunten*. Dat betekent dat de zelfevaluatie resulteert in duidelijk omschreven, realistische, tijdgebonden en aanvaardbare actiepunten waarvan de realisatie achteraf kan worden nagegaan.
3. ...*deze resulteert in concrete (en algemeen wenselijk geachte) resultaten*. Dat betekent dat ze de wil om kritischer tegen de zaken aan te kijken bevordert, dat ze resulteert in extra motivatie om kwaliteit te leveren, dat ze bijdraagt tot meer samenwerking en samenhang in de school, dat ze het bespreken van de eigen opvattingen over goed onderwijs stimuleert en de betrokkenheid van teamleden bij het schoolbeleid versterkt.

We weten dat deze drie indicatoren in sterke mate samenhangen. Men zou dus kunnen argumenteren dat informatie over één criterium meteen ook informatie geeft over de andere. Of een selectie van indicatoren wenselijk is, en welke indicatoren daarbij al dan niet de voorkeur verdienen, hangt af van de context van de meta-evaluatie. Een controlerende overheidsinstantie kan immers andere accenten leggen dan een school die zichzelf wil ontwikkelen inzake zelfevaluatie. We komen hierop terug in de eindconclusies.

3.1.2 Een operationeel antwoord

Er dringt zich verder nog een aantal bemerkingen op indien we de drie indicatoren effectief willen gaan gebruiken om de kwaliteit van zelfevaluaties te beoordelen. Het vervolg van deze paragraaf staat in het teken van zulk een beoordeling.

a. EÉN KWALITEIT OF VELE KWALITEITEN?

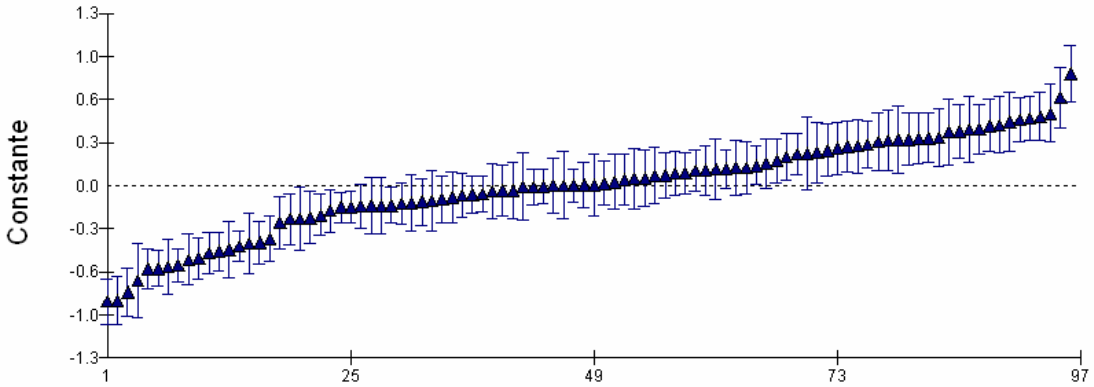
Uitspraken doen over de kwaliteit van zelfevaluaties wordt door een aantal zaken gecompliceerd. Zo blijkt het oordeel over de realisatie van schooloverstijgende verwachtingen sterk te variëren binnen scholen. Eerder dan van één kwaliteit blijkt er dus sprake van verschillende kwaliteiten. In het volgende staan we eerst stil bij de implicaties van een aantal vaststellingen. Die implicaties brengen ons tot vier mogelijke berekeningswijzen voor de drie kwaliteitsindicatoren. Deze berekeningswijzen worden hieronder gaandeweg beargumenteerd en in paragraaf (d) verder toegelicht.

b. VARIATIE IN BEOORDELINGEN

De meerniveau-analyses maakten het mogelijk om tot een decompositie te komen van het oordeel van respondenten over de kwaliteit van de zelfevaluatie in hun school. Er is een deel dat kan worden toegeschreven aan de school en een deel dat kan worden toegeschreven aan de individuele respondent. Dat geldt voor de drie indicatoren. We stelden telkens vast dat de verschillen in kwaliteitsbeoordelingen binnen scholen groter zijn dan de verschillen tussen scholen.

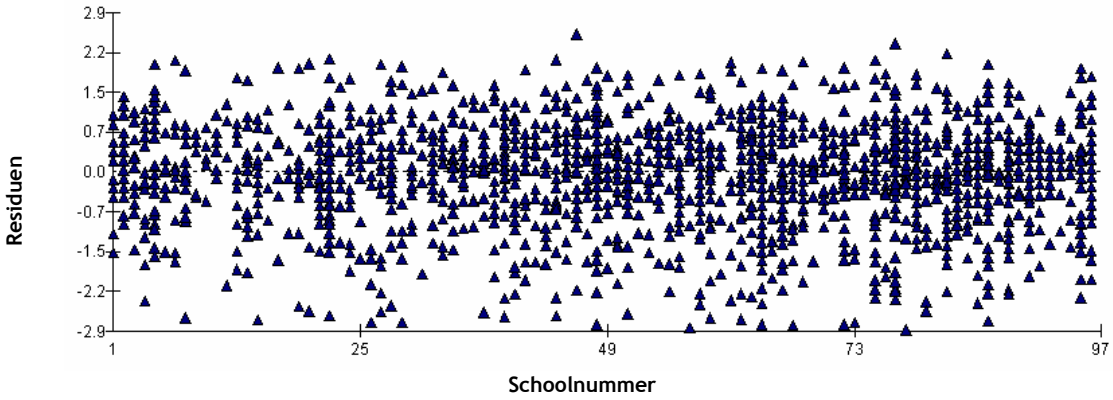
Men zou ‘de kwaliteit van een zelfevaluatie’ kunnen laten samenvallen met het aandeel op schoolniveau. Dit is immers de ‘kwaliteit’ die eigen is aan de school. Figuur 26 geeft voor de 96 deelnemende scholen hun scores op de indicator ‘resulteren in concrete resultaten’ weer. Ook het 95% betrouwbaarheidsinterval rond de ‘kwaliteit op schoolniveau’ opgenomen. Uit de figuur blijkt duidelijk dat er verschillen zijn tussen scholen. De school met de beste score op de kwaliteitsindicator scoort ruim 1.5 SD hoger op ‘resulteren in concrete resultaten’ dan de minst goede school. Men zou scholen op deze manier bijvoorbeeld kunnen indelen in drie kwaliteitsniveaus: minder goed dan gemiddeld, gemiddeld en beter dan gemiddeld. Op die manier kan de schoolscore een waardevolle indicator zijn voor de kwaliteit van zelfevaluaties.

Figuur 26: De 'kwaliteit' eigen aan de bestudeerde scholen



De meningen van al de teamleden in de school variëren sterk rond de schoolscore. Figuur 27 illustreert dit. Als iedereen het bij wijze van spreken eens zou zijn met de schoolscore zouden de afwijkingen van alle respondenten op nul moeten liggen. Een grote groep respondenten scoort echter aanzienlijk boven of onder de schoolscore. Eénzelfde schoolscore resulteert dus in zeer uiteenlopende scores bij de teamleden van die school. Dat gaat tot (ruim) 2 standaarddeviaties boven of onder de schoolscore.

Figuur 27: Afwijkingen van teamleden ten opzichte van de 'kwaliteit' eigen aan hun school



De variatie in beoordelingen roept de vraag op, wat uiteindelijk als meting van de kwaliteit van de zelfevaluatie beschouwd moet worden. Kiest men enkel voor de component die aan de school moet toegeschreven worden (dit is wat we zo dadelijk berekeningswijze 2 zullen noemen), of kent men ook aan de variatie tussen teamleden een rol toe? De schoolscore heeft als voordeel dat het een betrouwbare uitspraak mogelijk maakt over de kwaliteit van zelfevaluaties in scholen. Het nadeel is echter dat het om een (geschatte)

statistiek gaat en niet om een oordeel dat door teamleden op een unisono manier beleefd wordt. Er bestaat dus wel zoiets als 'dé kwaliteit van een zelfevaluatie' in een school, maar het gaat dan voor individuele teamleden om een abstract en statistisch begrip. En verder: wanneer men als onderzoeker enkel geïnteresseerd is in de schoolscore, dan zijn de verschillen tussen respondenten binnen scholen statistisch niet relevant. Zij vertellen ons immers niets extra over de schoolscore. In bepaalde contexten zou het echter een probleem kunnen zijn dat de schoolscore geschat werd en enkel door middel van meerniveauanalyses in kaart kan worden gebracht? Dan stelt zich de vraag of men ook niet simpelweg zou kunnen werken met de gemiddelde score van alle respondenten in een school? In dat geval neemt men zowel de variatie tussen scholen als de variatie tussen individuen binnen scholen mee. Naar deze laatste optie verwijzen we zo dadelijk als berekeningswijze 1.

c. SYSTEMATIEK IN DE VARIATIE

We stelden verder vast dat de variatie in de kwaliteitsindicatoren systematisch samenhangt met de functie van respondenten en met hoe respondenten de zelfevaluatiegebonden kenmerken percipiëren. Ook dat problematiseert het vinden van een objectieve maat om kwaliteit van zelfevaluatie te beoordelen.

➢ *De impact van de functie van respondenten*

De analyses hebben aangetoond dat de functie die men in scholen inneemt een sterke impact heeft op de gepercipieerde kwaliteit. Schoolleiders spreken zich in sterke mate positiever uit over de zelfevaluaties die in hun school plaatsvonden dan andere teamleden. Dat geldt zowel voor het proces als voor het product van de zelfevaluatie. Ook deze vaststelling maakt het moeilijk om eenduidig te bepalen hoe kwaliteitsvol zelfevaluaties verlopen zijn. Een rechtvaardige vergelijking van scholen zou dan ook rekening moeten houden met de functies die de respondenten in een school bekleden. Op die manier kan men vermijden dat scholen waar verhoudingsgewijs veel schoolleiders aan de bevraging deelnemen onterecht een hogere kwaliteit van de zelfevaluatie toegedicht krijgen dan scholen waar minder schoolleiders deelnamen. Men kan de schoolscores daarom ook controleren voor de functie van de respondenten. Dit is wat we berekeningswijze 3 zullen noemen.

➢ *De impact van de perceptie van zelfevaluatiegebonden kenmerken*

We hebben verder ook vastgesteld dat scholen hoger scoren op de drie kwaliteitsindicatoren naarmate de respondenten hoger scoren op 'houding ten aanzien van zelfevaluatie', 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad'. In de zoektocht naar een passende manier om de kwaliteitsindicatoren te berekenen kan men hier op twee manieren op reageren:

- Door de schoolscores te corrigeren voor de scores van respondenten op de zelfevaluatiegebonden kenmerken. Deze aanpak lijkt aangewezen wanneer men ervan uitgaat deze scores ook een indicatie zijn van een vorm van scepticisme of optimisme van de

respondenten. Scholen die bestaan uit een grote groep respondenten die zich eerder negatief uitspreken over de zelfevaluatiegebonden kenmerken zullen in deze logica onterecht ook een lagere schoolscore op de kwaliteitsindicatoren hebben. Deze respondenten schetsen dan immers een te negatief beeld van de kwaliteit van de zelfevaluatie. Het gaat als het ware over de ganse lijn om 'zwartkijkers'. Het uitgangspunt is dat de school als dusdanig hierover geen controle heeft en hier dan ook niet verantwoordelijk voor kan gehouden worden. Door de schoolscore te controleren voor de score op de zelfevaluatiegebonden kenmerken maakt men dan een meer rechtvaardige vergelijking mogelijk. In de volgende paragraaf zullen we dit berekeningswijze 4 noemen.

- Door de schoolscores niet te corrigeren voor de scores van respondenten op de zelfevaluatiegebonden kenmerken. In deze logica zijn de scores van respondenten op de zelfevaluatiegebonden kenmerken wezenlijke en betrouwbare indicatoren van de proceskenmerken van de zelfevaluatie. Een positieve/negatieve houding van respondenten ten aanzien van zelfevaluatie en een hoge/lage perceptie van het tegemoet komen aan de 'zelfevaluatie als beleidsdaad en onderzoeksdaad' zijn dan de verdienste van de school. Het zou dan onrechtvaardig zijn daarvoor te controleren. In deze logica blijft men bij berekeningswijzen 1 of 2.

d. MOGELIJKE BEREKENINGSWIJZEN VAN DE INDICATOREN VOOR DE KWALITEIT VAN ZELFEVALUATIES

In de opzet van een meta-evaluatie dringt zich de volgende vraag op: Wat is - gegeven de bovenstaande bemerkingen - een passende manier om de kwaliteitsindicatoren te meten? Er bleken immers verschillende manieren te bestaan om de drie indicatoren van de kwaliteit van zelfevaluaties te gaan meten. Wij maken een onderscheid tussen vier wijzen waarop de kwaliteitsindicatoren kunnen berekend worden.

- Berekeningswijze 1: De geaggregeerde schoolscore. Deze berekeningswijze gebruikt de gemiddelde score van alle respondenten in de school. Er wordt geen gebruik gemaakt van meerniveauanalyses.
- Berekeningswijze 2: De geschatte schoolscore. Deze berekeningswijze maakt gebruik van de schatting van de schoolscore aan de hand van het nulmodel in meerniveau-analyses.
- Berekeningswijze 3: De geschatte schoolscore met correctie voor de functie van respondenten. De schoolscore wordt hier gecontroleerd voor de functie van de respondenten. De schoolscore komt in berekeningswijze 3 in wezen overeen met de gemiddelde score van de teamleden in de school.
- Berekeningswijze 4: De geschatte schoolscore met correctie voor de perceptie van de zelfevaluatiegebonden kenmerken. De schoolscore wordt hier gecontroleerd voor de drie zelfevaluatiegebonden kenmerken. Deze berekeningswijze geeft de score van de school ervan uitgaande dat alle respondenten gemiddeld zouden scoren op de drie zelfevaluatiegebonden kenmerken.

Deze vier berekeningswijzen kunnen bij elk van de drie kwaliteitsindicatoren toegepast worden (cf. de 'schooleigen indicator', 'resulteren in gepaste actiepunten' en 'resulteren in concrete resultaten'). We illustreren de implicaties van de berekeningswijzen hier

enkel aan de hand van 'Resulteren in concrete resultaten'. We kijken daarbij eerst naar de samenhang tussen de vier berekeningswijzen. In Tabel 82 zijn zowel de correlaties tussen de absolute schoolscores als tussen de onderlinge rangordes van scholen weergegeven (Kendall's tau-b staat tussen haakjes). Hieruit blijkt dat de eerste drie berekeningswijzen in wezen weinig van elkaar verschillende schoolscores opleveren (zowel in absolute getallen als in rangorde van scholen).

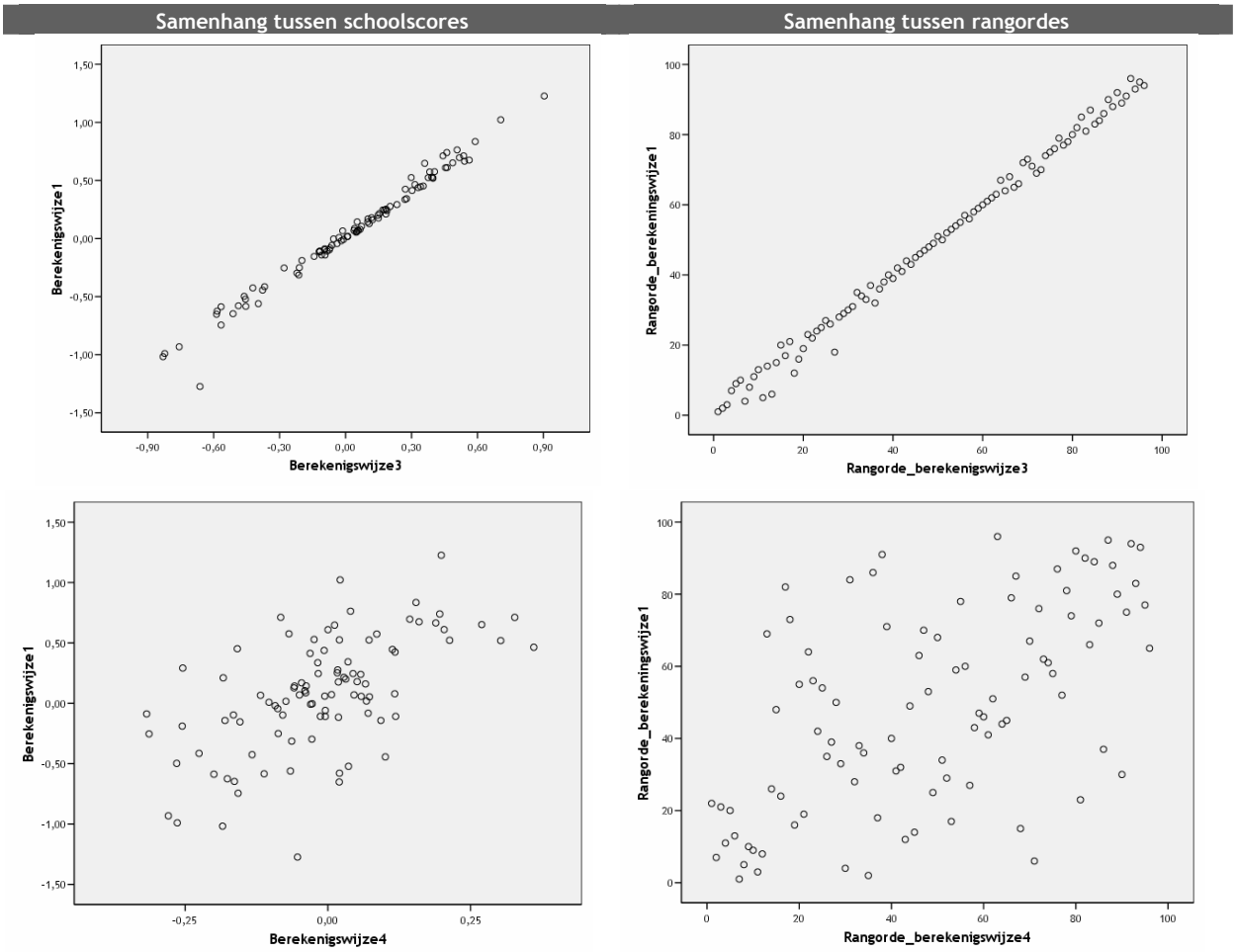
Tabel 82: Samenhang tussen de berekeningswijzen

	Berekenings-wijze 1	Berekenings-wijze 2	Berekenings-wijze 3	Berekenings-wijze 4
Berekeningswijze 1				
Berekeningswijze 2	.99 (.96)			
Berekeningswijze 3	.99 (.96)	.99 (1.0)		
Berekeningswijze 4	.59 (.41)	.61 (.42)	.62 (.42)	

Omdat de scores van individuele respondenten telkens deels boven en deels onder de geschatte schoolscore liggen, voegt meerniveau-analyse als dusdanig weinig toe wanneer men de kwaliteitsindicatoren wil meten zonder voor andere indicatoren te controleren. Het is statistisch perfect logisch dat de geaggregeerde schoolscore (berekeningswijze 1) aan dezelfde scholen een hoge, gemiddelde, en lage score toekent dan het nulmodel in de meerniveau-variant. Ook controleren voor de functie heeft echter geen noemenswaardig effect op de verhoudingen tussen scholen. Wellicht ligt de verhouding schoolleiders en teamleden in de verschillende scholen zo dicht bij elkaar dat de schoolscores van scholen er niet sterk door beïnvloed worden. De twee visualisaties bovenaan in Figuur 28 geven aan dat zowel de absolute schoolscores als de rangorde van scholen in berekeningwijze 1 en berekeningswijze 3 zo goed als identiek zijn.

Controleren voor de perceptie van zelfevaluatiegebonden kenmerken heeft wel een effect. Op de schalen van de Y-assen van de linkse deelfiguren lezen we bijvoorbeeld af dat de variatie tussen scholen veel kleiner is wanneer we berekeningswijze 4 gebruiken in plaats van berekeningswijze 3. Gesteld dat de zelfevaluatiegebonden kenmerken door alle respondenten hetzelfde zouden gescoord worden, zou er dus minder variatie zijn in de kwaliteitsindicatoren. De noemenswaardige verschillen in rangorde van scholen bij berekeningswijze 4 in vergelijking met berekeningswijze 1 zijn wellicht mede hieraan toe te schrijven. Verder zien we in de deelfiguren dat de schoolscore op basis van berekeningswijze 4 duidelijk minder sterk samenhangt met de scores op basis van berekeningswijze 1 dan de schoolscores op basis van berekeningswijze 3. Er blijft evenwel een behoorlijk positief verband bestaan. Dat wijst erop dat de samenhang tussen de scores van respondenten op de zelfevaluatiegebonden kenmerken enerzijds en hun oordeel over de kwaliteit van de zelfevaluatie anderzijds ook een werkelijke basis heeft. Toch maakt het voor individuele scholen soms een groot verschil welke berekeningswijze men gebruikt (zowel in absolute en relatieve termen). Deze analyses kaarten dus een relevante problematiek aan waarop wij hier niet verder kunnen ingaan omdat de gegevens om verdere uitspraken te doen ontbreken. Wij houden daarom niet vast aan deze berekeningswijze.

Figuur 28: De impact van het controleren voor functie en perceptie op 'zelfevaluatie gebonden kenmerken'



De suggestie dat sommige respondenten de zaken systematisch negatiever of positiever bekijken dan anderen wordt weliswaar niet ontkracht maar controleren voor de perceptie van de zelfevaluatiegebonden kenmerken is onvoldoende zuiver om dit op te lossen. Wij concluderen daarom dat een rechtvaardige meting van de kwaliteit van zelfevaluaties niet vereist dat er gecontroleerd wordt voor de scores op zelfevaluatiegebonden kenmerken. Op basis van deze studie kunnen we evenwel niet met zekerheid beargumenteren dat dit een terechte keuze is. Om dit pleit te beslechten zouden we bijvoorbeeld moeten afstappen van het gebruiken van percepties en zelfrapportage in het beoordelen van de kwaliteit van zelfevaluaties. Bijkomend onderzoek (o.a. de gevalstudies in deel 8) kan het mogelijk maken om te achterhalen of verschillende schoolteams zich al dan niet verschillend uitspreken over de kwaliteit en het proces van een zelfevaluatie die volgens een

objectieve waarnemer even kwaliteitsvol is verlopen. Daarnaast zou het interessant zijn om in het analyseren van de percepties van respondenten ook een indicator voor de mate van 'scepticisme' of van 'negativisme/positivisme' op te nemen. Onze hypothese blijft vooralsnog echter dat verschillen in de kwaliteit van zelfevaluaties in de eerste plaats toe te schrijven zijn aan het feit dat het zelfevaluatieproces in verschillende scholen op verschillende manieren verloopt.

We concluderen dus dat 'controle' niet tot een meer rechtvaardige meting van de kwaliteit van zelfevaluatie leidt. De keuze voor een passende berekeningswijze hangt verder ook af van de context waarin men uitspraken wil doen over de kwaliteit van zelfevaluaties. De geaggregeerde schoolscore heeft bijvoorbeeld het voordeel dat ze eenvoudig te meten is, meteen kan afgezet worden tegen de beoordelingsschaal wanneer men de scores niet standaardiseert (zie volgende paragraaf) en gebruikt kan worden bij het beoordelen van één enkele school. Bovendien is er geen bijkomende gegevensverzameling nodig om te corrigeren voor andere kenmerken. In het beantwoorden van de vraag hoe kwaliteitsvol de zelfevaluaties in de bestudeerde scholen zijn verlopen, maken we daarom gebruik van berekeningswijze 1.

3.2 Hoe kwaliteitsvol zijn zelfevaluaties in de bestudeerde scholen?

Zowel bij de kwaliteitsindicatoren als bij de zelfevaluatiegebonden kenmerken hebben we herhaaldelijk geconcludeerd dat respondenten zich gemiddeld gematigd positief uitspreken. De vraag dringt zich hier op of de vaststellingen die we in dit deel deden al dan niet wijzen op kwaliteitsvolle zelfevaluaties. Dit is een vraag naar interpretatie. Interpretaties zijn evenwel afhankelijk van het referentiekader dat men hanteert. Wat betekent het bijvoorbeeld dat zelfevaluaties volgens de respondenten gemiddeld 3.32 scoren op 'resulteren in concrete actiepunten'? In het kader van meta-evaluatie zijn ons inziens drie invalshoeken nodig om te bepalen hoe kwaliteitsvol de zelfevaluaties in de bestudeerde scholen zijn verlopen. Het gaat om een normgerichte, een criteriumgerichte én een schoolgerichte invalshoek.

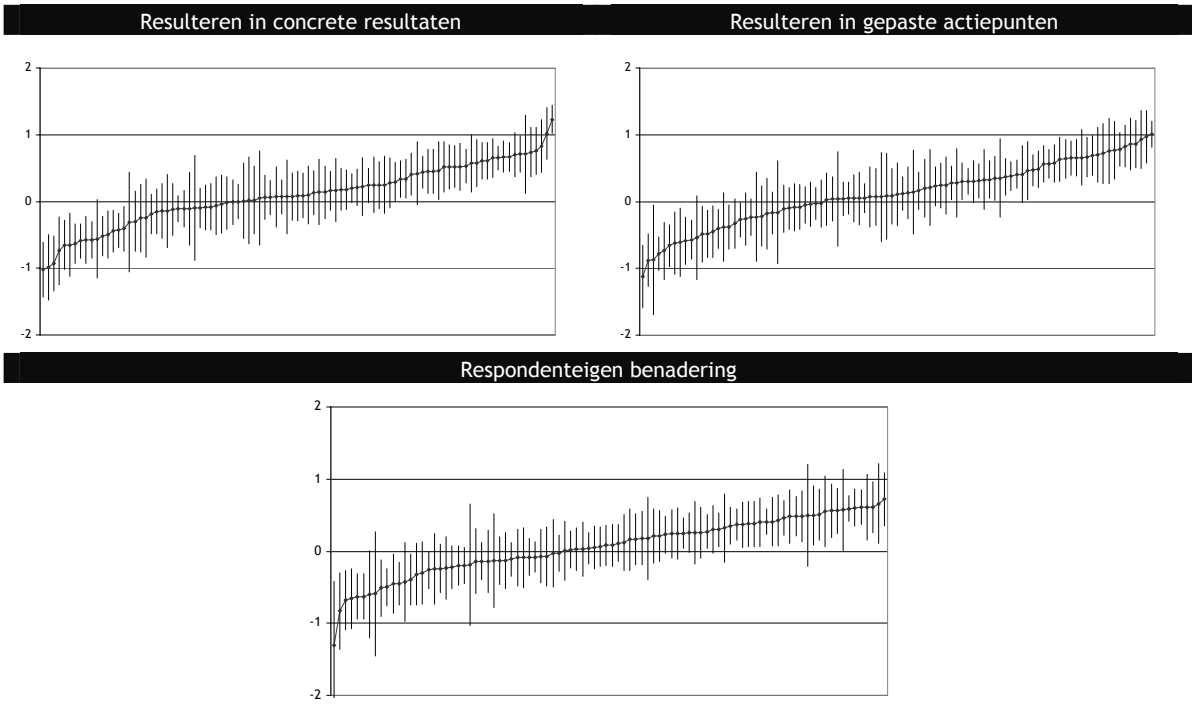
3.2.1 Een normgericht referentiekader

Bij de normgerichte (of relatieve) interpretatie van de resultaten worden de prestaties van een individuele school vergeleken met de prestaties van andere scholen. Er wordt als het ware een rangschikking gemaakt van de resultaten van alle zelfevaluaties. Vervolgens wordt aan de scholen met de beste scores een positieve beoordeling toegekend en aan de scholen met de minst goede score een negatieve beoordeling. Een zelfevaluatie is in deze logica kwaliteitsvol als ze even goed of beter is dan andere zelfevaluaties.

Wanneer we de 95%-betrouwbaarheidsintervallen rond de geaggregeerde schoolscore (zie berekeningswijze 1) leggen dan bekomen we voor de drie kwaliteitsindicatoren het resultaat in Figuur 29. Afgaande op de rangorde in de figuur kunnen we – uitgaande van de normgerichte benadering - bijvoorbeeld concluderen dat de kwaliteit van de zelfevaluatie in een 15-tal scholen 'zwak' is. Hun betrouwbaarheidsinterval ligt volledig onder de ge-

middelde gestandaardiseerde score (0) op de kwaliteitsindicatoren. De andere scholen 'bewijzen' immers dat het beter kan. Een 30-tal scholen scoort bij de drie kwaliteitsindicatoren statistisch significant beter dan gemiddeld. In de normgerichte benadering omschrijven we de kwaliteit van de zelfevaluaties in deze scholen als 'sterk'. De verschillen tussen scholen met een 'zwakke' en een 'sterke' score op de kwaliteitsindicatoren, lopen op tot anderhalf à twee standaarddeviaties. De verschillen zijn dus groot. Zelfevaluaties in scholen wiens betrouwbaarheidsinterval het nulpunt van de schaal omvat, omschrijven we als 'gemiddeld' kwaliteitsvol. Deze groep scholen bevat telkens ongeveer de helft van al de scholen.

Figuur 29: Normgerichte benadering van de drie kwaliteitsindicatoren (gestandaardiseerde scores)



Bij de normgerichte benadering wordt geen rekening gehouden met de vraag of alle zelfevaluaties misschien goed of slecht waren. Indien alle zelfevaluaties goed scoren, zal de zelfevaluatie met de minst goede resultaten toch een negatieve beoordeling krijgen. Een normgericht referentiekader is daarom enkel informatief als het aangevuld wordt met andere referentiekaders. Dat neemt echter niet weg dat het leerrijk is te weten hoe een school scoort in verhouding tot andere scholen. We leren op basis van het normgerichte referentiekader bijvoorbeeld dat zelfevaluaties in sommige scholen statistisch significant minder kwaliteitsvol zijn dan de zelfevaluatie in andere scholen. Bij het beantwoorden

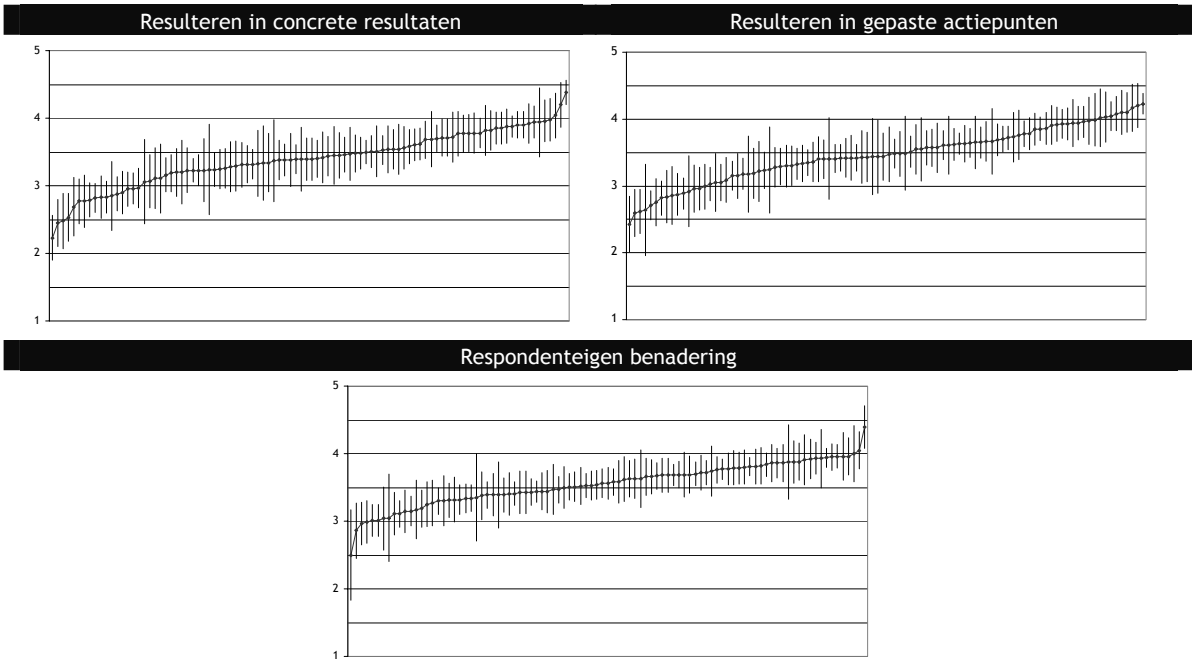
van de onderzoeksvraag 'Hoe kwaliteitsvol zijn zelfevaluaties in de bestudeerde scholen?' mag men de variatie in kwaliteit dan ook niet uit het oog verliezen.

3.2.2 Een criteriumgericht referentiekader

De data van scholen worden bij criteriumgerichte beoordelingen afgewogen tegen vooraf bepaalde criteria en standaarden. Hoge beoordelingen worden toegekend aan scholen die tegemoet komen aan, of beter zijn dan de vooropgestelde standaarden. Lage beoordelingen worden toegekend aan scholen die - afgezet tegen de criteria - onvoldoende kwaliteitsvol zijn. Illustratief voor deze werkwijze is de aanpak van de GOK-inspecties. Scholen werden daar allen tegen dezelfde meetlat gehouden. Realistische en door de verschillende betrokkenen gedragen standaarden om zelfevaluaties aan af te toetsen zijn momenteel evenwel nog niet beschikbaar. Wellicht mede hierdoor en door het beoogde ontwikkelingsperspectief hanteerde de GOK-inspectie een lage standaard (MVG, 2006). We stelden immers vast dat enkel een zeer kleine minderheid van de bestudeerde GOK-zelfevaluaties een onvoldoende kreeg. Wanneer men de standaard laag legt, zullen veel scholen tot kwaliteitsvolle zelfevaluaties komen.

Ook in analyses in het kader van 'zelfevaluatie binnenstebuiten' kunnen we standaarden vooropstellen. We zouden bijvoorbeeld een minimum schoolscore op de indicator 'resulteren in concrete resultaten' kunnen vooropstellen om de zelfevaluatie als kwaliteitsvol te bestempelen. Gegeven de gehanteerde beoordelingsschaal zou dit bijvoorbeeld de score '4' kunnen zijn. De schaal loopt van 1 tot 5 en de score '4' wijst dus op het eerder wel gerealiseerd zijn van de vooropgestelde resultaten. In de grafiek linksboven in Figuur 30 kunnen we aflezen hoeveel scholen aan de standaard voldoen. We leren daaruit dat er slechts één school is waarvan we met zeer grote zekerheid kunnen zeggen dat de zelfevaluatie aan de standaard voldeed. Bij twee bijkomende scholen is de kans groter dat ze eraan voldoen dan dat ze er niet aan voldoen (hun gemiddelde score ligt boven de waarde '4'). 77% van de scholen (74 in absolute aantal) voldoet zeker niet aan de standaard. Bij de resterende 20% (19 in absolute aantal) is de kans groter dat ze er niet aan voldoen dan dat ze er wel aan voldoen. Gegeven deze standaard zou dus slechts één bestudeerde zelfevaluatie met zekerheid het label 'kwaliteitsvol' mogen dragen. Indien we de standaard verleggen naar 3.5 op de schaal, dan zou 21% van de scholen met zekerheid het label 'kwaliteitsvol' mogen dragen. Indien we de standaard nog verlagen en verleggen naar het middenpunt van de schaal (d.i. 3), dan blijkt 63% van de scholen met zekerheid aan de standaard te voldoen. Gegeven deze standaard geldt voor 4 scholen dat de zelfevaluatie die er plaats vond niet kwaliteitsvol was. Voor de andere scholen (33%) is er geen zekerheid. Voor 13% geldt dat zij waarschijnlijk onder de standaard scoren, voor 20% geldt dat zij waarschijnlijk boven de standaard scoren.

Figuur 30: Criteriumgerichte benadering van de drie kwaliteitsindicatoren (niet gestandaardiseerde scores)



De standaard bepaalt in het criteriumgerichte referentiekader in welke mate de zelfevaluaties kwaliteitsvol waren. Het is op dit moment echter zeer moeilijk om te bepalen welke een passende standaard zou zijn. De bovenstaande cijfers tonen echter aan dat een standaard die de lat eerder hoog legt, impliceert dat de zelfevaluaties in een groot aantal scholen voor de indicator 'resulteren in concrete resultaten' als weinig of niet kwaliteitsvol moeten bestempeld worden. Om de implicaties van de keuze van een standaard verder te illustreren geven we in Tabel 83 de percentages scholen weer waarin de zelfevaluatie volgens de verschillende standaarden 'met zekerheid kwaliteitsvol', 'waarschijnlijk kwaliteitsvol', 'waarschijnlijk niet kwaliteitsvol' en 'met zekerheid niet kwaliteitsvol' zijn verlopen.

Tabel 83: Kwaliteit van de bestudeerde zelfevaluaties naargelang de gehanteerde standaarden

	Resulteren in concrete resultaten			Resulteren in gepaste actiepunten			Respondenteigen benadering		
Standaard	4	3.5	3	4	3.5	3	4	3.5	3
Met zekerheid kwaliteitsvol	1%	21%	63%	1%	26%	66%	1%	12%	79%
Waarschijnlijk kwaliteitsvol	2%	18%	20%	8%	29%	19%	2%	30%	17%
Waarschijnlijk niet kwaliteitsvol	21%	41%	13%	23%	20%	10%	27%	27%	4%
Met zekerheid niet kwaliteitsvol	76%	20%	4%	68%	25%	5%	70%	31%	0%

De standaarden die we hanteren zijn de scores 4, 3.5 en 3 op de gebruikte beoordelingsschaal. '4' betekent dat de respondenten in de school gemiddeld aangeven van mening te zijn dat de resultaten eerder wel van toepassing waren op de eigen zelfevaluatie. Bij '3' geeft men aan dat dit 'noch niet/noch wel' het geval was. '3.5' situeert zich tussen beide in.

3.2.3 Een schoolgericht referentiekader

Bij het schoolgericht referentiekader tenslotte worden de prestaties van scholen niet vergeleken met een extern referentiepunt, maar met een intern individueel vergelijkingspunt. Dit punt is gebaseerd op de individuele sterktes en zwaktes van de school, op de eigen verwezenlijkingen, vooruitgang, leerkansen, interesses, doelstellingen, achtergrond en andere factoren die van invloed kunnen zijn op de schooleigen prestatie. De interpretatie van de beschikbare data berust in deze strategie bijvoorbeeld op een vergelijking met de resultaten van vroegere zelfevaluatieprocessen. Eenzelfde kwaliteitsniveau kan in dit referentiekader dus bij verschillende scholen tot een verschillende beoordeling leiden. Concreet: voor een school die vroeger reeds succesvol zelfevaluaties uitvoerde, kan een bepaalde score een achteruitgang betekenen terwijl ze voor een andere school een vooruitgang is. Het gegeven dat het in bijna tweederde van de bestudeerde scholen om de eerste keer ging dat zij een geïnstrumentaliseerde zelfevaluatie uitvoerden, kan de conclusies uit het criteriumgericht referentiekader dus nuanceren. Om het schoolgerichte referentiekader volledig te kunnen benutten zouden we een goed zicht moeten hebben op de historiek van de school inzake zelfevaluatie en op de uitgangspositie waarmee de school aan de bestudeerde zelfevaluatie begon. Over die gegevens beschikken we in deze context niet. Dat maakt het moeilijk om uitspraken te doen over de kwaliteit van de zelfevaluaties vanuit de schoolgerichte invalshoek. Moest vervolgonderzoek het mogelijk maken om ook de zelfevaluaties te documenteren die de bestudeerde scholen in de toekomst zullen uitvoeren, dan kunnen de gegevens uit 'zelfevaluatie binnenstebuiten' een basis vormen om in de toekomst ook het derde referentiekader te gaan hanteren.

De drie referentiekaders maken duidelijk dat de vraag 'hoe kwaliteitsvol zelfevaluaties in de bestudeerde scholen verlopen zijn' (op dit moment) niet eenduidig te beantwoorden is. Enkel het normgerichte referentiekader stelt ons in staat om onbetwistbare uitspraken te doen over de kwaliteit van zelfevaluaties in individuele scholen. Er is vooralsnog immers geen algemeen aanvaarde standaard en bovendien hebben we geen zicht op de historiek van kwaliteitszorg in de bestudeerde scholen. De standaarden die we suggereerden geven echter aan dat de bestudeerde zelfevaluaties in een groot aantal scholen niet of enkel in beperkte mate tot de bevraagde resultaten leidden. De voorzichtigheid omtrent de verwachting dat scholen over de nodige capaciteiten beschikken om tot kwaliteitsvolle zelfevaluaties te komen (Rudd & Davies, 2000; Van Petegem e.a., 2005b) blijkt in het licht van de kwantitatieve analyses gerechtvaardigd. Het schoolgerichte referentiekader zou dan weer suggereren dat geïnstrumentaliseerde zelfevaluaties op dit moment, gegeven het prille karakter van deze vorm van kwaliteitszorg, in veel scholen een aanvaardbare tot goede kwaliteit hebben. Merk hierbij op dat deze conclusies enkel van

toepassing zijn op de door ons onderzochte scholen en niet zomaar kunnen veralgemeend worden naar het gehele Vlaamse leerplichtonderwijs.

3.3 Hoe kunnen de verschillen in het verloop van zelfevaluaties verklaard worden?

De kwantitatieve analyses hebben aangetoond dat het verloop van zelfevaluaties in de bestudeerde scholen sterk samenhangt met hoe de school functioneert en met de kwaliteit van de ondersteuning die scholen genieten. Daarnaast stelden we grote verschillen vast tussen schoolleiders en andere teamleden.

3.3.1 Het effect van het ruimere functioneren van de school

De manier waarop scholen vormgeven aan zelfevaluatieprocessen komt niet zomaar uit de lucht vallen. De analyses in dit deel hebben duidelijk aangetoond dat de manier waarop zelfevaluatieprocessen vorm krijgen, in sterke mate samenhangt met de kenmerken van het ruimere functioneren van scholen. Het is weinig waarschijnlijk dat respondenten uit scholen met een beperkt beleidsvoerend vermogen, een beperkte organisatie-effectiviteit of met een beperkt lerend karakter aangeven dat de zelfevaluatie die in hun school plaatsvond tegemoet komt aan de zelfevaluatieprocesgebonden kenmerken. Een aanzienlijk deel van de variatie in de 'houding ten aanzien van zelfevaluatie' en in de mate waarin zelfevaluaties volgens de respondenten tegemoet komen aan de dragers van 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad' kan op basis van schoolgebonden kenmerken verklaard worden. Aspecten van de schoolcultuur (bijvoorbeeld de aandacht voor menselijke relaties, voor duidelijke afspraken en procedures, en voor vernieuwende impulsen) en het lerend karakter (de mate van diepgaand leren en kenniscreatie, en de houding ten aanzien van het eigen leren) blijken samen te hangen met de beoordeling van de respondenten over het verloop van de zelfevaluaties. Welke kenmerken precies samenhangen met het verloop van de zelfevaluatie is afhankelijk van het specifieke zelfevaluatieproceskenmerk waarop men focust. In het volledige verklaringsmodel stelden we vast dat slechts drie schoolgebonden kenmerken een statistisch significant effect hebben wanneer de correlaties tussen de zelfevaluatieprocesgebonden concepten mee in het model opgenomen worden. Deze zijn de score op het effectiviteitperspectief 'rationele doel', de score op 'diepgaand leren' en de score op 'kenniscreatie'. Zowel verschillen in schoolscores als verschillen in respondentscores worden in belangrijke mate door deze kenmerken van het ruimere functioneren van de school verklaard. Dat betekent dat zelfevaluatieprocessen kwaliteitsvoller verlopen naarmate men in de school volgens de respondenten meer belang hecht aan resultaatgerichte inspanningen, naarmate men de eigen waarden, normen en regels meer durft in vraag stellen en naarmate er meer sprake is van (verschillende vormen van) kenniscreatie. Scholen blijken in sterke mate te verschillen op deze kenmerken. Men mag dan ook niet van alle scholen verwachten dat ze van de ene dag op de andere op een even kwaliteitsvolle manier vormgeven aan zelfevaluatieprocessen.

3.3.2 Het effect van kwaliteitsvolle ondersteuning

Het verloop van zelfevaluaties wordt ook sterk bepaald door de ondersteuning die men geniet. Respondenten die zich positief uitspreken over de kwaliteit van de ondersteuning die ze genoten, zijn van mening dat de zelfevaluatie in hun school beter aan de zelfevaluatiegebonden kenmerken tegemoet kwam dan respondenten die geen ondersteuning of verhoudingsgewijs minder kwaliteitsvolle ondersteuning genoten. Ondersteuning blijkt echter niet altijd een positief effect te hebben. De score op 'zelfevaluatie als onderzoeksdaad' ligt voor respondenten die geen ondersteuning genoten immers hoger dan voor respondenten die verhoudingsgewijs minder kwaliteitsvolle ondersteuning genoten. Ondersteuning komt het verloop van zelfevaluaties dus enkel ten goede als ze voldoende kwaliteitsvol is. De impact van ondersteuning is positiever naarmate de ondersteuning er volgens de respondenten meer in slaagt om hen te overtuigen van het belang van het uitvoeren van zelfevaluaties; om hun kennis over (het verloop van) zelfevaluaties uit te breiden; om de mate waarin zij verbanden kunnen leggen en samenhangen begrijpen tussen aspecten van zelfevaluatie te vergroten; om hen in staat stellen om in de praktijk een kwaliteitsvolle zelfevaluatie uit te voeren en om hen ertoe aan te zetten om zelfevaluaties spontaan te gaan toepassen.

3.3.3 Het effect van de functie in de school

Schoolleiders spreken zich in sterke mate positiever uit over zelfevaluatiegebonden kenmerken dan teamleden. Onze veronderstelling is dat dit niet (alleen) moet worden toegeschreven aan een positievere kijk van schoolleiders op wat er in hun school gebeurt. Het is niet ondenkbaar dat schoolleiders zich positiever uitspreken omdat zij sterker bij het zelfevaluatieproces betrokken zijn en omdat zij in een positie zitten die maakt dat zij bepaalde aspecten van de zelfevaluatie wel vaststellen terwijl leerkrachten dat niet doen. Zij hebben bijvoorbeeld mogelijks een beter zicht op de communicatie- en inspraakinitiatieven die in het kader van de zelfevaluatie worden op poten gezet of zijn beter geplaatst om de link tussen de zelfevaluatie en de overkoepelende schoolvisie te leggen. De schoolleiding kan bijvoorbeeld naar aanleiding van de zelfevaluatie voor zichzelf duidelijke actiepunten vooropstellen (en ter harte nemen) die de andere teamleden niet aan den lijve ervaren, of waarvan ze zelfs geen weet hebben. Het is dan niet dat schoolleiders de zelfevaluatie anders ervaren, zij ervaren bij wijze van spreken gewoonweg een andere zelfevaluatie. Een zelfevaluatie die kwaliteitsvol verloopt volgens het beleidsteam in een school, kan volgens de andere teamleden dus niet kwaliteitsvol verlopen zijn. De hoge scores van schoolleiders hoeven dus niet noodzakelijk 'verdacht' te zijn. Het hoeft niet noodzakelijk om artefacten te gaan zoals vermeld in de probleemstelling (Watling & Arlow, 2002). De scores kunnen perfect aansluiten bij de realiteit van zelfevaluatieprocessen. Het is echter ook niet uit te sluiten – verder onderzoek in de context van zelfevaluaties dringt zich hier op - dat zelfrapportage van processen die men zelf aanstuurt wel tot een subjectieve en te positieve beoordeling leidt. Kyriakides & Campbell (2004) bestempelen geldigheid als het allergrootste probleem van zelfevaluaties. Zelfrapportage over zelfevaluaties zou in dezelfde logica met een probleem van

geldigheid kunnen te kampen hebben. Het is verder dan ook niet ondenkbaar dat beoordelingen door teamleden over processen die men – in grote lijnen - niet aanstuurt maar ondergaat eveneens tot een subjectieve maar dan te negatieve beoordeling leiden.

3.4 Hoe kunnen de verschillen in de resultaten van zelfevaluaties verklaard worden?

De resultaten van de kwantitatieve analyses zijn in sterk mate een bevestiging van de bevindingen uit de Delphi-studie. We stelden immers vast dat de kwaliteit van zelfevaluaties in sterke mate verklaard kan worden door de manier waarop de zelfevaluatie verloopt en door de kwaliteit van de ondersteuning die scholen genieten. Ook over de functie van de respondenten dient hier kort iets te worden gezegd.

3.4.1 Het effect van de zelfevaluatiegebonden kenmerken

Zelfevaluaties scoren hoger op de vooropgestelde kwaliteitsindicatoren naarmate de respondenten een meer positieve houding hebben ten aanzien van zelfevaluatie en naarmate zij vinden dat de zelfevaluatie meer tegemoet komt aan de concepten 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad'.

De impact van 'zelfevaluatie als beleidsdaad' is het meest uitgesproken en geldt in het volledige verklaringsmodel voor de drie kwaliteitsindicatoren die we vooropstelden. Wanneer men zicht heeft op hoe respondenten denken over de dragers van 'zelfevaluatie als beleidsdaad' in hun school, kan men gefundeerde voorspellingen doen over hoe zij de kwaliteit van de zelfevaluatie percipiëren. Uit het meest spaarzame model bleek zelfs dat het gebruik van de verkorte versie van het instrument om 'zelfevaluatie als beleidsdaad' te meten, quasi evenveel variantie kan verklaren in de kwaliteitsindicatoren dan meer complexe modellen. Uit het belang van 'zelfevaluatie als beleidsdaad' concluderen we ook dat de verwachtingen die we vooropstelden met betrekking tot het belang van (de dragers van) beleidsvoerend vermogen in de context van zelfevaluaties in deze studie dus bevestigd worden. Het ruimere functioneren van de school blijkt wel degelijk samen te hangen met de kwaliteit van zelfevaluaties, maar het effect is indirect. De effectiviteitsperspectieven in scholen, het beleidsvoerend vermogen en het lerend karakter van scholen zijn van invloed op de kwaliteit van zelfevaluaties precies omdat deze kenmerken samenhangen met de manier waarop zelfevaluaties in scholen vorm krijgen. Dat werpt een enigszins fatalistisch licht op zelfevaluaties in scholen. Kort door de bocht geformuleerd: de beste resultaten worden als het ware geboekt door scholen die er het minst behoefte aan hebben. In het afsluitende deel van dit proefschrift komen we hierop terug.

We stelden verder vast dat zelfevaluaties ook kwaliteitsvoller verlopen naarmate meer aan de dragers van 'zelfevaluatie als onderzoeksdaad' tegemoet wordt gekomen. Dat betekent bijvoorbeeld dat men goed nadenkt over de te evalueren thema's en dat men de inhoud van de zelfevaluatie in een ruimer kader situeert. Daarnaast blijken zelfevaluaties hoger te scoren op de kwaliteitsindicatoren naarmate de validiteit en betrouwbaar-

heid van de gegevens volgens de respondenten meer gewaarborgd is. In de conclusies bij de Delphi-studie omschreven we deze zaken als het realiseren van systematiek doorheen de zelfevaluatie. Het mag op het eerste zicht wat verwonderen dat bijvoorbeeld de betrouwbaarheid van een gegevensverzameling samenhangt met het komen tot gepaste actiepunten. Het is echter niet onlogisch dat de beleidsvoering in scholen baat heeft bij informatie die niet alleen betrouwbaar is, maar ook relevant is en correct geïnterpreteerd wordt. Bovendien is het niet ondenkbaar dat respondenten de resultaten van een onbetrouwbare gegevensverzameling minder vertrouwen en daardoor minder snel geneigd zijn deze als uitgangpunt te gaan nemen voor het vooropstellen van actiepunten.

Tenslotte zijn zelfevaluaties ook kwaliteitsvoller naarmate respondenten een meer positieve houding hebben ten aanzien van zelfevaluaties. Dat kan twee zaken beteken. Ofwel dragen zelfevaluaties in bepaalde scholen werkelijk minder bij tot het inlossen van verwachtingen en tot het komen tot gepaste actiepunten omdat de negatieve houding van de respondenten tegenover heel het zelfevaluatiegebeuren dat bemoeilijkt. Het zou echter ook kunnen dat zowel respondenten met een negatieve als met een positieve houding ten aanzien van zelfevaluatie de resultaten gekleurd percipiëren. Hoewel de resultaten er in de feiten wel zijn, betekent dit dat deze resultaten door respondenten met een negatieve houding ten aanzien van zelfevaluatie niet aan de zelfevaluatie toegeschreven worden. Daardoor wordt vervolgens het negatieve beeld van de zelfevaluatie bevestigd. Voor respondenten met een positieve houding geldt dan het omgekeerde. Ongeacht de achterliggende verklaring, de link tussen de houding van respondenten en de gepercipieerde kwaliteit maakt opnieuw duidelijk dat het niet vanzelfsprekend is tot kwaliteitsvolle zelfevaluaties te komen in scholen met een doorgaans negatieve houding ten aanzien van zelfevaluatie. Het Vlaamse onderwijs lijkt op dat punt met een probleem te kampen. In de beschrijvende analyses omtrent de houding ten aanzien van zelfevaluatie bleek dat zelfevaluatie bij de meeste teamleden niet geliefd is. Daarmee worden bevindingen van eerder onderzoek in Vlaamse scholen bevestigd (Devos & Verhoeven, 2003; Van Petegem e.a., 2005b; MVG, 2006). Men gelooft wel dat het proces tot waardevolle resultaten kan leiden maar men ziet het proces als dusdanig niet goed zitten. In het verbeteren van de kwaliteit van zelfevaluaties mag men zich daarom niet alleen richten op de vaardigheden en inzichten die een passend verloop van de zelfevaluatie vergen, ook de houding van mensen is cruciaal.

Zelfevaluaties kunnen maar een rol spelen in een kwaliteitszorgsysteem wanneer ze met een gewaarborgde kwaliteit plaatsvinden. Het bovenstaande impliceert dat dit ook duidelijke consequenties heeft naar de houding ten aanzien van zelfevaluaties en naar het verloop van de zelfevaluatie.

3.4.2 Het effect van kwaliteitsvolle ondersteuning

Respondenten die aangeven genoten te hebben van kwaliteitsvolle ondersteuning geven meer aan dat de zelfevaluatie in hun school kwaliteitsvol is verlopen. We concludeerden hoger reeds dat kwaliteitsvolle ondersteuning een indirect effect heeft op de kwaliteit van zelfevaluatie. Het volledige verklaringsmodel bracht echter ook aan het licht dat

ondersteuning ook een direct effect heeft, los van het effect via de zelfevaluatiegebonden kenmerken. Respondenten die ondersteuning genoten die bovengemiddeld kwaliteitsvol was, geven aan dat de zelfevaluatie in hun school hoger scoorde op 'resulteren in gepaste actiepunten' en 'resulteren in concrete resultaten' dan zij die aangeven dat de ondersteuning niet kwaliteitsvol was. Doorgaans scoren respondenten die geen ondersteuning genoten niet slechter dan zij die ondergemiddeld kwaliteitsvolle ondersteuning genoten. Het directe effect van kwaliteitsvolle ondersteuning kan wellicht verklaard worden door de grotere aandacht die mede onder impuls van de ondersteuning gegeven wordt aan het komen tot actiepunten en aan het benadrukken van en integreren van deze actiepunten in het bestaande schoolbeleid. Deze stap maakte in onze operationalisering immers niet expliciet deel uit van de zelfevaluatiegebonden kenmerken.

3.4.3 Het effect van de functie in de school

Schoolleiders spreken zich ook positiever uit over de kwaliteitsindicatoren van de zelfevaluatie dan teamleden. In wezen kunnen we hier dan ook dezelfde argumentatie opbouwen als bij het verklaren van verschillen in het verloop van de zelfevaluaties. Er is immers een sterk bruto-effect van de functie van respondenten. Onder meer in het volledige verklaringsmodel bleek evenwel dat het effect van functie op de kwaliteitsindicatoren verdwijnt wanneer gecontroleerd wordt voor zelfevaluatiegebonden kenmerken. Dat betekent dat schoolleiders die dezelfde opvatting hebben over het verloop van zelfevaluaties dan teamleden, niet hoger scoren op de kwaliteitsindicatoren dan deze teamleden. Wanneer schoolleiders en teamleden dezelfde houding hebben ten aanzien van zelfevaluatie en wanneer ze dezelfde mening delen over het verloop van de zelfevaluatie als beleidsdaad en onderzoeksdaad, zullen zij ook dezelfde mening toegedaan zijn over de kwaliteit van de zelfevaluatie. We moeten dus eigenlijk concluderen dat er in wezen geen netto-effect is van de functie die men in de school bekleedt.

4 AFRONDEND

Met dit 7de deel ronden we het bespreken van het kwantitatieve luik van deze studie af. In het verdere verloop ligt het accent op kwalitatieve gegevensverzameling en –analyse.

DEEL 8

GEVALSTUDIES

Om de variatie in zelfevaluaties verder te bestuderen en om de bevindingen uit de kwantitatieve analyse beter te kunnen interpreteren, werd ervoor gekozen om ook gevalstudies in de onderzoeksopzet op te nemen. In drie secundaire scholen bestudeerden we de wijze waarop zelfevaluatieprocessen vorm krijgen en de resultaten waartoe deze processen leiden. De gevalstudies (of ook casestudies) focussen op het verloop van zelfevaluaties. De onderzoeksvragen zijn daar doelbewust op toegespitst. Ze luiden: (1) 'Aan de hand van welke indicatoren kan men de kwaliteit van een zelfevaluatie operationaliseren?'; (2) 'Hoe kwaliteitsvol verlopen zelfevaluaties in Vlaamse scholen?'; en (3) 'In welke mate worden de resultaten van een zelfevaluatie bepaald door (a) het verloop van de zelfevaluatie en (b) de ondersteuning die scholen genieten?'

In de volgende paragrafen lichten we eerst de keuze voor deze onderzoeksmethode en het verloop van de gevalstudies toe. Daarna volgen de bevindingen. Daarbij maken we een onderscheid tussen een verticale en een horizontale analyse. We ronden af met een aantal conclusies.

1 WAAROM GEVALSTUDIES?

De aard van de onderzoeksactiviteiten tijdens gevalstudies en de resultaten waartoe deze activiteiten leiden, maken dat we van deze onderzoeksmethode een unieke bijdrage verwachten om tot inzicht in het verloop van zelfevaluaties te komen. We willen immers gedetailleerde kennis verzamelen over gedragingen, argumenten, motieven, veronderstellingen, opinies, interpretaties en ervaringen van mensen die men niet via schriftelijke vragenlijsten op het spoor kan komen. Met het uitvoeren van gevalstudies wilden we dan ook in de eerste plaats aandacht besteden aan de complexe sociale interacties en aan de betekenis die de deelnemers in de bestudeerde systemen geven aan elkaar en aan de gebeurtenissen binnen het zelfevaluatieproces. Swanborn (2000, p. 25) geeft aan dat er in gevalstudies aandacht besteed wordt aan de zogenaamde 'meervoudige realiteiten'. Deze zijn de verschillende en mogelijks zelfs tegengestelde visies van betrokkenen op sociale systemen en de verschillende interpretaties van gebeurtenissen die daarmee gepaard gaan.

Gevalstudies dragen bij tot het omvatten van de complexiteit van zelfevaluatieprocessen in scholen. De 'eenvoudige' modellen die ten grondslag liggen aan statistische analyses doen de complexiteit van het functioneren van scholen immers te kort. De meningen van respondenten worden in gevalstudies dan ook niet opgeteld, maar wel op hun onderlinge gelijkenissen en verschillen onderzocht. Een probleem bij het onderzoeken van cases is immers de generaliseerbaarheid van de resultaten. Cases worden namelijk niet op toevallige basis geselecteerd vanuit een bekende populatie (Yin, 1994). Het doel van het uitvoeren van gevalstudies is dus niet het generaliseren naar een populatie. NiederKofler (1991, in Swanborn, 2000, p. 67) schrijft in dit opzicht: "*The case-study investigator's goal is not to demonstrate the validity of an argument for statistical populations or universes. Rather, he aims to create and expand rich theoretical frameworks that should be useful in analyzing similar cases*." Als we op basis van deze gevalstudie willen generaliseren dan is dit enkel op het niveau van de theorievorming. Veralgemenende uitspraken hebben betrekking op relevante variabelen en verbanden.

In concreto streven we drie doelen na met het uitvoeren van de gevalstudies. In de eerste plaats willen we een aantal van de onderzoekshypothesen die we aan het eind van deel 4 vooropstelden toetsen aan empirische evidentie. In de conclusie bij de gevalstudies wordt daarom doelbewust vertrokken van de onderzoeksvragen van deze studie. Daarnaast willen we het theoretisch kader verrijken. We willen de validiteit van het theoretisch kader toetsen en ook aanpassingen of aanvullingen suggereren. Dit uit zich in een aantal theoretische reflecties aan het eind van dit deel. Tenslotte willen we met de gevalstudies 'vlees en bloed' geven aan het theoretisch kader en aan de bevindingen uit het survey onderzoek. De beschrijving van de resultaten van de gevalstudie 'illustreert' wat eerder in dit proefschrift uitgewerkt werd.

2 HET VERLOOP VAN DE GEVALSTUDIES

De omschrijvingen van wat een gevalstudie is, zijn - omwille van de verscheidenheid in mogelijke toepassingen - talrijk. Wij baseren ons op de definitie van Swanborn (2000, p. 22) die een aantal sleutelelementen samenbrengt. Deze definitie levert meteen ook een structurerend kader aan om de concrete invulling van de gevalstudie te beschrijven. Volgens Swanborn is een gevalstudie het bestuderen van een sociaal verschijnsel (1) bij één (of enkele) drager(s) van het verschijnsel, (2) in een natuurlijke omgeving, (3) over een bepaalde periode, (4) waarbij meestal diverse databronnen worden gebruikt, (5) gericht op een gedetailleerde beschrijving van tal van variabelen en het op het spoor komen van verklaringen van processen. In de volgende paragrafen bespreken we de invulling van deze kenmerken in de uitgevoerde gevalstudies. Aansluitend verduidelijken we hoe de onderzoeksgegevens geanalyseerd werden.

2.1 De selectie van casussen

In deze studie is er sprake van meervoudige parallelle casestudies. Drie secundaire scholen werden gelijktijdig en op dezelfde manier bestudeerd tijdens het uitvoeren van een zelfevaluatie. We hebben ervoor gekozen om de cases diepgaand te bestuderen en over de periode van het zelfevaluatieproces op te volgen. Met het opnemen van drie scholen is men in staat een onderscheid te maken tussen schoolspecifieke en algemene aspecten van het verloop van zelfevaluatie (Braster, 2000). Daarmee is echter niet gezegd dat het saturatiepunt met zekerheid reeds bereikt werd.

De keuze van deze drie cases is gebaseerd op theoretische overwegingen. In principe werden de cases geselecteerd uit de populatie van scholen die in het schooljaar 2005-2006 een geïnstrumentaliseerde zelfevaluatie planden uit te voeren. Om zo homogeen mogelijke scholen te zoeken, werd gekozen voor drie scholen voor secundair onderwijs en meer bepaald voor technisch en beroepsonderwijs. We hebben ervoor gekozen om de drie cases – binnen deze homogene groep - verder zodanig te selecteren dat er heterogeniteit op de onafhankelijke variabelen zou zijn. Omwille van de verwachte impact van het beleidsvoerend vermogen van de school (zie theoretisch kader), hebben we ervoor gekozen om scholen op deze parameters te laten variëren. Creswell (2005, p. 205) spreekt in dat verband van theoretische of conceptuele selectie. We onderscheiden drie gradaties in het concept 'beleidsvoerend vermogen': een sterk, een middelmatig en eerder zwak. Op die manier willen we sterker staan in onze conclusies met betrekking tot deze onafhankelijke variabele (Swanborn, 2000). Scholen werden op hun beleidsvoerend vermogen beoordeeld na overleg met een begeleider van deze scholen. Meer objectieve gegevens om het 'beleidsvoerend vermogen' van scholen te beoordelen, waren immers niet beschikbaar. Merk op dat we de verdeling in gradaties niet als een harde garantie beschouwden om te verzekeren dat de vooropgestelde driedeling ook op een zuivere manier zou teruggevonden worden. De bedoeling was te pogen om variatie in de gevalstudies in te bouwen. Uit een lijst van een 8-tal mogelijke cases werden uiteindelijk 3 scholen geselecteerd. De drie eerste scholen die wij contacteerden waren bereid in dit onderzoek mee te stappen.

2.2 De natuurlijke omgeving

We bestudeerden het verloop van zelfevaluaties in hun natuurlijke omgeving. Op die manier wilden we zoveel mogelijk vermijden dat eventuele relevante aspecten van het te bestuderen verschijnsel door het onderzoeksdesign van opname uitgesloten worden (zoals dat bij survey-onderzoek het geval kan zijn). We wilden een grote openheid verzekeren naar onverwachte variabelen en verbanden. Het respecteren van de natuurlijke omgeving kan echter moeilijk absoluut toegepast worden. Door respondenten uit hun context weg te halen en in een interviewsetting te brengen, wordt bijvoorbeeld al aan authenticiteit ingeleverd. Het komt er in wezen dan ook op neer de natuurlijke omgeving zoveel mogelijk in ere te houden. Zoveel mogelijk, want als onderzoeker maakt men altijd keuzes, kijkt men vanuit een bepaalde invalshoek en reduceert men de werkelijkheid tot die zaken die bijdragen tot het beantwoorden van de onderzoeksvraag (Swanborn, 2000). En toch, hoewel we de gevalstudies benaderen vanuit een referentiekader moet de toets aan

de natuurlijke omgeving ons als onderzoekers in staat stellen om dat referentiekader eventueel aan te passen en aan te vullen. Daarom hanteren we in deze gevalstudies een holistische aanpak. Dit betekent dat er vooraf zo weinig mogelijk geselecteerd wordt binnen de aspecten die mogelijk een rol zouden kunnen spelen; dat de werkelijkheid niet té vereenvoudigd voorgesteld wordt; en dat niet voorbijgegaan wordt aan het feit dat mensen betekenis geven aan de gebeurtenissen om zich heen (Swanborn, 2000).

2.3 De onderzochte periode

Het gaat in deze gevalstudies niet om een één-moment-opname. De cases werden gedurende de duur van een zelfevaluatieproces gevolgd. De duur van de onderzocht periode varieert echter in de verschillende cases. Zelfevaluaties verschillen immers in tijdsduur. De interviews startten in het najaar van 2005. In één school is de geplande zelfevaluatie uiteindelijk niet volledig doorgegaan. In deze school vond de tweede en laatste interviewronde reeds plaats in het voorjaar van 2006. Omdat in deze school het schooljaar voordien ook al een beperkte zelfevaluatie plaatsvond, werd in de gegevensverzameling ook op deze vroegere ervaringen gefocust. In de twee andere scholen vonden de vierde en laatste gesprekken plaats tussen december 2006 en maart 2007. De beschikbare middelen maakten het om praktische redenen helaas niet mogelijk om nadien nog gegevens te verzamelen. Merk overigens op dat de gevalstudies en het survey-onderzoek elkaar in de tijd overlapten.

2.4 De databronnen

De voornaamste bron van data waren de mondelinge diepte-interviews (Hutjes & Van Buuren, 1992; Denzin & Lincoln, 2000). De (semi-gestructureerde) diepte-interviews met schoolleiders en leerkrachten werden gepland bij de start, twee- of driemaal tijdens en eenmaal zo laat mogelijk in het zelfevaluatieproces. In de school waarin de plannen voor het uitvoeren van de zelfevaluatie stil vielen werden slechts twee interviewrondes ingelast. Naast de schoolleider of diens adjunct werd er telkens ook een leerkracht (of een lid van het middenkader) geselecteerd die sterk betrokken was bij het zelfevaluatiegebeuren en een leerkracht die er niet direct bij betrokken was. De interviews werden gehouden met het oog op het verzamelen van zowel algemene informatie over de school en de zelfevaluatie die er plaatsvond als van informatie over de eigen ervaringen en opinies van de respondenten. De interviewleidraad was identiek voor de verschillende groepen respondenten.

De casestudies werden verder ook geïnformeerd door een oriënterende analyse van relevante documenten. Tijdens één van de gesprekken werd aan de schoolleider de vraag gesteld welke documenten relevant zouden kunnen zijn om een beeld te schetsen van de school als geheel, van de organisatorische en culturele kenmerken van de school en van het verloop van de zelfevaluatie. Ook tijdens het verdere verloop van de gevalstudies werden uiteenlopende documenten verzameld. Daarnaast werd bij de onderwijsinspectie het meest recente doorlichtingsverslag van de scholen opgevraagd. Bij de selectie van op

te nemen documenten stond de vraag of deze een licht zouden werpen op de onderzoeksvragen centraal (Braster, 2000).

2.5 De onderzochte variabelen

Het onderzoeken van een sociaal verschijnsel houdt in dat bepaalde aspecten wel en anderen niet in het onderzoek worden betrokken. Toch starten we deze gevalstudies niet vanuit één expliciete theorie. We maken weliswaar gebruik van inzichten die in eerdere deelonderzoeken van deze studie zijn gegroeid. Het doel daarvan is het maken van gefundeerde keuzes in de vaststellingen die zich aandienen en het komen tot een passend interpretatiekader voor de onderzoeksbevindingen. Het gaat dus enerzijds om variabelen die het mogelijk maken om de (kwaliteit) van de zelfevaluatie te beschrijven en om variabelen waarvan we verwachten dat ze van invloed zijn op de kwaliteit van de zelfevaluatie. We focussen daarbij doelbewust op het verloop van de zelfevaluatie. Schoolgebonden kenmerken – zoals bijvoorbeeld het beleidsvoerend vermogen van de scholen - werden enkel in de rand van de zelfevaluatie en dus niet expliciet bevraagd. Aandacht voor schoolgebonden kenmerken zou immers ten koste gaan van de diepgang bij het bestuderen van het zelfevaluatieproces.

2.6 Analyse

Kenmerkend aan het gebruik van gevalstudies is dat de verzamelde gegevens zeer verscheiden en eerder onoverzichtelijk zijn. Over de analysetechniek moet dan ook zorgvuldig nagedacht worden.

2.6.1 Achterliggende uitgangspunten

Zo leek het ons bijvoorbeeld aangewezen niet eerst alle gegevens te verzamelen en pas dan te analyseren, maar ervoor te zorgen dat er een wisselwerking zou optreden tussen reflectie, waarneming en analyse (Wester, 1996, p. 70). Een methode die hier goed bij aansluit, is de *Grounded Theory*-benadering. Deze onderzoeksstrategie is er in wezen op gericht om nieuwe theorie te formuleren, en niet zozeer om de waarde van reeds geformuleerde theorie te toetsen.

De algemene lijnen van de Grounded Theory werden 4 decennia geleden opgesteld door Glaser en Strauss (1967). De naam 'Grounded Theory' verwijst naar het genereren van theorie uit data. Na enige tijd ontstonden er echter twee stromingen waarvan één onder leiding van Strauss en Corbin en de andere onder leiding van Glaser (Strauss & Corbin, 1990; Glaser, 1992). De invulling van Glaser is veel strikter dan deze van Strauss en Corbin. In hun meer recent werk beklemtonen deze laatsten het belang van de verschillende types van kennis die een onderzoeker reeds heeft voor de aanvang van het onderzoek en wordt aangegeven dat theorie analyses ook kan inspireren (Struebing, 1999). Glaser daarentegen houdt sterk vast aan de oorspronkelijke methodologische uitwerking van de Grounded Theory. Hij pleit voor een flexibel onderzoek, dat volledig geleid wordt vanuit

de informanten en hun sociaal geconstrueerde sociale realiteit (Babchuck, 1997) i.p.v. vanuit theorie om aldus theoretisch etnocentrisme in onderzoek te vermijden. Hij benadrukt de interpretatieve, contextuele en emergente aard van theorie-ontwikkeling (Goulding, 1999). Deze gevalstudies sluiten nauwer aan bij de visie van Strauss en Corbin, aangezien de dataverzameling en –analyse zowel geïnspireerd wordt door een (oriënterend) theoretisch kader als door de respondenten zelf, en dus niet louter en alleen door de inbreng van de respondenten, zoals Glaser vooropstelt. Grounded Theory wordt in deze studie opgevat als een benadering om systematisch naar kwalitatieve gegevens te kijken met als doel te komen tot verdere theorieontwikkeling.

2.6.2 Het analyseproces

Alle interviews werden op band opgenomen en uitgetypt in een tekstverwerkingsprogramma. De codes werden door twee onderzoekers toegekend. Dit gebeurde aan de hand van het softwareprogramma ATLAS.TI. Deze software maakt het mogelijk om codes toe te kennen aan interviewfragmenten, om de gegevens te ordenen en om de interviewdata later systematisch te doorzoeken. Zo wordt het mogelijk de gegevens te analyseren per thema, per school of per groep respondenten. Een collega-onderzoeker werkte zich eerst in het theoretisch kader in en deed het voorbereidend codeerwerk. Zij kende in een eerste ronde codes toe en contacteerde bij twijfel de eigenlijke onderzoeker in dit project. Deze laatste heeft de toegekende codes in een tweede ronde allemaal geverifieerd en deed in een beperkt aantal gevallen aanpassingen. Alvorens dit te doen werden alle interviews door hem opnieuw integraal doorgelezen. Elk interviewfragment werd dus in twee rondes en in wezen door twee onderzoekers gezamenlijk gecodeerd. Er werd voor deze aanpak gekozen om pragmatische redenen. Het vermoeden bestond dat de eerste onderzoeker inhoudelijk te weinig vertrouwd was met de onderzochte verschijnselen om een volledig gescheiden coderingsproces te verantwoorden. Door op de beschreven manier te werk te gaan was het echter toch mogelijk om tijdswinst te realiseren door de inzet van meer dan één codeerder.

De gehanteerde codes werden afgeleid van het theoretisch kader. Het theoretisch kader was meteen ook het analysekader. Om overlap en redundantie te vermijden werd gekozen voor eerder omvattende codes. De codes kunnen ondergebracht wordt in vijf families en soms in tussenliggende categorieën. De schuingedrukte begrippen in Tabel 84 geven de gebruikte codes weer. Tussen de haakjes is opgenomen hoe vaak een bepaalde code toegekend werd.

Tabel 84: Gehanteerde codes in de analyse van de interviews

1. *Zelfevaluatie als beleidsdaad (12)*
   - *Reflectief vermogen (34)*
     - *Houding ten aanzien van zelfevaluatie (51)*
   - *Gezamenlijke doelgerichtheid (116)*
     - *Initiatief voor de zelfevaluatie (32)*
     - *Verwachte resultaten en doelen (53)*

- *Doeltreffende communicatie (168)*
- *Onderst relaties en samenwerking (44)*
- *Gedeeld leiderschap (108)*
- *Responsief vermogen (19)*
- *Innovatief vermogen (8)*
- *Geïntegreerd beleid (12)*

2. *Zelfevaluatie als onderzoeksdaad (47)*
   - *Vastleggen inhoud (55)*
   - *Keuze zelfevaluatie-instrument (32)*
   - *Gegevensverzameling (97)*
     - *Bewaken betrouwbaarheid en validiteit (80)*
     - *Efficiëntie (34)*
   - *Interpretatie van de gegevens (37)*
   - *Komen tot actiepunten (68)*
3. *Resultaten van de zelfevaluatie (80)*
   - *Gerealiseerde resultaten en doelen (9)*
   - *Kwaliteit van de zelfevaluatie (53)*
4. *Ondersteuning (45)*
   - *(nagaan) behoefte ondersteuning (25)*
   - Gebruik van ondersteuning
     - *Beschrijving van de ondersteuning (28)*
     - *Beoordeling van de ondersteuning (33)*
5. (Achtergrond)kenmerken van de school en de respondent
   - *Functieomschrijving respondent (15)*
   - *Kwaliteitszorg in de school (32)*
   - *Vroegere ervaringen met zelfevaluatie (99)*
   - *Beleidsvoerend vermogen van de school (517)*
   - *De school als lerende organisatie (24)*

Het codeerproces is in eerste instantie een vorm van datareductie. De codering liet ons toe de informatie in overzichtelijke schema's samen te brengen. Op basis van deze dataoverzichten (per code) werden vervolgens conclusies getrokken (Miles & Huberman, 1994). Het uitschrijven van de bevindingen was geen louter technisch proces. Ook het komen tot een coherente aanpak en het zoeken naar een geschikte manier om betekenissen aan de lezer over te brengen, waren belangrijke aandachtspunten. We kiezen daarom voor verschillende perspectieven bij de verticale en horizontale analyses en voor een soms eerder verhalende aanpak.

3 BEVINDINGEN

In het analyseren van de bevindingen uit de gevalstudies hanteren we twee opeenvolgende benaderingen. In eerste instantie werd voor de drie cases telkens een samenvatting gemaakt met de belangrijkste vaststellingen. Per school werd dus eerst over de verschillende variabelen heen een beschrijving gemaakt. We noemen dit in navolging van Miles & Huberman (1984) de verticale analyse (cf. within-site analysis). Daarna werden de horizontale beschrijvingen over de scholen heen geanalyseerd (cf. cross-site analysis). Ook in het rapporteren hanteren wij deze structuur. Paragraaf 4 bevat de verticale analyse van

de bestudeerde scholen en is voornamelijk bedoeld als illustratie. In paragraaf 5 rapporteren we vervolgens over de horizontale analyse.

De aanpak van de rapportering is in sterke mate bepaald door onze wil om anonimiteit te waarborgen. Daarom gaan we eerst kort in op het probleem van anonimiteit dat opduikt bij het rapporteren van de resultaten van gevalstudies. Staessens (1991) maakt een onderscheid tussen externe en interne anonimiteit. De externe anonimiteit wordt bewaakt door de scholen door middel van pseudoniemen te 'benoemen' en te werken met fictieve namen. Daarnaast wordt contextuele informatie die de anonimiteit van de school in gevaar zou kunnen brengen niet opgenomen. Het bewaken van de interne anonimiteit (binnen de school) is veel moeilijker, ondanks het feit dat het slechts betrekking heeft op een beperkt aantal personen. In de bestudeerde scholen werden telkens drie personen met verschillende functies bevraagd. Dat maakt het zeer moeilijk om te garanderen dat de informatie die de respondenten aan de onderzoekers toevertrouwden ook effectief anoniem blijft wanneer een verticale analyse gerapporteerd wordt. Wolcott (1973, in Staessens, 1991) adviseert daarom om als onderzoeker vanuit het standpunt van de respondenten te bekijken of het opnemen van herkenbare informatie aanvaardbaar is. Alle informatie die de relaties van de respondenten met collega's zou kunnen bemoeilijken of die tegen de respondenten zou kunnen gebruikt worden, wordt dan ook best weggelaten (LeCompte, 1993). Wij zijn van oordeel dat het in deze context deontologisch ongepast zou zijn om in de verticale analyses met rechtstreekse verwijzingen of citaten te werken. In de gerapporteerde verticale analyse geven we om deze redenen een eerder feitelijke beschrijving van het verloop van de zelfevaluaties. Letterlijke en traceerbare verwijzingen naar de inbreng van individuele respondenten worden (zoveel mogelijk) vermeden. In de horizontale analyses – waar we over de drie scholen heen analyseren – zijn deze er wel. De citaten kunnen hier in principe immers uit meerdere scholen afkomstig zijn.

4 VERTICALE ANALYSE

In deze paragraaf geven we een beschrijving van de drie bestudeerde zelfevaluatieprocessen. Het gaat om de zelfevaluatie in het Technisch Instituut Voldegem, in het Technisch Instituut Baldegem en in het Sint-Odrada Instituut. Wij hebben er in deze fase voor geopteerd om zo dicht mogelijk aan te leunen bij de gegevensverzameling. Om een valide beschrijving te maken van het zelfevaluatieproces moet immers zo sterk mogelijk aangesloten worden bij het 'verhaal' van de respondenten. De gerapporteerde gegevens, opinies, interpretaties en waardeoordelen zijn telkens afkomstig van bevraagde respondenten of geraadpleegde documenten. Voor het Technisch Instituut Voldegem is de beschrijving volledig uitgewerkt (zie paragraaf 4.1). Na de uitgebreide verticale analyse van deze case presenteren we een synthese voor de twee andere cases. Deze beschrijvingen geven aan dat onze hypothese over de variatie in het verloop van geïnstrumentaliseerde zelfevaluatie bevestigd wordt (cf. hypothese 1).

4.1 Uitgebreide beschrijving van de zelfevaluatie in het Technisch Instituut Voldegem

De aandacht gaat in het volgende in de eerste plaats uit naar het beschrijven van het zelfevaluatieproces in de school. We hanteren daarbij voor deze illustratie een chronologische aanpak. Dat betekent dat we de beschrijving van de zelfevaluatie beginnen met het ontstaan van het zelfevaluatie-initiatief en eindigen met een reflectie op de resultaten. Aan de chronologische beschrijving van de aanpak van de zelfevaluatie gaan evenwel nog twee zaken vooraf. Eerst geven we een korte beschrijving van de achtergrondkenmerken van de bestudeerde school. Daarmee willen we de lezer voorzien van informatie over de context waarin de zelfevaluatie zich afspeelt. Om de anonimiteit te bewaken wordt een teveel aan details echter vermeden. Daarna staan we stil bij een aantal kenmerken van het functioneren van de school, voor zover deze relevant zijn om het verloop van de zelfevaluatie beter in kaart te brengen.

4.1.1 (Achtergrond)kenmerken van het Technisch Instituut Voldegem

Het Technisch Instituut Voldegem (TIV) is een school met een BSO/TSO-onderwijsaanbod en omvat de eerste drie graden van het secundair onderwijs. De schoolgebouwen bevinden zich naast de ring van Voldegem, dicht bij het centrum van de gemeente en in de buurt van een bekend bedrijf. De laatste jaren schommelt het globale leerlingenaantal rond 750. Ondanks inspanningen van de school om meisjes te rekruteren bestaat de leerlingenpopulatie voor 95% uit jongens. Er worden zowel voor TSO als BSO studiegebieden zoals auto, grafische technieken en mechanica-elektriciteit ingericht.

Het beleid van de school wordt uitgetekend en uitgevoerd door de directeur (algemeen schoolbeleid en administratief beleid), de adjunct-directeur (pedagogisch- en leerlingenbeleid), de technisch adviseur coördinator (technisch beleid en ondersteuning van het preventiebeleid) en het middenkader (vijf technisch adviseurs, preventieadviseur, graadcoördinatoren en een zorgcoördinator). Dit team wordt bijgestaan door de stafmedewerkers. Al deze betrokkenen zijn verenigd in wat men de 'beleidsgroep' noemt. Er worden regelmatig overlegmomenten ingelast. De directeur van het Technisch Instituut Voldegem (TIV) is tevens coördinerend directeur van de scholengemeenschap. De scholengemeenschap doet enkel wat ze wettelijk moet doen. De individuele instellingen en hun respectievelijke inrichtende machten werken volledig autonoom.

De visie van het Technisch Instituut Voldegem is verwoord in een algemeen geformuleerde missietekst die terug te vinden is in het schoolreglement en op de website van de school. De school wil een 'huis van vertrouwen' zijn door waarden voor te leven en het geven van kwaliteitsonderwijs. De school heeft volgens een recent doorlichtingsrapport het imago dat zij zich richt op alle leerlingen en ook de zwakkeren zal opvangen. Bovendien staat de school goed aangeschreven bij de industrie en het bedrijfsleven.

4.1.2 Het ruimere functioneren van de school

De school komt van een situatie waarin alles door het directieteam beslist wordt, maar wil evolueren naar een meer participatieve aanpak. Dit blijkt ook uit de interviews met schoolleiding en leerkrachten. Beide geven aan dat de openheid vanuit de directie er is, maar dat leerkrachten er (nog) te weinig gebruik van maken. Toch streeft men naar actieve betrokkenheid van alle inspraakorganen en participanten om een groot draagvlak te creëren. Volgens de directie is het echter, ondanks alle inspraakmogelijkheden, niet evident om iedereen te motiveren en blijft een duidelijke sturing vanuit het beleid noodzakelijk. De interviews leren dat het volgens hen niet volstaat om een perfect communicatie- en participatieplan te hebben. Zolang de 'wil' er bij sommige teamleden niet is, zal je hen – volgens de schoolleiding - toch niet kunnen bereiken.

De schoolcultuur laat veel ruimte voor een vrije en informele werking. Uit de interviews blijkt dat de mondelinge communicatie zeer open wordt gevoerd. De communicatie is direct en in het geval van conflicten vaak ook hard. De school benut heel wat informatiekanalen en media om haar participanten te bereiken, waaronder het communicatieplatform van een elektronische leeromgeving, het schoolkrantje en een verzorgde schoolwebsite. Elke dag wordt het team in het leraarslokaal ad valvas geïnformeerd over wat er te doen is; welke vergaderingen, bezoeken of studiereizen er zijn. Leerkrachten kunnen dit ook elektronisch raadplegen. Ouders en leerlingen worden over het dagelijkse schoolleven geïnformeerd via de schoolagenda's, de oudercontacten, persoonlijke briefwisseling en de website. Beleidsitems op het organisatorische en materiële vlak worden door de leraren in de vakgroepen besproken. Hun standpunten stromen door naar de pedagogische raad en de directie. Op personeelsvergaderingen wordt het schoolbeleid geregeld toegelicht. De school onderhoudt goede contacten met de lokale pers en werkt aan haar imago via een dynamisch PR-beleid. Ze komt zoveel mogelijk naar buiten met de initiatieven die ze neemt (bvb. de mini-onderneming).

De personeelsleden loven de aangename werksfeer op school en waarderen de open communicatie en de aanspreekbaarheid van de directie. Er heerst een collegiale sfeer. Het onderwijzend personeel ervaart voldoende inspraakmogelijkheden. Als sterk punt van de school geven zij de zorg voor de leerlingen aan, als zwak punt het soms minder verzorgde taalgebruik. Ook de ouders spreken zich naar verluidt positief uit over de school en wijzen op haar goede reputatie.

Door deel te nemen aan de vergaderingen van een aantal interne participatieorganen, kan het beleidsteam de complementariteit van de beleidsstructuren mee bewaken. Het gestructureerd overleg neemt toe maar er is enige overlapping van opdrachten en bevoegdheden, wat de transparantie niet ten goede komt. Ook over de verschillende studiegebieden heen is er niet altijd sprake van een geïntegreerd beleid. Personeelsleden in de verschillende studiegebieden lijken onderling op elkaar maar verschillen van de anderen (bijvoorbeeld 'hout' versus 'metaal'). Door de verschillen in aanpak ontstaan spanningen die samenwerking in de weg staan. De verschillende studiegebieden zijn ook fysiek van elkaar gescheiden en daardoor is er een traditie van verschillende groepen ontstaan.

De school probeert deze situatie door gezamenlijke pauzes te doorbreken maar de effecten daarvan laten vooralsnog op zich wachten. Een respondent geeft aan dat er eerder sprake is van een aantal stukken dan van een geheel.

Vernieuwingen in de school gaan bijna altijd gepaard met weerstanden. Een grote groep leerkrachten wil gerust gelaten worden en hun eigen ding kunnen doen. Na enige tijd schuiven ze wel mee op in de richting van het vernieuwingsdoel, maar daar gaat veel tijd overheen. Achteraf zijn ze dan vaak wel overtuigd van de meerwaarde van de nieuwe aanpak. Het komt er nogal eens op neer dat een grote groep volgt wanneer een aantal individuen overtuigd zijn. Daarnaast is er echter ook een kleine groep 'hardnekkigen', die niet (meer) mee willen. Dat geven zowel de schoolleiding als de leerkrachten aan. Deze kleine groep krijgt van collega's de stempel van 'bang voor verandering, lui en gemakzuchtig'.

Zolang leerkrachten hun twijfels, tekorten en fouten niet in groep moeten kenbaar maken, is men naar eigen zeggen bereid kritisch naar zichzelf te kijken. Leerkrachten reflecteren individueel of samen met (een) naaste collega('s). In een vakwerkgroep ligt het nog moeilijker, daar is men voorzichtig over wat men zegt. Daar durft men moeilijker over het eigen falen spreken. Er heerst volgens de respondenten in de school niet echt een cultuur die uitnodigt om te reflecteren op het eigen functioneren. Er is in de eerste plaats sprake van handelen en doen, en minder van reflectie.

4.1.3 Kwaliteitszorg en vroegere ervaringen met zelfevaluatie

Er worden in de school initiatieven genomen inzake kwaliteitszorg, onder meer ook door middel van bevragingen rondom verscheidene thema's (beleidsdag, briefing, enquête door de directiesecretaris, wensenformulier, jaarlijks evaluatieformulier in opdracht van het schoolbestuur, …). Dit alles gebeurt echter zonder systematiek en samenhang. De respondenten in de school erkennen deze beperkte systematiek in het kwaliteitszorgbeleid. Er gebeurt doorgaans geen gestructureerde analyse, interpretatie en evaluatie van de verzamelde gegevens. Men neemt wel initiatieven maar de opvolging van deze initiatieven laat nogal eens te wensen over. Er zit geen sterke visie achter, maar dit wordt niet als een probleem ervaren. Doordat leerkrachten zich goed voelen in de school, ziet men geen nood en geen behoefte om hier verandering in te brengen. Daardoor wordt vaak ook vanuit een probleemsituatie naar een oplossing gezocht.

Verschillende respondenten verwijzen bij de vraag naar vroegere ervaringen inzake zelfevaluatie naar een initiatief van een vier- à vijftal jaar geleden. De verantwoordelijke voor de personeelsadministratie voerde toen, een jaar voordat hij met pensioen ging, gesprekken met het voltallige personeel. Het was een man die het vertrouwen van de collega's genoot en waar men – naar eigen zeggen - veel tegen vertelde. Hij heeft daar toen een lijvig dossier van gemaakt, waaruit werkpunten werden gehaald die de basis waren voor toekomstige beleidsvoering. Het belangrijkste actiepunt dat uit deze zelfevaluatie voortvloeide was het verbeteren en uitbreiden van de bestaande overleg- en communicatiekanalen.

De houding van leerkrachten ten aanzien van zelfevaluatie zou volgens de respondenten niet negatief zijn, maar eerder afwachtend. De personeelsleden zijn er in principe niet tegen, maar de verwachting was dat zij het geheel van op afstand met een kritische blik zouden gaan volgen. Enthousiaste reacties werden dus bij de start niet verwacht.

We vermelden tenslotte ook nog dat de school nadat ze haar zelfevaluatieproces had opgestart vernam dat ze zou doorgelicht worden. De doorlichting en de zelfevaluatie overlapten elkaar dus in de tijd.

4.1.4 Het verloop van de zelfevaluatie

Deze paragraaf doorloopt chronologisch de verschillende stappen in de zelfevaluatie die plaatsvond in het Technisch Instituut Voldegem (TIV).

a. VAN EEN PRIL IDEE NAAR EEN BESLISSING

Het initiatief voor de zelfevaluatie kwam er naar aanleiding van een aanvoelen bij de directie dat het wenselijk zou zijn op een aantal domeinen wat bij te sturen. Het doel was te achterhalen wat de teamleden vinden van de school, te kijken waar men staat en te bepalen welke passende actiepunten zouden zijn. Dat gebeurt wel meer in de school, maar vaak worden daarbij geen of slechts een beperkt deel van de leerkrachten betrokken. Met het uitvoeren van een systematische zelfevaluatie bij alle teamleden wilde de schoolleiding verzekeren dat iedereen zijn stem zou kunnen laten horen. Op die manier hoopte men zicht te krijgen op wat er leeft bij alle teamleden en welke hun verzuchtingen zijn. Het was de wens van de schoolleiding om eens te bekijken of zij de zaken in hun school wel correct inschatten. Ook het sterk geëvolueerde personeelsbestand, met veel jonge nieuwkomers, was een stimulans om eens systematisch te gaan kijken wat er met de schoolcultuur gebeurd was nu de personeelsgroep zo anders is samengesteld dan vijf jaar geleden. In de bewoording van de respondenten ging het in de eerste plaats om de spiegelfunctie van zelfevaluatie. De school wou kijken 'waar ze stonden' en kijken wat er voor verbetering vatbaat was. Ook bevestiging krijgen van wat goed gaat was daarbij een verwachting. De doelen die men vooropstelde met de zelfevaluatie werden nergens formeel neergeschreven.

De school is dus duidelijk zelf de opdrachtgever voor de zelfevaluatie. De formele beslissing om een zelfevaluatie-instrument te gaan gebruiken werd in eerste instantie genomen door de directie, buiten de beleidsgroep om. Later werd de beleidsgroep met succes van deze keuze overtuigd. Er wordt door verschillende respondenten aangegeven dat de vraag naar een zelfevaluatie nooit spontaan vanuit het leerkrachtenteam zou ontstaan. De zelfevaluatie is in dat opzicht niet anders dan de andere beleidskeuzes die in de school genomen worden. Het is ook daar nog steeds de schoolleiding die in belangrijke mate aanstuurt.

De doorlichting was volgens de schoolleiders geen argument om de zelfevaluatie te starten. Men heeft er echter ook niet voor gekozen om de zelfevaluatie stop te zetten. De directeur en de adjunct-directeur keken uit naar de resultaten van beide vormen van

kwaliteitszorg. Ze geven aan de complementariteit van 'Wat vinden wij zelf?' en 'Wat vindt de inspectie?' waardevol te vinden. Niet alle teamleden zijn er evenwel van overtuigd dat beide van elkaar losstaan. Het luidt bij sommigen dat het wel héél toevallig is dat beide dit jaar samenvallen. De school kon bovendien vermoeden dat er een doorlichting zat aan te komen omdat ze in de tweede doorlichtingsronde – die in het schooljaar van de zelfevaluatie zou eindigen - nog niet aan bod was gekomen.

b. VAN EEN BESLISSING NAAR EEN INSTRUMENT

De keuze om een zelfevaluatie uit te voeren viel in het Technisch Instituut Voldegem zo goed als samen met de keuze voor DISO[33] (Diagnose Instrument Secundair Onderwijs). De school koos doelbewust voor een bevraging van het hele team aan de hand van een schriftelijke vragenlijst. Deze aanpak werd onder meer ingegeven door de twijfel over de weinig systematische werkwijze bij de vorige zelfevaluatie. Tot kort voor de beslissing om een zelfevaluatie te gaan uitvoeren was de schoolleiding niet op de hoogte van het bestaan van gebruiksklare zelfevaluatie-instrumenten. Toen de directeur op een studiedag toevallig vernam van het bestaan van DISO, was de keuze snel gemaakt. Het gebruik van dat instrument bleek immers samen te gaan met een begeleidingstraject waarop het TIV gratis kon intekenen. Bovendien bleek dat andere scholen er succesvol gebruik van maakten. De directeur en diens adjunct hebben deze keuze gemaakt. De mogelijkheid om met andere instrumenten te werken werd niet overwogen. Andere teamleden geven aan geen zicht te hebben op de motieven die de schoolleiding tot deze keuze voor DISO bracht.

c. DEELNAME AAN HET INLEIDENDE BEGELEIDINGSTRAJECT EN DE SAMENSTELLING VAN DE 'STUURGROEP ZELFEVALUATIE'

Van zodra de school besliste om zich te engageren voor het uitvoeren van een zelfevaluatie heeft zij zich ingeschreven voor een DISO-nascholingsproject. Dat was een doelbewuste keuze omdat men het gevoel had niet over de nodige expertise te beschikken om een zelfevaluatieproces zelfstandig tot een goed einde te brengen. In eerste instantie bestond de ondersteuning uit het bijwonen van nascholingssessies. De school nam daar met twee mensen aan deel. Beide geven achteraf aan daar niet veel aan te hebben gehad. Bij een volgende zelfevaluatie zouden ze daar niet meer aan deelnemen.

---

[33] DISO is een instrument en methodiek voor zelfevaluatie die aangeboden wordt door het VSKO (Vlaams Secretariaat van het Katholiek Onderwijs) maar in principe kunnen ook niet VSKO-leden er in mee stappen. Het instrument geeft geen concreet ingevuld kwaliteitsconcept maar is een instrument dat scholen helpt bij het in kaart brengen van het eigen gehanteerde kwaliteitsconcept. DISO wordt omschreven als een instrument dat op een aantal items het kwaliteitsconcept kan expliciteren en communiceren om het door de school te laten evalueren. Het instrument richt zich op de volgende 12 thema's: pedagogische visie, visie op leerlingen en leren, rol van de leraar, leerlingbegeleiding, onderwijsinhouden en didactische aanpak, klasorganisatie, evalueren en rapporteren, vakwerkgroepen, reglement, schoolontwikkeling, communicatie en directie, bestuur, beheer. Daarnaast biedt DISO ook nog plaats voor 20 schooleigen uitspraken.

Het begeleidingsproject omvatte verder onder meer ook een intake gesprek. In eerste instantie had de school gedacht de zelfevaluatie vorm te laten geven door de directeur, diens adjunct en een aantal technisch adviseurs (d.i. de beleidsgroep). De ondersteuners drongen er echter op aan om ook leerkrachten in dit proces te betrekken. Daarom werd een oproep gelanceerd naar geïnteresseerde leerkrachten. Daarop was in eerste instantie niemand afgekomen. In een volgende stap werden leerkrachten daarom persoonlijk aangesproken; ook leerkrachten die door het leven gaan als eerder 'stil'. Men wou ook weten wat zij van de opzet zouden denken. In de school klinkt het dat de stuurgroep bestaat uit 'Chinese vrijwilligers'.

Een uitgebreide werkgroep werd samengesteld, inclusief leerkrachten. Deze werkgroep ging door het leven als de 'stuurgroep zelfevaluatie' en zou instaan voor het ontwerpen van het verdere verloop van de zelfevaluatie. In de stuurgroep is iemand van de begeleidingsdienst de opzet van het werken met DISO komen toelichten. Deze toelichting werd inhoudelijk sterk gewaardeerd. Het verduidelijkte aan de ruimere groep wat de bedoeling was en hoe DISO zou gebruikt worden. Er wordt in de school opgemerkt dat de sturende maar gewaardeerde begeleiding maakte dat bij heel wat stappen geen fundamentele keuzes meer gemaakt moesten worden. Daardoor twijfelen sommigen zelfs aan het nut van de stuurgroep in het opzetten van het zelfevaluatieproces. De stuurgroep is overigens maar een beperkt aantal keer samengekomen tijdens het zelfevaluatieproces en leden geven zelf aan dat er weinig echt inhoudelijk werk moest verricht worden. Toch is men anderzijds ook wel tevreden dat men de raadgevingen van de DISO-begeleiders gevolgd heeft. Het beleidsteam geeft aan sommige zaken vanzelfsprekend te vinden terwijl deze dat voor leerkrachten niet zijn. Op zulke momenten is het belangrijk dat alle teamleden een plaats krijgen in het vormgeven aan de zelfevaluatie.

Alvorens de vragenlijst definitief te gaan gebruiken, heeft men in de stuurgroep wel overwogen om zelf een aantal bijkomende vragen te ontwikkelen. DISO biedt die mogelijkheid aan scholen. Uiteindelijk is men tot het besluit gekomen om dit niet te doen. Hoewel dit de verschillende betrokkenen zeer waardevol leek, is men er niet op ingegaan. Het instrument werd dus gebruikt zoals het standaard wordt aangeboden. De school kon zich sowieso vinden in de inhoudelijke invulling van DISO en vreesde het proces om zelf items te gaan ontwikkelen vermits men daar nog geen ervaring mee had. Bovendien is één en ander bij de start van het project nogal snel verlopen. Indien er meer tijd zou zitten tussen de keuze om aan zelfevaluatie te doen en het afnemen van de vragenlijst, dan had men van deze mogelijkheid wellicht wel gebruik gemaakt.

d. INFORMEREN EN SENSIBILISEREN VAN HET SCHOOLTEAM

Om het personeel te informeren over de zelfevaluatie werd een informatiemoment ingericht. Dat gebeurde voornamelijk op initiatief van de begeleiders en tijdens een personeelsvergadering die door iedereen verplicht was. Tijdens de informatievergadering ging men voornamelijk in op wat de school met het uitvoeren van de zelfevaluatie hoopte te bereiken. Er werd onder meer uitgegaan van de vroegere kwalitatieve benadering van zelfevaluatie (door de interviews) en de wens om nu een meer systematische werkwijze

te gebruiken. Er werd ook nog een brief rondgestuurd met informatie. Later laste men nog een extra informatiemoment in omdat bleek dat de school doorgelicht zou worden. De schoolleiding wou duidelijk aangeven dat er tussen de doorlichting en de zelfevaluatie een strikte scheiding zou zijn.

Teamleden die niet bij de planning van de zelfevaluatie betrokken waren, geven aan dat ze over de zelfevaluatie niet veel tot niets weten. Het thema kwam wel ter sprake in vergaderingen, maar buiten dat ze verwacht werden een vragenlijst in te vullen, bleef daar naar eigen zeggen niet veel van hangen. Teamleden hebben niet het gevoel dat men probeerde hen warm te maken voor deze zelfevaluatie. Van echte sensibilisering was geen sprake. Sommige leerkrachten geven aan dat er in hun school in het geheel geen communicatie gevoerd werd omtrent de vragenlijst in het kader van de DISO-zelfevaluatie, hoewel dat wel degelijk het geval was. Teamleden die actief deel uitmaken van werkgroepen in de school waren blijkbaar beter geïnformeerd.

Als er in het TIV al sprake was van informeren en sensibiliseren over de zelfevaluatie is dat in sterke mate op initiatief van de begeleiders geweest. Zonder begeleiding zouden de leden van het schoolteam niet zo sterk op de hoogte geweest zijn van de doelen en het verloop van de zelfevaluatie. Men vindt dat in de school naar eigen zeggen ook niet echt nodig. Zolang reflectie en zelfevaluatie zich richten op het schoolniveau verwachtte men immers geen problemen. De open cultuur in de school en de ervaring met enquêtes, maakten meer sensibilisering volgens de schoolleiding niet nodig. Net als andere leden van de stuurgroep zelfevaluatie geven zij aan dat teveel informatie niet hoeft. Ook teamleden die verder afstaan van de zelfevaluatie stellen dat er niet meer energie dan nodig moet worden gestoken in het informeren van het team.

De schoolleiding sluit niet uit dat er leerkrachten zijn die het gevoel hebben dat de zelfevaluatie er enkel komt omdat de directie dat wil. Niet alle leerkrachten staan er volgens de schoolleiding achter. Leerkrachten zouden daar echter 'geen spel van maken'. Zij doen dan gewoon wat van hen gevraagd wordt zonder hun ongenoegen te uiten.

e. DE EIGENLIJKE GEGEVENSVERZAMELING

De gegevensverzameling binnen een DISO-zelfevaluatie bestaat uit twee op elkaar aansluitende luiken: het invullen van de vragenlijst en de verhelderingsgesprekken.

- *Het invullen van de vragenlijst*

Het TIV heeft er voor gekozen om alle teamleden op dezelfde moment de vragenlijst te laten invullen. Men had het gevoel dat er één en ander zou kunnen mislopen als mensen de kans zouden krijgen om dit thuis te doen. De eigenlijke gegevensverzameling vond daarom collectief plaats op een vrijdagmorgen, de laatste dag van het eerste trimester. De leerlingen waren toen niet op school aanwezig. Men wou dit doelbewust niet op een avond doen en de laatste dag voor de vakantie leek ideaal om te zorgen voor een gepaste gemoedsrust bij het team. Eerst werd een half uur voor toelichting voorzien om duidelijk te maken hoe de vragenlijst ingevuld diende te worden. Omdat de meeste teamleden met de computer zouden werken, bleek dit nodig. Enkel de secretariaatsmedewerkers vulden

hun vragenlijst op papier in. Voor de afname werd de werkwijze ook al per brief toegelicht. Tijdens de afname waren er een aantal interne begeleiders voorzien die de teamleden konden bijstaan bij het invullen van de vragenlijst. Daarbij werd er op gelet dat de anonimiteit gewaarborgd bleef. Het afnemen van de vragenlijst bleek vlot te verlopen. Variërend van persoon tot persoon zou dit 30 minuten tot een klein uurtje geduurd hebben. Het instrument bleek duidelijk en de vragen waren helder geformuleerd. Er werd voor gekozen om de afwezige teamleden niet uit te nodigen om later nog deel te nemen.

De verwerking van de antwoorden gebeurde grotendeels door de DISO-begeleiders. Zij stelden een rapport samen met alle gegevens. Dit rapport beschrijft de resultaten per koppelvraag in volzinnen.

> Bij de koppelvraag...
>
> Het is op onze school (in mijn graad, afdeling, …) op dit ogenblik zo dat (a) leraren Technische Vakken en Praktische Vakken elk een eigen manier van omgaan met de leerlingen hebben en (b) alle leraren dezelfde visie hebben op het omgaan met leerlingen.
>
> Schrijft het verslag...
>
> Volgens een meerderheid van de leraren derde graad hebben leraren Technische Vakken en Praktische Vakken elk een eigen manier van omgaan met de leerlingen. In de 2de graad Mechanica gaat een minderheid hiermee akkoord, in de 2de graad Handel de helft van de leraren. Het is niet zo dat alle leraren dezelfde visie hebben op het omgaan met leerlingen.

Na een tweetal maanden werd het volledige rapport aan de school bezorgd. Tijdens deze periode werd er in de school niet verder gewerkt rond de zelfevaluatie. Eenmaal beschikbaar werd het rapport besproken in de stuurgroep van de zelfevaluatie. Daar werd bekeken hoe dit rapport de basis zou kunnen worden van de geplande verhelderingsgesprekken.

- *De verhelderingsgesprekken*

Verhelderingsgesprekken maken standaard deel uit van het werken met DISO en zijn een vorm van survey-feedback. De DISO-begeleiders spreken zelf van communicatieve validering. Dat betekent dat de resultaten van de vragenlijst aan de respondenten teruggekoppeld worden met de vraag of zij er zich in kunnen vinden. Daarbij gebruikt men de metafoor van een foto. De vragenlijst heeft als het ware geleid tot een foto die nog wat wazig is en het doel van het gesprek is deze foto te verscherpen.

Verhelderingsgesprekken gebeuren in kleine(re) groepen van maximaal 25 teamleden. Over de groepsamenstelling voor de verhelderingsgesprekken werd in de stuurgroep goed nagedacht. Er werd voor gekozen om de verhelderinggesprekken voor de mensen die algemene vakken geven uit te splitsen per graad. Voor de anderen werd per studiegebied gewerkt. Ook de schoolleiding en de opvoeders kregen een eigen verhelderingsgesprek. Het team werd over de doelen en de opzet van de verhelderingsgesprekken in een personeelsvergadering en per brief geïnformeerd.

De verhelderingsgesprekken vonden ruim vier maanden na afname van de vragenlijst plaats, op een moment dat de leerkrachten lesvrij waren. De gesprekken werden geleid door mensen van de DISO-begeleiding en duurden twee tot tweeënhalf uur. Alle vragen werden één voor één voorgelezen en er werd nagegaan of het beeld dat uit de schriftelijk bevraging naar voren kwam naar het oordeel van de subgroep in kwestie werkelijkheidsgetrouw was. Men kreeg dan de kans om één en ander te verduidelijken. Het was daarbij niet de bedoeling om nieuwe informatie te verzamelen of om verklaringen en beoordelingen te geven. Wanneer dit wel gebeurde en er discussie dreigde, grepen de begeleiders in. Er waren alleen leden van de subgroep en de begeleider aanwezig. De directie werd dus enkel in het eigen verhelderingsgesprek betrokken.

De leden van de stuurgroep verwachtten vooraf veel van deze verhelderingsgesprekken, hoewel men sterk opzag tegen het pak organiseerwerk dat ermee gepaard ging. Uiteindelijk hebben de verhelderingsgesprekken deze verwachtingen ook ingelost. Zowel de leerkrachten als de schoolleiding geven aan dat ze in de gesprekken bevindingen uit de vragenlijst recht hebben kunnen zetten. Men kon verduidelijken wat men precies bedoelde en kreeg als groep de kans om nuances aan te brengen. In sommige groepen leidde de verhelderingsgesprekken voornamelijk tot een bevestiging, in andere was er wel degelijk bijsturing van het tussentijdse verslag. Er werd zowel in de positieve als in de negatieve richting bijgestuurd.

De respondenten geven dus aan dat de verhelderingsgesprekken in hun opzet slaagden. Toch wordt aangehaald dat de termijn tussen de gegevensverzameling en de verhelderingsgesprekken iets te lang was. Men weet vaak niet meer precies wat men antwoordde. Bij het invullen van de vragenlijst zou men soms immers snel over items heen gaan. Positief was dan weer de achtergrond van de begeleiders. Die is volgens de respondenten van groot belang was om tot goede verhelderingsgesprekken te komen. Het zou belangrijk zijn om in de groep te passen. In een school als het Technisch Instituut Voldegem geldt dat bijvoorbeeld voor begeleiders met een achtergrond in het technisch- en beroepsonderwijs.

Op basis van de resultaten van de verhelderingsgesprekken werd door de begeleiders een nieuw, definitief verslag gemaakt. Dat nieuwe verslag voegde een eerder kwalitatief (narratief) luik toe aan de reeds bestaande cijfergegevens. Onder het thema 'Visie op leerlingen en leren (het onderwijspeil en omgaan met leerlingen)' lezen we bijvoorbeeld.

> Een grote meerderheid van de leraren vindt dat het onderwijspeil bepaald wordt door het peil van de leerlingen en de samenstelling van de klas. Een eerder kleine meerderheid van de leraren gaat ervan uit dat de leerlingen zich aan ons onderwijspeil aanpassen. In de 1ste graad en AV 2de/3de graad is het onduidelijk. Het leerplan schrijft een basispeil voor. Dat trachten alle leraren te respecteren. Volgens de leraren in de 1ste graad hangt veel af van het vak en het onderwerp. Wiskunde kent bijvoorbeeld een basispakket (met uitbreidingsmogelijkheden) voor alle leerlingen. Deze leraren maken ook een onderscheid tussen AV (waar leerlingen zich eerder aan het peil aanpassen) en TV/PV (waar het peil eerder aan de leerlingen wordt aangepast).

> Volgens een meerderheid van de leraren hebben leraren Technische Vakken en Praktische Vakken elk een eigen manier van omgaan met de leerlingen. Bij de leraren Handel is het onduidelijk. De leraren hebben een eigen lesstijl, persoonlijkheid. Dat vertaalt zich in verschillende werkvormen, evaluatievormen… Sommige leraren zijn soepeler dan anderen. Veel hangt af van de persoon (en zijn opleiding), de studierichting (TSO/BSO), het vak (hoofdvak of niet), de grootte van de groep, het aantal uren, het lokaal…

Bij elk thema geeft het verslag ook een aantal aanzetten voor reflectie, een aantal suggesties om rond het thema in kwestie te werken en een aantal inspirerende en verdiepende literatuurverwijzingen.

De school heeft aangedrongen op een snelle reactie bij het schrijven van het verslag. Ze zou namelijk graag beschikken over het verslag alvorens de doorlichting plaats zou vinden (mei 2006) en de examens zich zouden aandienen. De bedoeling was nog 'iets' met de resultaten te kunnen doen tijdens het lopende schooljaar. Uiteindelijk is dit niet gelukt en werd de interpretatie van de resultaten en van de planning van veranderings- en vernieuwingsactiviteiten uitgesteld tot het nieuwe schooljaar. In ieder geval beëindigde dit verslag van de DISO-begeleiders de fase van de gegevensverzameling. Voor de stuurgroep en de schoolleiding lagen op dat moment naar eigen zeggen twee belangrijke vragen op tafel: Wat moeten we hieruit leren en hoe gaan we dat terugkoppelen naar de rest van het team? Beide vragen worden hieronder verder uitgewerkt.

f. INTERPRETATIE EN PLANNING VAN VERANDERINGS- EN VERNIEUWINGSACTIVITEITEN

Over de manier waarop men van de zelfevaluatieresultaten zou komen tot een interpretatie en tot concrete actiepunten werd in de school niet met de teamleden gecommuniceerd. De vraag hoe deze stap aan te pakken heeft men zich bij de start en tijdens het begin van de zelfevaluatie ook niet gesteld. De volgende aanpak is dus eerder ad hoc ontstaan. Hoewel beide niet volledig kunnen losgekoppeld worden, maken we een onderscheid tussen de interpretatie enerzijds en het komen tot actiepunten anderzijds.

- ***De interpretatie van de zelfevaluatieresultaten***

Het zelfevaluatierapport speelde een cruciale rol in het interpretatieproces. We gaan daarom eerst in op de wijze waarop het verslag in de school onthaald werd. Daarna staan we stil bij de wijze waarop men ermee aan de slag is gegaan.

- ***Reflecties bij het zelfevaluatieverslag***

Over het zelfevaluatieverslag kwamen er verschillende bemerkingen. Het verslag is volgens de respondenten in de eerste plaats een eerder droge rapportering van de resultaten, zonder een verhaal er achter. Het is naar hun mening uitgebreid, omslachtig en langdradig geschreven. Bovendien zijn de zelfevaluatierapporten zeer beschrijvend en moet de stap naar interpretatie in de eigen schoolcontext nog gezet worden. Er wordt geopperd dat de begeleiding op deze manier niet veel toevoegt. Het verslag was volgens sommigen niet meer dan het in standaardzinnetjes uitschrijven van wat er in de tabellen

staat. Uit deze kritiek blijkt volgens ons dat de respondenten wellicht geen onderscheid maken tussen de twee verslagen die de school ontving (één voor en één na de verhelderingsgesprekken). Anderzijds wordt aangegeven dat het logisch is dat de begeleiders niet voor verklaringen en interpretaties kunnen zorgen. Daarvoor staan zij te ver van de school. In deze logica moet dit verslag in de eerste plaats betrouwbaar zijn en is het aan de school om er valide conclusies uit te trekken.

Hoewel het volledige eindrapport reeds snel op het interne netwerk beschikbaar was gemaakt, heeft geen enkele respondent weet van teamleden die het spontaan grondig bekeken hebben. Het was dan ook – voor schoolleiding en leden van de stuurgroep - duidelijk dat er nog een selectie van informatie moest plaats vinden alvorens dit naar de grote groep teamleden teruggekoppeld zou kunnen worden.

- *Het verloop van het interpretatieproces*

De school heeft het eindverslag opnieuw in de stuurgroep besproken. Dit gebeurde in mei 2006. Iemand van de DISO-begeleiding heeft de stuurgroep uitleg gegeven over hoe het verslag te lezen. De leden van de stuurgroep hadden het verslag vóór de stuurgroepvergadering volledig doorgelezen. Tijdens de stuurgroep schreef iedereen voor zichzelf neer wat er volgens hemzelf bij twee thema's uit het verslag het meest relevant was voor het huidig en toekomstig schoolbeleid. Deze werkwijze was een suggestie van de begeleiding. De leden van de stuurgroep vonden dit een moeilijke opdracht. In de stuurgroep werd verder ook nagedacht over hoe men de terugkoppeling van de resultaten naar het leerkrachtenteam zou kunnen verzorgen. Verschillende mogelijkheden werden afgetoetst. De intentie was om het DISO-verslag in de stuurgroep te interpreteren en om in de stuurgroep tot prioriteiten te komen. Het plan was om vanaf september 2006 wekelijks samen te komen om twee thema's te bespreken. Die strategie zou moeten leiden tot een degelijke samenvatting.

Het initiatief om het verslag te gaan vertalen en terug te koppelen naar de gehele school heeft sindsdien veel vertraging opgelopen, onder meer door de overlap met de doorlichting. Doordat de zelfevaluatie langer duurde dan gepland waren de resultaten van de doorlichting sneller beschikbaar dan deze van de zelfevaluatie. Beide evaluaties hebben veel informatie opgeleverd en het werd moeilijk om keuzes te maken. Gezien de aard en de herkomst van de gegevens heeft de schoolleiding prioriteit gegeven aan het doorlichtingsverslag. Het verder werken met de zelfevaluatieresultaten heeft daardoor langer op zich laten wachten. Het luidt dat vele teamleden hun handen al vol hadden met de resultaten van de doorlichting. Mentaal is dan ook in eerste instantie aandacht besteed aan de bevindingen van de inspectie. De zelfevaluatie kwam daardoor op de tweede plaats terecht. Reeds tijdens het proces van het interpreteren van de zelfevaluatieresultaten vreesde men dus dat een aantal bevindingen pas op langere termijn zouden kunnen worden opgenomen. De doorlichting en het feit dat het einde van het schooljaar naderde, legden een grote druk op de schoolleiding en het team waardoor belangrijk geachte stappen met betrekking tot de zelfevaluatie noodgedwongen een aantal maanden uitgesteld dienden te worden.

Een andere reden voor de vertraging was wellicht ook dat de leden van de stuurgroep de voorgestelde intensieve manier van werken niet goed zagen zitten. Uiteindelijk nam de adjunct-directeur het herschrijven van het rapport dan ook voor zijn rekening. Sommige leden van de stuurgroep waren daar lange tijd niet van op de hoogte en verwachtten nog steeds dat ze een rol zouden moeten spelen. Volgens de schoolleiding werden daar echter duidelijke afspraken rond gemaakt. De leerkrachten in de stuurgroep hebben nooit aangestuurd op acties van hunnentwege en geven aan blij te zijn dat de adjunct-directeur het werk voor zijn rekening nam. Ze zijn van mening dat een collectieve interpretatie een enorm complex proces zou worden.

De adjunct-directeur vertaalde het DISO-rapport van 40 à 50 pagina's in een dossier van een 6-tal pagina's. Het doel was te komen tot een bondige samenvatting van het beeld dat men heeft over de eigen school. De samenvatting bevatte ook bemerkingen en suggesties voor toekomstig beleid (zie Tabel 85). De samenvatting werd gaandeweg ook aan de stafmedewerkers voorgelegd om te kijken of het beeld strookte met wat volgens anderen het 'echte' verhaal van DISO was. De adjunct-directeur wou op die manier bewaken dat het niet alleen om zijn perceptie ging. Bovendien werd het middenkader meteen ook geïnformeerd over de belangrijkste bevindingen. De stuurgroep werd hier niet meer bij betrokken.

Tabel 85: Illustratie van de samenvatting voor 2 van de 12 thema's

Samenvatting verslag	Bemerkingen	Suggesties
Visie op leerlingen en leren		
• Onderwijspeil: peil wordt bepaald door de leerlingen en de samenstelling van de klas (meerderheid van de lkr zegt dit). Er is een groot verschil tussen lkr AV en PV/TV. • Omgaan met leerlingen: PV/TV lkr hebben een andere band met de lln doordat ze er meer uren komen. Ze hebben niet allemaal dezelfde visie op omgaan met leerlingen. • Doelen van het onderwijs: PV/TV lkr is vakkennis het hoofddoel. Dit wil ook zeggen vaardigheden voor het leven.	• PV lkr kunnen anders omgaan met de leerlingen.	• Overleg tussen collega's is hier zeker zinvol. Deze verschillende opmerkingen moeten daar aan bod komen. • Visie achter de leerplannen ontdekken. Externe hulp desnoods inschakelen. • Weten we van elkaar waar we mee bezig zijn. Verticale en horizontale leerlijnen. • Gedifferentieerd lesgeven. Kunnen we van elkaar leren?
Evalueren en rapporteren		
• Grote meerderheid vindt dat evaluatie gaat over kennis, vaardigheden en attituden. • Toetsen is een middel om de lln aan te zetten om te studeren. Deze geeft ook informatie over de leervorderingen. • Toetsen= reproductie van leerstof. • Afspraken zijn er binnen vakwerkgroepen. • Rapporteren gebeurt met cijfers en eventueel commentaar. • Het rapport wordt aangevuld na een klassenraad. De rest komt van de vakleraren.	• Als het meestal gaat over kennis en verwerking van geziene leerstof, waar gaan de lln dan zelfstandiger leren werken? • Rapportcommentaar: kan een ouder of lln met die opmerking iets doen (bvb. "X is lui, …" Worden er redenen gegeven?	• De lkr vinden het normaal dat men belang moet hechten aan vakinhouden en attituden, vaardigheden en inzichten. Maar moeten er dan ook geen afspraken komen hoe men die attitudebeoordeling moet gebruiken?

Over de kwaliteit van deze samenvatting durft de adjunct-directeur zich niet uit te spreken. Omdat het de eerste keer was dat hij een dergelijke interpretatie maakte, had hij het resultaat graag eens aan de ondersteuners voorgelegd. Hijzelf en het beleidsteam waren er weliswaar tevreden over, maar toch zou het geen kwaad kunnen moest de begeleiding ook bij de stap naar interpretatie en het komen tot actiepunten een actieve rol spelen. De DISO-begeleiders hadden wel aangegeven beschikbaar te zullen op vraag van de school, maar zij werden niet meer betrokken in deze fase van de zelfevaluatie. Het zou volgens de adjunct-directeur zinvol zijn om aan de interpretatie meer aandacht te besteden tijdens de voorbereidende samenkomsten. Of men zou een samenkomst kunnen uitstellen tot op het moment dat men in de scholen toe is aan het interpreteren van de resultaten.

Op een personeelsvergadering werd aangegeven dat de samenvatting van het zelfevaluatierapport op het elektronisch leerplatform te vinden zou zijn. Inhoudelijk werd het niet toegelicht volgens de teamleden terwijl de schoolleider aangeeft dat de samenvatting inhoudelijk volledig besproken werd. De oorspronkelijke plannen om de resultaten van de zelfevaluatie op een ludieke manier toe te lichten werden niet uitgevoerd. Er werd ook geen ruimte voorzien voor reactie en discussie over de resultaten. De schoolleiding was van mening dat teamleden in de verhelderinggesprekken reeds de kans kregen om te reageren en dat zij een tweede ronde overbodig zouden vinden.

- *Van geïnterpreteerde resultaten naar actiepunten*

Ook de stap naar actiepunten werd in eerste instantie gezet door de adjunct-directeur. Mocht hij zich er niet achter gezet hebben, dan zou DISO waarschijnlijk uit de aandacht verdwenen zijn. Om die reden zou men in de school nooit bewust kiezen voor een overlap tussen de zelfevaluatie en een doorlichting. Na de doorlichting heeft men naar eigen zeggen genoeg werk voor de volgende jaren. De resultaten van de zelfevaluatie zijn vrijblijvender en worden daarom minder snel ter harte genomen. Bovendien zijn de implicaties van de doorlichting voor de teamleden volgens de schoolleiding niet alleen concreter maar ook dwingender.

De suggesties naar aanleiding van de zelfevaluatie en van de doorlichting werden uiteindelijk door de adjunct-directeur geïntegreerd in stappenplannen. De schoolleiding vond dit aangewezen, gegeven de overlap tussen beide. Bij de start van de zelfevaluatie werd vooropgesteld dat de prioriteitenbepaling geen activiteit louter van de directie alleen mocht worden. De verantwoordelijkheid zou in de eerste plaats toekomen aan de beleidsgroep (met de (adjunct-)directeur en de technisch adviseurs). Uiteindelijk hebben zij inderdaad de grote lijnen van het stappenplan uitgetekend. Dat was eigenlijk in lijn met wat men volgens de adjunct-directeur vooraf van plan was. Men had zich voorgenomen om na de interpretatie van de zelfevaluatieresultaten weer te werken met bestaande raden. De stuurgroep die de zelfevaluatie vorm gegeven heeft, heeft dus ook in het komen tot actiepunten geen rol meer gespeeld.

Het stappenplan zet voor de komende twee schooljaren een aantal prioritaire taken voorop. Het gaat bijvoorbeeld om het organiseren van nascholing voor nieuwe leerkrach-

ten om hen in te wijden in het pedagogische project van de school, het ontwikkelen van studiewijzers voor de geïntegreerde proef, het organiseren van overleg over de nieuwe proefwerkenregeling, het organiseren van functioneringsgesprekken en de herziening van de werking van de cel 'leerlingenbegeleiding'. Naast de opsomming van deze taken geeft het plan ook een timing en een taakverdeling aan.

Het stappenplan zou kort na de laatste interviewronde op de personeelsvergadering door de adjunct-directeur voorgesteld worden. We spreken in deze alinea daarom verder over verwachtingen en plannen. Het plan was niet om bij de presentatie duidelijk aan te geven dat het stappenplan mede het resultaat was van de zelfevaluatie die heeft plaatsgevonden. Teamleden zouden immers niet wakker liggen van de herkomst van de prioriteiten. Op basis van het stappenplan plande men de bestaande werkgroepen (zoals beleidsgroep, de vakwerkgroepen, groepen die werken rond veiligheid en investeringen) de vraag voor te leggen welke actiepunten zij zouden kunnen ondernemen. Zij zouden allemaal de opdracht krijgen om ook zelf een aantal lessen te trekken uit de samenvatting en het stappenplan. Hun voorstellen zullen dan opnieuw naar de schoolleiding teruggekoppeld worden. Enerzijds om goedkeuring te geven, anderzijds om eventueel nog bij te sturen. Daarna zouden de werkgroepen van de schoolleiding groen licht moeten krijgen om aan de vooropgestelde plannen de komende 2 à 3 jaar prioriteit te geven. En dan zou het opnieuw aan de schoolleiding zijn om toe te zien op de realisaties. Want vanuit de doorlichting en een beetje uit de zelfevaluatie kwam naar voren, dat men van de directie verwacht dat zij daar nauwgezetter dan voorheen op zouden toezien.

De schoolleiding en de leden van de werkgroep wilden streven naar concrete actiepunten. Zij wilden bewaken dat de leerpunten niet té algemeen geformuleerd zouden worden. In welke mate de actiepunten op een specifieke, meetbare, acceptabele, realistische en tijdsgebonden manier uitgewerkt werden, konden we helaas niet meer in kaart brengen binnen de voorziene onderzoeksperiode.

4.1.5 De resultaten van de zelfevaluatie

De onderzoeksperiode laat het niet toe om ook de resultaten van het zelfevaluatieproces volledig in kaart te brengen. We beperken ons hier dan ook tot de resultaten op 16 maanden na de afname van de schriftelijke vragenlijst en op ruim 10 maanden nadat de school het zelfevaluatierapport ontving van de DISO-ondersteuners. We maken een onderscheid tussen de resultaten door de ogen van de teamleden en de resultaten door de ogen van de adjunct-directeur.

- *De resultaten door de ogen van de teamleden*

In de ogen van de teamleden, ook in deze van het teamlid dat deel uitmaakte van de stuurgroep, zijn de resultaten zeer beperkt. 16 maanden na het afnemen van de vragenlijst heeft de grote groep leerkrachten en ondersteunend personeel naar eigen zeggen nog geen elementen opgevangen waarvan ze vermoeden dat deze het resultaat zijn van de zelfevaluatie. Hoewel zij dit een gemiste kans vinden, geven ze aan niet wakker te

liggen van de beperkte resultaten. Toch blijkt het uitblijven van verdere stappen frustraties op te roepen. Ondertussen leeft de zelfevaluatie niet meer bij de teamleden. Er wordt verder niet meer over gepraat. De geïnterviewde teamleden verwachten dat de zelfevaluatie op den duur helemaal op de achtergrond gaat verdwijnen. Het duurt volgens hen allemaal veel te lang.

De teamleden zijn zich er niet van bewust dat de schoolleiding, weliswaar met enige vertraging, achter de schermen werkte aan een samenvatting en een stappenplan. De schoolleiding geeft aan dat de teamleden weldra wel het gevoel zullen hebben dat de zelfevaluatie iets heeft opgeleverd. De presentatie van het stappenplan stond immers kort na het laatste gesprek geagendeerd op de personeelsvergadering. Niet dat teamleden zich daardoor perfect zullen herinneren welke de resultaten van de zelfevaluatie zijn, maar ze zullen zich volgens de adjunct-directeur wel herkennen in de actiepunten die de beleidsgroep naar voren wil schuiven.

- *De resultaten door de ogen van de adjunct-directeur*

De adjunct-directeur daarentegen ziet wel reeds een aantal concrete realisaties van de zelfevaluatie. Dankzij de zelfevaluatie heeft hij een aantal eerder onverwachte signalen ontvangen omtrent zaken die niet optimaal verlopen. Daarnaast was er ook een bevestiging van een aantal vermoedens. De directie heeft naar eigen zeggen op basis daarvan en los van de bovenstaande stappenplanaanpak zelf reeds lessen getrokken uit de zelfevaluatie. Zaken die zijzelf direct konden bijsturen, werden niet via het stappenplan of bestaande werkgroepen aangepakt. Dat geldt bijvoorbeeld voor het creëren van eenvormigheid in het schrijven van verslagen en het opnieuw leven inblazen van de ICT-werkgroep. Dat gebeurde op de achtergrond en werd niet doelbewust onder de aandacht van andere teamleden gebracht als een resultaat van de zelfevaluatie.

Daarnaast zijn er ook een aantal andere bevindingen en resultaten die maken dat heel het zelfevaluatieproces volgens de adjunct-directeur de moeite waard is geweest. Hij vindt het belangrijk dat de teamleden hun stem hebben kunnen laten horen en dat ze zich hebben kunnen uitspreken over de pijnpunten die er volgens hen in hun school zijn. De zelfevaluatie heeft de aandacht gevestigd op nieuwe prioriteiten. Men wist bijvoorbeeld niet dat er zich inzake zorgbeleid problemen voordeden. De teamleden hebben volgens adjunct-directeur dankzij de zelfevaluatie meer het gevoel dat er naar hen geluisterd wordt. De zelfevaluatie heeft voor de schoolleiding daarenboven duidelijk gemaakt in welke mate bepaalde klachten al dan niet algemeen gedragen zijn. Naast de pijnpunten waardeert de adjunct-directeur ook dat er bevestiging is gekomen van de zaken die goed gaan in de school. Hij verwijst naar de losse informele schoolcultuur, de goede contacten tussen schoolleiding en teamleden en de goede band van leerlingen met hun leerkrachten. Nog voor het stappenplan in werking is getreden zijn er volgens de adjunct-directeur een aantal concrete verwachtingen, die hij bij de start van de zelfevaluatie vooropstelde, gerealiseerd.

4.1.6 De meta-evaluatie

Het Technisch Instituut Voldegem heeft in de voorbereiding van de zelfevaluatie niet stilgestaan bij de vraag hoe de kwaliteit van de zelfevaluatie zou kunnen worden beoordeeld. Zo werd bijvoorbeeld niet systematisch in kaart gebracht welke doelen de school wilde bereiken met het uitvoeren van de zelfevaluatie. Deze doelen konden dan ook niet als evaluatiecriteria gebruikt worden. Dat is doorgaans ook niet de manier van werken in de school. De vraag naar evaluatie stelt zich in alle beleidsactiviteiten pas op het eind, als ze zich al stelt. Aan het eind van de zelfevaluatie gekomen, geeft men aan ook in deze context aan zulk een reflectie geen boodschap te hebben. De schoolleider zou het wel eens in kleine groep willen proberen, maar vreest dat de teamleden daar niet mee lastig gevallen willen worden.

4.1.7 De toekomst

Het zelfevaluatieproces in het Technisch Instituut Voldegem was nog niet helemaal afgerond. Toch speelt de schoolleiding nu reeds met de idee om het proces over twee à drie jaar te herhalen. De adjunct-directeur vermoedt dat hij dan opnieuw zelf het initiatief zal nemen; en ook zal moeten nemen overigens. De teamleden zouden daarbij dan niet anders reageren dan deze eerste keer. Zijn overtuiging is dat de kaarten om een zelfevaluatie uit te voeren in de nabije toekomst niet noemenswaardig beter zullen liggen dan bij de start van deze zelfevaluatie, maar ook niet noemenswaardig slechter.

4.2 Synthese van de zelfevaluatie in het Technisch Instituut Baldegem en het Sint-Odrada Instituut

Op een gelijkaardige manier zouden wij voor de twee andere cases een verticale analyse kunnen presenteren. Uitgaande van het unieke karakter van elk zelfevaluatieproces zou dit overigens een leerrijk iets zijn. Om de omvang van de rapportering te beheersen, hebben we er evenwel voor gekozen om ons te beperken tot een korte synthese. De volgende twee paragrafen geven een korte beschrijving van de zelfevaluatie in het Technisch Instituut Baldegem en in het Sint-Odrada Instituut.

4.2.1 De zelfevaluatie in het Technisch Instituut Baldegem

Het Technisch Instituut Baldegem is een school in een grootstedelijke omgeving met een eerste graad en een TSO-BSO-bovenbouw. De school verzorgt onder meer de opleidingen verzorging, restaurant-keuken en chemie. Het leerlingenpubliek bestaat voornamelijk uit meisjes, vaak van allochtone origine.

Het was de adjunct-directeur van het Technisch Instituut Baldegem die het initiatief nam om in de school (opnieuw) een zelfevaluatie op te starten. Op eigen initiatief had hij het jaar voordien bij een beperkt aantal teamleden proefgedraaid met IZES (Instrument Zelfevaluatie van scholen). De spiegel- en beleidsvoorbereidende functies van zelfevaluatie

stonden daarbij centraal. In deze volgende fase wilde hij de zelfevaluatie meer uitvoerig aanpakken en er zoveel mogelijk personeelsleden – zonder verplichting - bij betrekken. Op die manier wilde de adjunct-directeur tot een meer representatief beeld komen en wilde bij evoluties in kaart brengen. Zelfevaluatie maakt in de ogen van de schoolleiding deel uit van een groter geheel van kwaliteitszorgactiviteiten in de school (men spreekt over zaken als visieontwikkeling, coaching, functiebeschrijvingen, lesobservaties). Evaluatie richt zich in de school op het globale beleid, op de schoolleiding en op de leerkrachten. De systematiek van het gebruik van een zelfevaluatie-instrument is volgens de schoolleiding complementair aan het informele van andere vormen van evaluatie.

De adjunct-directeur had kennis gemaakt met IZES tijdens een opleiding en vond dit een informatief en gebruiksvriendelijk instrument. In deze fase werd het inzetten van andere instrumenten niet overwogen, onder meer omdat er groot vertrouwen is in de degelijkheid van IZES. Voor het ontwikkelen van een eigen instrument vindt men dat de knowhow en de tijd ontbreekt. Deze keer zou overigens de elektronische versie van het IZES-instrument gebruikt worden. De adjunct-directeur had zich met het vooruitzicht op de eerste IZES-zelfevaluatie overigens ingeschreven in een nascholingstraject. Hij woonde de sessies tezamen met een collega bij. Bij de vorige zelfevaluatie werd voornamelijk de nazorg gewaardeerd. Daarmee doelt men op het gebundeld rapport van de resultaten met een aanzet tot interpretatie en met een duiding van de bevindingen. Ook de ervaringen van andere scholen konden de adjunct-directeur boeien tijdens de nascholingssessies.

Over de intentie om een nieuwe zelfevaluatie uit te voeren werd door de adjunct-directeur geen overleg gepleegd, ook niet met de algemeen directeur. Er was wel sprake van het informeren van het middenkader en de teamleden, zij het in het zeer beperkte mate. Ondanks de intenties om een uitvoerige geïnstrumentaliseerde zelfevaluatie uit te voeren, werden er in het Technisch Instituut Baldegem uiteindelijk geen concrete stappen in die richting gezet. Ook dat was niet het resultaat van een gezamenlijke besluitvorming. De adjunct-directeur heeft het proces, hoewel aangekondigd bij de teamleden, op eigen initiatief niet verder laten doorgaan. In die keuze speelden zowel de resultaten van de eerdere zelfevaluatie als toevalligheden een rol. De resultaten van de eerder uitgevoerde maar beperkte zelfevaluatie werden binnen het directiecomité besproken en aan het personeel gepresenteerd. Die resultaten bleken in het algemeen positief waardoor volgens de adjunct-directeur de behoefte om meer mensen te betrekken minder aan de orde was. Een toevalligheid was dat het bespreken van de eerdere zelfevaluatie door meer dwingende thema's naar de agenda van een latere personeelsvergadering moest worden geschoven, waardoor de vooropgestelde timing in het gedrang kwam. Een bespreking aan het eind van het schooljaar wilde men bovendien ten alle tijden vermijden. Daarnaast rapporteerde de adjunct problemen met het ICT-netwerk in de school waardoor de installatie van het zelfevaluatieprogramma niet kon plaatsvinden. Tenslotte ervoer de adjunct-directeur ook de competitie voor de beschikbare tijd. Daardoor kan volgens hem in de school te weinig tijd vrijgemaakt worden voor fundamentele reflecties. Deze redenen maakten dat de adjunct-directeur besloot de zelfevaluatie uit te stellen. De intentie daarbij was in de toekomst te komen tot een cyclische aanpak met het gebruik van een zelfevaluatie-instrument elke 2 à 3 jaar.

De leerkrachten in het Technisch Instituut Baldegem waren zich niet bewust van de intenties van de adjunct-directeur om opnieuw een zelfevaluatie uit te voeren. Naast het ter kennisgeving op een personeelsvergadering werden nog geen initiatieven genomen om over de opzet te communiceren en om het team te sensibiliseren. Moest de zelfevaluatie hebben plaatsgevonden dan was men wel van plan hier aandacht aan te besteden. Eerdere zelfevaluatie-ervaringen in de school leren volgens de respondenten immers dat de startfase cruciaal is. De bevraagde teamleden dragen initiatieven zoals het uitvoeren van een zelfevaluatie. Illustratief is dat er in sommige afdelingen eigen initiatieven bestaan om aan zelfevaluatie te doen. Die initiatieven zijn wel nog onvoldoende op elkaar afgestemd. Er in zowel inhoudelijk als qua timing sprake van overlap.

Het Technisch Instituut Baldegem is dus een school met ervaring inzake het uitvoeren van zelfevaluaties aan de hand van schriftelijke vragenlijsten. Eerdere zelfevaluatieresultaten werden volgens de schoolleiding bijvoorbeeld gebruikt tijdens pedagogische studiedagen; er werden werkgroepen opgericht en de beleidskeuzes die eruit voortvloeiden werden ingeschreven in het schoolwerkplan. Gegeven deze eerdere ervaringen werd in de gesprekken noodgedwongen gefocust op de zelfevaluatie die eerder had plaatsgevonden. De geplande zelfevaluatie die de school tot een prima casus maakte voor deze studie werd immers niet uitgevoerd. Deze zelfevaluatie was bovendien ook onvoldoende gekend bij de teamleden om er diepgaand op in te gaan. De case is onvolledig en kan verder niet zoals de anderen op een systematische manier beschreven worden.

4.2.2 De zelfevaluatie in het Sint-Odrada Instituut

Het Sint-Odrada Instituut is een school voor technisch en beroepsonderwijs gelegen in een stedelijke omgeving. De school heeft enkel een bovenbouw. Onder meer de opleidingen verkoop, verzorging, handel, en sociale en technische wetenschappen worden er aan geboden. De school rekruteert dan ook voornamelijk meisjes.

De school heeft 4 jaar geleden een zelfevaluatie uitgevoerd aan de hand van DISO (Diagnose Instrument Secundair Onderwijs). Het is daarbij altijd de bedoeling geweest om deze zelfevaluatie na een aantal jaren opnieuw te doen. Daar kwam bovenop dat men ook op het niveau van de scholengemeenschap rond zelfevaluatie begint te werken. Beide elementen waren in wezen een aanleiding om opnieuw een zelfevaluatie uit te voeren. Bovendien was de schoolleiding wat 'uitgekeken' op de prioriteiten waartoe de vorige zelfevaluatie had geleid. Het doel van de zelfevaluatie was dus naast de spiegelfunctie ook beleidsvoorbereidend. De schoolleider benadrukte ook herhaaldelijk zijn intentie om de teamleden op hun eigen verantwoordelijkheden te wijzen. Hij wilde de tekorten die de teamleden aangeven meteen koppelen aan de vraag welke rol zij voor zichzelf zien in de oplossing van het probleem. Zelfevaluatie werd dus ook in een legitimerende rol gebruikt. De schoolleiding wil zich immers richten op elementen die '*leerkrachten zelf aangekaart hebben*'. De zelfevaluatie werd door de adjunct-directeur op de agenda gezet. Moest hij dat niet gedaan hebben dan zou er volgens hem geen zelfevaluatie gekomen zijn.

Werken met DISO was in de vorige ronde volgens de schoolleider goed meegevallen. Zowel het instrument, de manier waarop het gebruikt werd als de begeleiding werden positief gewaardeerd. Bovendien wilde men opnieuw een zo representatief als mogelijk beeld van de school verkrijgen en gaan met het gebruik van DISO geen kosten gepaard. Het gebruik van een eigen of een ander instrument of van een andere methodiek werd daarom niet overwogen. Wel werden er aan het standaardinstrument een aantal eigen vragen toegevoegd. Dat gebeurde in eerste instantie op initiatief van de begeleiders. De uitwerking ervan gebeurde in de werkgroep die de zelfevaluatie coördineerde.

Bij de vorige DISO-zelfevaluatie hadden de directeur en de adjunct-directeur van de school een nascholingsprogramma gevolgd. Investeren in een nieuwe nascholing waarin een algemeen kader gepresenteerd wordt, leek hen deze keer geen meerwaarde te hebben. De school heeft daarom enkel een beroep gedaan op de on-site begeleiding. De begeleiders verzorgden het intakegesprek, gaven impulsen om de vragenlijst inhoudelijk aan te passen, zorgden voor de aansturing van de verhelderingsgesprekken en stelden de rapporten samen.

De school kent 12 werkgroepen die zich elk op een specifiek aspect van de school richten. Eén daarvan houdt zich doelbewust met kwaliteitszorg bezig. De coördinatie van de zelfevaluatie werd door deze groep waargenomen. Het is een werkgroep die bestaat uit vrijwilligers en uit mensen die er omwille van hun rol in de school deel van behoren uit te maken. De impact van deze stuurgroep is evenwel volgens de betrokkenen beperkt gebleven.

Teamleden die niet bij de 'zelfevaluatie-werkgroep' betrokken waren, hadden geen duidelijk zicht op de motieven om de zelfevaluatie uit te voeren. Zij werden tijdens het proces in wezen enkel geïnformeerd over de praktische zaken die het uitvoeren van een zelfevaluatie met zich meebrengt. Bij de eerdere zelfevaluatie was dit wel gebeurd, en daarom had men het aanvoelen dat dit niet meer nodig zou zijn. Relevante verslagen en rapporten zijn echter beschikbaar op aanvraag of op de elektronische leeromgeving, maar daar wordt door de teamleden geen gebruik van gemaakt. De zelfevaluatie is niet echt een thema dat leeft bij de teamleden. Zij voelen het aan als een activiteit van de schoolleiding waarin van hen een bijdrage verwacht wordt. Dat neemt niet weg dat de bevraagde teamleden wat hun eigen klaspraktijk betreft regelmatig vormen van zelfevaluatie inlassen, bijvoorbeeld binnen de vakgroepwerking.

De gegevensverzameling volgde het stramien zoals door DISO vooropgesteld werd. Dat betekent dat alle teamleden eerst een schriftelijke vragenlijst invulden. Er werd daarvoor een moment gereserveerd waarop alle leerkrachten op de school dienden te zijn, net voor de krokusvakantie. Er werd gebruik gemaakt een papieren versie van de vragenlijst. De verzamelde gegevens werden door de school zelf geturfd en aan de begeleiders bezorgd. Het verslag dat daar uit voortvloeide, vormde een drietal maanden later het uitgangspunt voor de verhelderinggesprekken. Daarbij werden teamleden in kleinere groepen onderverdeeld om de vaststellingen uit de vragenlijst te bespreken en eventueel bij te sturen. Beide vormen van gegevensverzameling zijn volgens de respondenten probleemloos verlopen. De opmerkingen die gemaakt werden hebben betrekking op de ver-

deling van teamleden in passende groepen en op de verwachte impact van een groep 'harde roepers' op het zelfevaluatieverslag. De beide gegevensverzamelingen vonden overigens plaats op momenten waarvan de teamleden aangaven dat ze ongelukkig gekozen waren. Daardoor was de betrokkenheid bij het geheel volgens hen beperkt.

Het syntheseverslag werd in eerste instantie besproken op een samenkomst van de begeleider met de directeur en de adjunct-directeur van de school. Daar werd beslist om het rapport op de laatste schooldag van het jaar op een personeelsvergadering toe te lichten. Op die samenkomst werden door de begeleider de resultaten van de zelfevaluatie uitvoerig toegelicht. Hoewel het de respondenten in het SOI verstandig leek om die resultaten door een intern iemand te laten presenteren, gebeurde dat omwille van tijdsgebrek niet. De leden van de stuurgroep bedankten voor deze taak. De presentatie van het verslag werd door de teamleden niet enthousiast onthaald. Dat is volgens hen toe te schrijven aan de inhoud van de presentatie (met teveel aandacht voor de 'droge' cijfers en te weinig voor verklaringen en mogelijke actiepunten) en aan de ongelukkige timing (iedereen was reeds in een vakantiestemming). De keuze om te focussen op de resultaten en niet op mogelijke actiepunten voor de toekomst was gemaakt door de schoolleiding. Deze wilde vermijden dat de functie van de coördinerende werkgroep ondermijnd zou worden. Het rapport werd verder ook integraal aan alle teamleden bezorgd, inclusief aan de leden van de stuurgroep en het werd ook ter beschikking gesteld op het elektronisch leerplatform van de school. De verschillende respondenten in het Sint-Odrada Instituut geven echter aan dat hier weinig of gaan gebruik van werd gemaakt en dat dit niet tot een vorm van interactie heeft geleid.

De intentie was om het rapport na de zomervakantie terug op te nemen in de stuurgroep die de zelfevaluatie coördineerde. Daarbij zou men op dezelfde manier tewerk gaan als bij de vorige zelfevaluatie. Dat houdt in dat men wilde komen tot een grondige bespreking en tot het vooropstellen van prioriteiten en werkpunten. Deze zouden dan vervolgens naar andere werkgroepen in de school worden doorgeschoven. De stap naar de bespreking in de stuurgroep werd in de eerste acht maanden van het nieuwe schooljaar nog niet gezet. Op dat moment werd de laatste interviewronde ingelast, wat impliceert dat de zelfevaluatie niet langer door de onderzoekers kon opgevolgd worden. Dat het rapport zo lang (10 maanden inclusief de zomervakantie) stil gelegen heeft is toe te schrijven aan het feit dat de adjunct-directeur geen initiatief heeft genomen om ermee aan de slag te gaan. Hij was naar eigen zeggen niet echt 'overtuigd' van de resultaten. Volgens hem bleek uit de resultaten te weinig dat er rond bepaalde thema's in de voorbije jaren hard gewerkt was en werden door de teamleden prioriteiten naar voren geschoven die hijzelf niet deelde. Daardoor was de motivatie om een initiatief te nemen er niet. De leden van de stuurgroep namen ondertussen een afwachtende houding aan. Voor de andere teamleden heerste er omtrent de zelfevaluatie volledige windstilte. Tijdens de laatste interviewronden leefde de verwachting dat men nog iets van de zelfevaluatie zou vernemen niet meer. Toch was de intentie om de resultaten grondig te gaan bespreken nog steeds aanwezig bij de schoolleiding. Vermits die stap voor het eind van de gevalstudies nog niet gezet werd, kunnen we over het proces van het komen tot actiepunten en over de concrete actiepunten hier echter niet rapporteren. Ook over een strategie om eventuele

actiepunten naar het team te communiceren en om de zelfevaluatie die had plaatsgevonden te gaan evalueren, werd in het Sint-Odrada Instituut nog niet doelbewust nagedacht.

5 HORIZONTALE ANALYSE

In de verticale analyses werden scholen als aparte en unieke eenheden benaderd. De horizontale analyse vergelijkt de scholen onderling. Onze keuze voor scholen voor technisch- en beroepssecundair onderwijs maakt dat de onderlinge homogeniteit voldoende groot is om onderlinge vergelijkingen mogelijk te maken. Door deze vergelijking willen we thema's en processen identificeren die in de verschillende zelfevaluaties terugkomen en willen we patronen in deze gegevens identificeren (Miles & Huberman, 1984).

Inhoudelijk ligt de focus in deze horizontale analyse op het verloop van de zelfevaluaties. Bevindingen over het ruimere functioneren van de school worden enkel opgenomen wanneer ze tot een beter inzicht leiden in de zelfevaluaties die hebben plaatsgevonden. In het rapporteren over het verloop van de zelfevaluaties maken we een onderscheid tussen 'zelfevaluatie als beleidsdaad', 'zelfevaluatie als onderzoeksdaad', 'resultaten van zelfevaluatie' en ondersteuning. Pas in de conclusies bij dit deel wordt expliciet verwezen naar de onderzoeksvragen. Toch was de vraag of informatie bijdraagt tot het beantwoorden van de onderzoeksvragen (zie einde deel 3) het fundamentele principe in het selecteren van inhouden.

De bewijskracht van de bevindingen berust voornamelijk op gegevens die tijdens de interviews verzameld werden. We werken daarom veelvuldig met citaten. De authenticiteit van citaten is een belangrijke indicatie van de validiteit van de gegevens (Kelchtermans, 1994). De citaten zijn daarom niet of minimaal bewerkt. Dat laatste gebeurde enkel om de leesbaarheid te verbeteren. De interviewfragmenten zijn steeds schuin gedrukt. Elk citaat wordt vergezeld door een code. Deze code verduidelijkt aan welke groep respondenten een citaat moet toegeschreven worden. Er werden immers drie groepen in de bevraging betrokken: schoolleiding (SL), bij de zelfevaluatie betrokken teamleden (TLZEV) en niet bij de zelfevaluatie betrokken teamledenleden (TL).

5.1 Zelfevaluatie als beleidsdaad

De verschillende dragers van 'zelfevaluatie als beleidsdaad' zijn niet mooi van elkaar te onderscheiden. Een levensecht en gecompliceerd proces zoals het uitvoeren van een zelfevaluatie is een holistisch gebeuren. Hoewel we de verschillende dragers in het volgende afzonderlijk presenteren, zal blijken dat een correct begrip enkel mogelijk is als men het geheel beschouwt. Omdat dit de structuur in de resultaten ten goede komt, presenteren we de dragers in een andere volgorde dan voorheen. Achtereenvolgens staan we stil bij reflectief vermogen, gezamenlijke doelgerichtheid, doeltreffende communicatie, ondersteunende relaties en samenwerking, gedeeld leiderschap, geïntegreerd beleid, responsief vermogen en innovatief vermogen.

5.1.1 Reflectief vermogen: omtrent de houding ten aanzien van reflectie en zelfevaluatie

Zelfevaluatie werd hoger omschreven als een systematische vorm van reflectie. We stelden de bereidwilligheid om zich wat betreft het eigen handelen en het handelen als team tot reflectie te engageren voorop als een voorwaarde om tot kwaliteitsvolle zelfevaluatie te kunnen komen. In deze paragraaf rapporteren we over de mate waarin de bestudeerde zelfevaluaties aan deze voorwaarde tegemoet komen.

a. HOUDING TEN AANZIEN VAN REFLECTIE EN ZELFEVALUATIE

De interviews met de verschillende leden van schoolteams leren ons in de eerste plaats dat de houding van schoolleiders ten aanzien van zelfevaluatie doorgaans (veel) positiever is dan deze van leerkrachten. De positievere houding van schoolleiders blijkt ook uit de grotere verwachtingen die ze vooropstellen bij het uitvoeren van zelfevaluaties (zie 'Gezamenlijke doelgerichtheid). We stellen vast dat de bevraagde schoolleiders er vaak ten onrechte van uitgaan dat hun teamleden het met hen eens zijn over de mogelijkheden van een zelfevaluatie. Hun veronderstelling is dat teamleden het uitvoeren van een zelfevaluatie wel genegen zijn en het waarderen dat ze bij de gegevensverzameling betrokken worden.

> *SL: 'Euhm, ik denk dat mensen dat wel op prijs stellen dat ze de kans krijgen om eens om de zoveel jaar, om eens mee het beeld van de school te bepalen. En is dat zoals zij vinden dat de school zou moeten zijn? Ik denk dat dat wel goed is, dat mensen daar hun zeg eens over kunnen doen. Dus in die zin, dat ze ook niet tegen een zelfevaluatie zijn.'*

Het vooruitzicht van een zelfevaluatie blijkt bij de bevraagde teamleden zelf echter andere reacties op te roepen.

> *TLZEV: 'Ik denk dat leerkrachten zoals ik ze ken, heel afwachtend, heel kritisch zullen reageren euhm… ja, zo een beetje afstand nemen. Ze gaan zich negatief opstellen denk ik. Dus niet direct van enthousiast.'*

> *TL: 'Leerkrachten hebben niet het gevoel dat ze daar zelf beter van worden, bijvoorbeeld, van het uitvoeren van die zelfevaluatie? Nee, het is zoiets verplicht.'*

Een positieve houding ten aanzien van zelfevaluatie gaat volgens de respondenten samen met een positieve houding ten aanzien van reflectie in het algemeen. Verschillende respondenten geven aan dat ze zich ervan bewust zijn dat er in hun school te weinig sprake is van reflectie. Uit de volgende citaten blijkt dat dit zowel te maken heeft met een negatieve houding ten aanzien van reflectie als met gewoontevorming.

> *TL: 'Reflectie, dat gebeurt hier praktisch nooit eigenlijk. Zo lang dat we goed bezig zijn, zo lang dat het draait en bolt, stelt niemand zich vragen. En iedereen werkt al jaren op dezelfde manier en dat is nooit niet een keer. (…) We houden ons alleen bezig met wat praktische punten eigenlijk. Maar echt zoiets diepgaands… of onze manier van werken, dat wordt ook nooit,…dat ligt zo ge-*

> *voelig, delicaat, we willen niemand voor de kop stoten denk ik. Ook de directie houdt zich daar een beetje vanaf zo.'*
>
> *SL: 'Denken, ja, dat beschouwen wij nogal eens als tijdsverspilling, en ja, dat is misschien toch verkeerd. Dus ik denk, we moeten ergens tussenin gaan zitten. Wij zijn te extreem daarin …'*

b. EEN CULTUUR VAN HANDELINGSGERICHTHEID

De bestudeerde scholen zijn scholen voor technisch- en beroepsonderwijs. Blijkbaar is er in de onderzochte scholen een sterke handelingsgerichtheid. Leerkrachten stellen het werken met hun leerlingen duidelijk centraal. Dat heeft prioriteit. Activiteiten op schoolniveau vinden de respondenten maar waardevol wanneer deze een duidelijke impact hebben op de activiteiten die met de leerlingen gedaan worden. Voor fundamentele reflecties en ook zelfevaluaties geldt dat volgens hen onvoldoende. Van een reflectiecultuur die zou bijdragen tot een positieve houding ten aanzien van zelfevaluatie is bij de bevraagde teamleden dus weinig of geen sprake. De houding van teamleden kan binnen een school echter sterk variëren. Het is niet zo dat een heel team pro, contra of onverschillig is. De houding ten aanzien van zelfevaluatie lijkt dus eerder een kenmerk van de teamleden dan van de school als geheel. Dat illustreren we aan de hand van de uitspraken van twee teamleden uit eenzelfde school.

> *TL: 'Ik denk dat we dat doen, omdat dat moet. En dat we daar dan wel eerlijk in zijn. Alhoewel, ik denk dat of ge dat nuttig vindt of niet, dat dat heel persoonlijk is. Ik heb er nu echt geen idee van hoe het lerarenkorps daarover denkt. Ik denk dat er mensen zijn, die zeggen dat dat wel nuttig zal zijn. Maar dat er evengoed zijn die zeggen, dat is tijdverspilling en dat zal zo 50/50 zijn. denk ik.'*
>
> *TL: 'Nee, nee. Dat leeft niet. Ze vinden dat niet belangrijk en ik denk dat ook bijna niemand echt geloofd dat er iets gaat veranderen. Denk ik. Het overgrote deel hé. Als je zo het woord DISO laat vallen, dan is dat van pfff … Dan is dat zo van ja, dat iedereen de indruk heeft dat ze (de schoolleiding) dat doen omdat ze dat… omdat dat goed staat en omdat ze dat moeten, allez ja, zo… ik denk dat daar niet veel belang aan gehecht wordt… de mensen zijn daar niet echt mee bezig.'*

In elk van de bestudeerde scholen is er een zeer kleine kern van leerkrachten die reflectie en vooruitgang bemoeilijken. Ze zijn zeer sterk gehecht aan hun eigen autonomie en aanvaarden het niet dat anderen ideeën lanceren die hen verplichten de eigen werkwijze in vraag te stellen, laat staan te veranderen. Het is volgens schoolleiders een kleine groep met een grote impact.

We geven ten slotte nog mee dat de handelingsgerichtheid van scholen zich ook uit bij het uitvoeren van een zelfevaluatie. Reflectief vermogen houdt immers ook in dat men ook over het zelfevaluatieproces als dusdanig reflecteert. Binnen de bevraagde onderzoeksperiode heeft geen enkele school echter een systematische evaluatie van de zelfevaluatie gedaan (cf. meta-evaluatie). Het zou schoolleiders echter geen slecht idee lijken dit wel te doen.

SL: 'In principe zouden we misschien nu een soort van evaluatie van het proces moeten doen. Geen zelfevaluatie IZES opnieuw, maar van kijk we hebben nu die IZES gedaan, wat vind je daar nu van? Terug een onderzoekje naar leerkrachten… van wat kunnen we daarin verbeteren? En zou je nog verder meewerken ? En zo van die dingen.'

c. *WAT HOUDT DEZE CULTUUR IN STAND?*

De gevalstudies zetten ons op het spoor van een aantal elementen die de afwachtende houding van de bevraagde teamleden ten aanzien van zelfevaluatie en reflectie kunnen verklaren. Zelfevaluaties blijken door de bevraagde leerkrachten bijvoorbeeld nog vaak als een vreemd iets te worden ervaren. Bovendien is dat gevoelen er niet alleen wanneer een zelfevaluatie voor de allereerste keer uitgevoerd wordt. Ook in een school met zelfevaluatie-ervaring stellen we dit vast.

TL: 'Wat ik gemerkt heb is, dat de zelfevaluatie, dit keer nog minder gezien wordt als iets dat van de leerkrachten is. Ok, die groep (de stuurgroep van de zelfevaluatie), ik denk dat die dat gevoel misschien nog wel hebben. Maar als je daar buiten gaat, dat mensen eerder zo iets hadden van: "Goh, nu moeten we daar nog voor naar school komen, na de uren. Om voor de directie weer een vragenlijst in te vullen.'

Ook de perceptie van zelfevaluatie als 'bedreigend' draagt bij tot de eerder negatieve houding. Sommige respondenten vrezen dat ze door een zelfevaluatie te zeer persoonlijk verantwoordelijk gehouden zullen worden voor de aanpak en het resultaat van hun activiteiten.

TL: 'Je zal toch nog een heel deel collega's hebben die eigenlijk dan hetzelfde zullen reageren als leerlingen. Allez ja, dat die het heel moeilijk hebben met zichzelf te evalueren. Die zoiets hebben van wat kan je daar nu mee aanvangen, wat ben je daar nu mee? Ze zien het nut er niet altijd van in… Het is ook, ik denk dat het op zich niet zo vanzelfsprekend is om aan zelfevaluatie te doen. Mensen zouden de vaardigheden moeten hebben om het te doen. Maar is dat het geval? Niet bij iedereen waarschijnlijk.'

Opvallend is dat men uitspraken zoals de bovenstaande doorgaans doet over anderen. De respondenten geven zelf aan geen of weinig problemen te hebben met reflectie en het erkennen van eigen fouten. Het zijn namelijk de 'anderen' die daarmee worstelen. De schoolleiders 'verwijten' op dat punt bijvoorbeeld de leerkrachten, wat de leerkrachten de schoolleiders 'verwijten'. Er bestaan op dat punt duidelijk vooroordelen. Reflectie zou evenwel voornamelijk moeilijk liggen wanneer het gaat om reflectie in groep zoals dat bij zelfevaluaties op schoolniveau vaak het geval is. Voor reflectie op individueel niveau of reflectie tijdens persoonlijke informele contacten blijkt de bereidwilligheid er bij de bevraagde teamleden wel te zijn. Dat heeft te maken met de gepercipieerde nood aan openheid en vertrouwen. Die is er in de bestudeerde scholen doorgaans wel bij bepaalde collega's of in bepaalde settings. Een uitgebreide zelfevaluatie verruimt deze 'veilige' settings echter waardoor men afwachtend of zelfs afwijzend reageert. Andere respondenten geven dan weer aan dat zelfevaluaties zich vaak richten op thema's op schoolniveau die voor de leerkrachten best veilig zijn. SL: '*Bij een zelfevaluatie praat je eigenlijk vaak*

zeer algemeen, en dan is dat allemaal geen probleem. Dat willen ze allemaal wel doen.' De houding hangt dus ook samen met de thema's die in de zelfevaluatie behandeld worden. Uitspraken doen over het leiderschap en de communicatie in de school zijn veel veiliger thema's dan de didactische aanpak van individuele leerkrachten.

Verder blijken ook praktische zaken de reflectie te bemoeilijken. Schoolleiders geven aan dat de dagelijkse beslommeringen geen tijd overlaten voor reflectie. *SL: 'Er is nooit tijd om te vergaderen'. SL: 'De tijd ontbreekt er voor.'* Toch zou één en ander langzaamaan aan het kantelen zijn.

De houding ten aanzien van zelfevaluatie wordt uiteraard ook bepaald door vroegere ervaringen met zelfevaluatie. De bevraagde respondenten geven bijvoorbeeld aan dat zelfevaluaties in het verleden niet tot voldoende concrete veranderpunten leidden en daardoor door het team als een vorm van tijdverlies geïnterpreteerd worden. Kwaliteitsvolle zelfevaluaties dragen volgens de respondenten bij tot een positieve houding; niet kwaliteitsvolle zelfevaluaties dragen bij tot een negatieve houding. Het is belangrijk hierbij te vermelden dat uit de interviews blijkt dat de perceptie omtrent het al dan niet resulteren in actiepunten sterk kan variëren. Een zelfevaluatie die voor de schoolleiding een aantal scherp afgelijnde aandachtpunten opleverde, kan door het leerkrachtenteam toch als een proces zonder resultaat herinnerd worden. Het uitvoeren van een zelfevaluatie draagt volgens de respondenten dus enkel bij tot een betere houding ten aanzien van zelfevaluatie wanneer men er in slaagt de resultaten van dit proces zichtbaar te maken voor het hele schoolteam.

5.1.2 Gezamenlijke doelgerichtheid: omtrent initiatieven, functies en verwachtingen

Alvorens stil te staan bij 'gezamenlijke doelgerichtheid' focussen we op het ontstaan van het initiatief om een zelfevaluatie uit te voeren in de bestudeerde scholen. Pas daarna onderzoeken we of men - en zo ja welke - verwachtingen men bij de zelfevaluatie vooropstelt. We onderscheiden twee invalshoeken. Eerst gaan we na of scholen doelgericht te werk gaan bij het uitvoeren van een zelfevaluatie. We brengen daarbij in kaart welke functies en concrete verwachtingen men bij de start van de zelfevaluatie koestert. Daarna onderzoeken we of deze doelgerichtheid gezamenlijk is.

a. HET INITIATIEF OM EEN ZELFEVALUATIE UIT TE VOEREN

Uit de analyses blijkt in eerste instantie dat het in de drie bestudeerde scholen telkens de schoolleiding was die het initiatief nam om een zelfevaluatie uit te voeren. *SL: 'Uiteindelijk zijn wij begonnen met een zelfevaluatie, omdat dat ons als directie een goed idee leek.'* SL: *'En wie heeft die beslissing genomen om dat op de agenda te zetten? Ikke.'* Het initiatief tot de zelfevaluatie is duidelijk niet van onderuit gegroeid. Het verhaal van zelfevaluatie start met een beslissing van de directeur of diens adjunct, al dan niet in onderling overleg. In de school waar het initiatief tot de zelfevaluatie uiteindelijk niet werd uitgevoerd, bleek de algemeen directeur geen weet te hebben van de motieven

die bij zijn adjunct tot het initiatief geleid hadden. *SL: 'Dat is Leo (naam van de adjunct) zijn initiatief geweest… Ik ben daar niet betrokken bij geweest.'*

De bevraagde schoolleiders geven aan dat het initiatief in ieder geval moet uitgaan van de school zelf. De school zelf moet volgens hen voldoende sterk gemotiveerd zijn om de zelfevaluatie tot een goed einde te brengen. Als de schoolleiding de zelfevaluatie enkel uitvoert om administratief in orde te zijn, zal dat volgens de schoolleiders negatieve implicaties hebben voor de zelfevaluatie. Hoewel schoolleiders uit de bestudeerde scholen aangeven dat het initiatief van hen kwam, zijn ze zich naar eigen zeggen bewust van het belang om ook de rest van de school voor het initiatief warm te maken. Dat betekent bijvoorbeeld dat de coördinerende stuurgroep of een bestaande werkgroep betrokken wordt bij deze keuze. Verder ging de betrokkenheid van overige teamleden bij het komen tot het initiatief van de zelfevaluatie in de bestudeerde scholen niet. De vraag naar zelfevaluatie zal ook volgens de teamleden zelf niet snel vanuit de leerkrachtengroep komen.

> *SL: 'Ja, de betrokkenheid van het korps bij heel dat proces, dat wil ik echt niet overschatten hoor. (...) Als ik euhm, het initiatief niet neem, de vergadering niet plan, als ik de agenda niet maak dan komt het er niet hoor. Dus allez ja, daarom vind ik het ook wel echt een moeizaam proces. Ge moet er aan sleuren, sleuren, sleuren.'*
>
> *TL: 'Ik denk niet dat leerkrachten dat uit zichzelf gaan doen. Dat denk ik niet. Ik denk dat dat opgelegd moet worden. Want anders doet ge voort en ge denkt dat ge goed bezig bent en je doet maar voort hé. Dat denk ik dus niet, dat dat iets is dat spontaan gaat groeien bij leerkrachten.'*

De vaststelling dat het de schoolleiding is die in de school het initiatief nam, impliceert niet dat de eigenlijke impuls in onze cases ook altijd van binnen de school kwam. Zo blijkt men op het niveau van de scholengemeenschappen ook aan zelfevaluatie te beginnen werken. In één school ging de impuls duidelijk van die hoek uit. Ook de link met een doorlichting blijkt niet onbelangrijk in het initiëren van een zelfevaluatie. Het is in sommige scholen een argument om de zelfevaluatie naar het personeel toe te verantwoorden. Dat stelt bijvoorbeeld de directeur van een school waar niet onmiddellijk een doorlichting gepland was.

> *SL: 'Wij hebben de leerkrachten toen ook gezegd dat zelfevaluatie iets is, dat door de inspectie, door de doorlichting ook echt verwacht wordt. Ik denk dat mensen dat ook al wel meegenomen hebben van, allez ja, tegenwoordig als school moet je dat ook wel doen.'*

In de twee andere scholen is de schoolleiding van mening dat de link met de doorlichting er niet is. In één van beide school vielen zelfevaluatie en doorlichting nochtans samen. Hoewel de schoolleider van die school de link met de doorlichting als motivatie om tot zelfevaluatie te komen ontkent, vindt de basis dit naar eigen zeggen 'verdacht'.

> *TL: 'Ik dacht wel dat het een beetje verband had. De doorlichting vraagt het waarschijnlijk ook dat we onze eigen school eens in vraag stellen. Dus daarvoor dat het nu misschien rap nog wat wordt georganiseerd. Dat heeft er wel verband mee natuurlijk.'*

Voor leerkrachten die verder van de zelfevaluatie afstaan, is het vaak gissen naar wat de aanleiding om in hun school een zelfevaluatie uit te voeren. Het initiatief komt volgens hen van 'bovenaf', zonder dat het volledig kan gekaderd worden.

b. DE DOELGERICHTHEID VAN ZELFEVALUATIES

Los van het initiatief staan de doelen en verwachtingen van respondenten ten aanzien van de zelfevaluatie. Deze doelen en verwachtingen werden in de bestudeerde scholen niet systematisch geïnventariseerd en neergeschreven. In de interviews komen ze echter wel ter sprake, en dit op twee manieren. Enerzijds in een eerder abstracte invulling. We spreken dan van de functies van zelfevaluatie. Anderzijds zijn er de concrete verwachtingen. Hoewel we beide hieronder afzonderlijk beschrijven, is er een inhoudelijke overlap tussen beide.

- *De functies van zelfevaluatie*

De respondenten verwijzen voornamelijk naar de spiegelfunctie en de beleidsvoorbereidende functie wanneer hen gevraagd wordt naar welke doelen ze bij de start van de zelfevaluatie vooropstelden. Men wil aan de hand van informatie over de eigen situatie zicht krijgen op het eigen functioneren en dat van collega's. De spiegel of foto waarin de zelfevaluatie resulteert, behoort volgens de respondenten een meer objectief oordeel over de school mogelijk te maken. Dat objectieve oordeel heeft men nodig om er vervolgens de beleidsvoering op te enten.

> *SL: 'Als wij dagdagelijks rondlopen, weten wij wel ongeveer waar dat het mis gaat. Maar dat is zowat meer intuïtie. En met deze zelfevaluatie wordt het een stukje formeler en systematischer vastgelegd. (...) Initieel was het wel zoiets van wij willen een beleid maken. Een beleid dat niet alleen geënt is op intuïtie of informele gesprekken links en rechts.'*

De beleidsvoorbereidende functie van zelfevaluatie komt sterk tot uiting in het belang dat respondenten hechten aan het komen tot verbeterpunten. Zelfevaluaties worden in de cases niet opgezet om te blijven focussen op wat er fout loopt. Ze zijn bedoeld om het in de toekomst beter te doen. Herhaaldelijk wordt aangegeven dat zelfevaluatie ook de positieve elementen van een school dient bloot te leggen. De metafoor van de spiegel impliceert niet dat men zich enkel naar de minder positieve aspecten moet richten. Ook het borgen van het goede is van belang volgens de respondenten.

> *TL: 'Want dat gebeurt ook weinig hé. Vooral het negatieve wordt besproken. Dus wat moet verbeterd worden, dus ja misschien… want wij hebben inderdaad ook sterke punten en die komen te weinig aan bod. Bijvoorbeeld dat de leerlingen hier graag naar school komen. Dat zou misschien ook wel interessant zijn, dat die ook eens ter sprake komen. Dat dat ook blijkt uit die zelfevaluatie.'*

Gezien de nadruk op de beleidsvoorbereidende functie van zelfevaluatie hoeft het niet te verbazen dat er verschillen zijn tussen scholen met en scholen zonder vroegere zelfevaluatie-ervaringen. De 2 scholen met zelfevaluatie-ervaring bekijken zelfevaluatie eerder als een onderdeel van een cyclisch proces waarbij op een gegeven moment nagegaan

wordt in welke mate de prioriteiten uit de vorige zelfevaluatie al dan niet gerealiseerd werden. Zij willen tot nieuwe actiepunten komen omdat de vorige actiepunten deel zijn gaan uitmaken van het beleid en/of omdat men er wat op uitgekeken is. Bij de school zonder ervaring is deze cyclische component minder duidelijk aanwezig en moet deze als het ware nog op gang komen. Beleidsvoering staat hier eerder in het teken van komen tot actiepunten dan van evalueren van actiepunten.

Volgens sommige teamleden hebben zelfevaluaties verder ook een legitimatiefunctie. Schoolleiders zelf zijn niet geneigd verantwoordingsgerichte functies toe te kennen aan een zelfevaluatie, tenzij wanneer die zelfevaluatie verplicht zou zijn door de overheid. Andere teamleden doen dat wel. Zij geven aan dat de functie van de zelfevaluatie in hun school ook bestaat uit het geven van rekenschap aan de onderwijsinspectie.

> *TLZEV: 'Ik heb ook de in druk dat die zelfevaluatie nu een beetje gebeurt - niet toevallig - omwille van de doorlichting. Dat ze ook kunnen zeggen, we zijn daar mee bezig. (...) Ja, er zijn nog een aantal zaken, waar we toevallig dit jaar mee starten, om dan toch goed te scoren eigenlijk.'*

- *Concrete verwachtingen bij de zelfevaluaties*

Onze invulling van kwaliteit is voldoen aan de verwachtingen. De drie functies van zelfevaluatie - spiegel, beleidsvoorbereidend, en legitimatie - worden door de respondenten omschreven in de vorm van een reeks uiteenlopende concrete verwachtingen. We formuleren hieronder een aantal criteria die door de respondenten zelf naar voren geschoven worden om als het ware de kwaliteit van de zelfevaluatie in hun school aan af te toetsen.

- Komen tot actiepunten. De verwachtingen gaan in de eerste plaats duidelijk in de richting van het komen tot actiepunten. Er wordt aangegeven dat een zelfevaluatie kwaliteitsvol is als concrete acties niet achterwege blijven. Een schoolleider antwoordt het volgende op de vraag hoe hij als buitenstaander de kwaliteit van een zelfevaluatie zou beoordelen. SL: '*Goh, ik zou dan zeker wel eens gaan kijken... niet naar het blinkend verslag van die zelfevaluatie, maar naar het hetgeen dat er mee gebeurt. En je kan een bevraging doen, maar uiteindelijk doe je dat ook maar in functie van acties die dat er uit gaan volgen. Als die achterwege blijven, ja dan, werpt dat toch wel wat ander licht op heel dat zelfevaluatieproces.'*

  Ook leerkrachten en het ondersteunend personeel benadrukken het belang van concrete actiepunten. Er moet iets met de resultaten gebeuren, dat is volgens hen cruciaal.

  > *TL: 'Er zou achteraf toch hier en daar een verandering moeten te zien zijn, want anders… Het is al zo dikwijls gebeurd dat we in de school een evaluatie doen. Maar daar blijft het dan ook bij. Er worden weinig concrete punten uitgehaald en dat zou in dit geval niet mogen gebeuren - denk ik - anders zou dat weer een slag in het water zijn. … Dus ik hoop dat er deze keer iets praktisch wordt besproken en dat de resultaten merkbaar zijn. Anders gaat men de mensen daar geen 2de keer warm voor kunnen maken. Dat is ook typisch denk ik. Leerkrachten zijn heel kritisch hé.'*

- Het gesprek aangaan met het gehele team. De respondenten – en de schoolleiders in het bijzonder - zien in zelfevaluaties een instrument om een communicatieproces met de andere teamleden op gang te brengen. Een grote groep van het schoolteam zou immers niet spontaan uiting geven aan ontevredenheid of zou suggesties voor veranderingen niet aan de schoolleiding kenbaar maken. Men ziet zelfevaluatie dus ook als een proces dat mensen inspraak geeft en verwacht dat het een aanzet is om te werken aan gezamenlijke doelgerichtheid in de school.

  SL: 'Op zich krijgen wij wel voldoende feedback, vanuit het middenkader, vanuit het directieteam. Maar vanuit de leerkrachten daar krijg je weinig, allez weinig… ze zullen wel eens zeggen dat loopt scheef of dat loopt scheef, maar de grote lijnen eigenlijk, krijg je daar niet van. Daar is een zelfevaluatie, denk ik, wel een nuttig instrument voor.'

- Komen tot legitieme en gedragen beleidskeuzes. Een zelfevaluatie geeft teamleden de kans om aan te geven hoe zij het schoolbeleid ervaren, wat zij belangrijk vinden en welke veranderingen aangewezen lijken. Schoolleiders geven dan ook aan de resultaten van zelfevaluatie te gebruiken als een legitimatie voor acties of innovaties. Zulke verwachtingen bij de schoolleiding ontgaan ook de andere teamleden niet.

  SL: 'Een zelfevaluatie, als directie, is ook wel in die zin iets krachtig, omdat ge kunt zeggen, het is wel iets wat jullie gevraagd hebben hé. Wij zitten als directie hier niet zomaar te zeuren, pedagogische studiedag na pedagogische studiedag, over bijvoorbeeld vernieuwende werkvormen.'

  TLZEV: 'Ik denk dat men van de directie uit ook wist dat er heel wat onderwijsvernieuwingen op komst waren. En ik denk dat men hier dus die boot niet wilde missen. En dat men dus ook ergens voor een stukje een medium heeft gezocht, denk ik, om dat zo druppelsgewijs in feite te laten doordringen.'

- Aantonen welke rol de verschillende teamleden kunnen spelen in het schoolbeleid. Men verwacht van de zelfevaluatie dat ze duidelijk maakt aan de betrokkenen dat ze zelf ook verantwoordelijkheid te dragen hebben in het ontwikkelen van het schoolbeleid. De zelfevaluatie moet helpen duidelijk maken dat men het initiatief tot verandering niet alleen bij anderen mag leggen. Zelfevaluatie wordt door een aantal respondenten dus ook gezien als een instrument om teamleden op hun eigen verantwoordelijkheden te wijzen.

  SL: 'Het gaat dus niet alleen om dingen die door de directie moeten gedaan worden, maar om 'wat kunnen wij als leerkracht daar in doen' ? Dat is voor mij wel echt het voornaamste. Als dat zou zijn, dan zou ik heel content zijn. (...) Niet gewoon aan de kant blijven staan, een lijstje invullen en aan de directie geven. Dat kan volgens mij niet de bedoeling zijn.'

- Oplossen van concrete problemen. Een concreet probleem is bijvoorbeeld de vaststelling dat het secretariaat niet meer naar behoren functioneert omwille van persoonlijke spanningen. Respondenten zijn van mening dat een zelfevaluatie informatie kan opleveren die kan bijdragen tot een oplossing. Toch worden op dat punt geen

onmogelijk zaken verwacht. Er is realisme wat betreft het oplossen van diepgewortelde problemen door het uitvoeren van een zelfevaluatie.

Zelfevaluaties moeten volgens de respondenten in de eerste plaats resulteren in concrete actiepunten. Daarover is er consensus. Daarnaast blijken verschillende respondenten toch ook een aantal specifieke verwachtingen te koesteren. Dat geldt voornamelijk voor de leden van de schoolleiding. Andere teamleden blijken doorgaans geen of enkel een beperkt zicht te hebben op de verwachtingen die men in hun school met het uitvoeren van een zelfevaluatie wil realiseren. Dat roept de vraag op in welke mate de doelgerichtheid algemeen gedragen is.

c. DE GEZAMENLIJKE DOELGERICHTHEID VAN ZELFEVALUATIES

In deze paragraaf gaan we na welke initiatieven men in de bestudeerde scholen nam om een draagvlak te creëren voor de vooropgestelde zelfevaluatiedoelen. Daarbij valt in eerste instantie op dat er weinig of enkel beperkte initiatieven genomen werden om de doelen ruimer kenbaar te maken. Doorgaans is er sprake van een toelichting op de personeelsvergadering en wordt de zelfevaluatie aangekondigd in de communicatiekanalen waarover de school beschikt. Deze initiatieven blijken bij de teamleden die niet bij de coördinatie van de zelfevaluatie betrokken zijn echter niet te leiden tot een duidelijk zicht op de vooropgestelde doelen.

> *TLZEV: ' Nu dat is een beetje naar mijn gevoel ongelukkig gegaan, in de zin het was het einde van het schooljaar euhm, de adjunct-directeur, is dat kort komen toelichten. Heeft dat daar neergelegd en dan was de vraag aan leerkrachten van vul dat in.'*
>
> *TL: '[vraag] Heb jij een idee van wat de redenen zouden kunnen zijn, van het directieteam om de beslissing te nemen om de zelfevaluatie te gaan doen?' Nee, eigenlijk niet. Nee, daar heb ik geen zicht op. Kan ik niet zeggen.'*
>
> *TL: 'Nee, ik denk niet dat de leerkrachten een goed zicht hadden op de doelen van deze zelfevaluatie. Ik denk eigenlijk niet dat ze goed geïnformeerd zijn over de… dat is wel op de personeelsvergadering meegedeeld, maar... omdat het hen niet interesseert denk ik. Ze zijn daar niet echt mee bezig zo, ver van mijn bedshow. Dat zijn dingen die moeten gebeuren, maar of ze er echt mee bezig zijn… ik heb niet de indruk.'*

De gebruikelijke kanalen blijken in de bestudeerde scholen dus niet effectief in het creëren van een gezamenlijke doelgerichtheid. Meer creatieve initiatieven om over de doelen van de zelfevaluatie te communiceren worden in de cases niet aangetroffen. Sommige schoolleiders vinden het ook niet noodzakelijk om hun bedoelingen bij een zelfevaluatie op een omslachtige manier naar het ruimere schoolteam te communiceren.

> *SL: 'Eerlijk gezegd voor onze school, denk ik zelf niet dat dat nodig zou zijn. Als we moesten zeggen van wij hebben een werkinstrument, wij leggen dat uit als directie. Wij leggen uit van mensen kijk, we gaan een zelfevaluatie doen. Daar moeten wij hier op school eigenlijk geen werkgroep voor samenstellen. Dat zou niet nodig zijn voor ons. (...) Ik ben er eigenlijk vanuit gegaan, dat men dat wel doorheeft. Zonder meer eigenlijk ja.'*

Voor sommige teamleden gaat deze redenering inderdaad op. Zij ervaren het niet als een probleem dat de schoolleiding doelen nastreeft die hen niet meteen aanbelangen. Deze situatie hoeft niet noodzakelijk negatief te zijn. Men mag niet uit het oog verliezen dat wat volgens één iemand een goede inhoudelijke invulling is van een zelfevaluatie dat niet hoeft te zijn voor anderen in de school. Verschillen in doelgerichtheid kunnen dus simpelweg toe te schrijven zijn aan zowel uiteenlopende interesses of aan de verschillende functies die teamleden in een school waarnemen.

> *TL: 'Omdat ik zeg van, voor de moment het is zelfevaluatie op schoolniveau hé. Mij interesseert het meer op klassikaal niveau. Dat zeg ik persoonlijk, daar heb ik persoonlijk zelf veel meer aan voor mee te werken. Om uw eigen te kunnen verbeteren, maar op dat niveau, ik denk niet dat ik daar in mee speel.'*

> *SL: 'Maar rond leerlingenbeleid bijvoorbeeld – omdat dat mijn taak is binnen de school – vind ik dat natuurlijk heel belangrijk. Dus daar vind ik mijn ding wel in terug, maar ik kan me voorstellen een technisch adviseur, dus eigenlijk de man die het technisch beleid doet… Die gaat daar minder zaken terug vinden, denk ik.'*

Andere teamleden zijn wel gevoelig voor de mate waarin de doelen van de zelfevaluatie door henzelf gedragen worden. Het onderscheid tussen beide groepen teamleden is ook hier te verklaren door hun mate van betrokkenheid. Zij geven aan dat het proces onvoldoende gedragen wordt door de grote groep. Dat heeft volgens hen gevolgen voor de mate waarin de actiepunten die uit de zelfevaluatie voortkomen door de teamleden ter harte zullen worden genomen. Ook schoolleiders komen, eenmaal het proces op gang getrokken is, tot deze vaststelling. Sommigen gaan hierdoor ook twijfelen aan mogelijke uitkomsten van de zelfevaluatie.

> *SL: 'Het is geen proces van de leerkrachten hoor, zeker niet. Het is eigenlijk een proces van de directie, met een paar voor de zelfevaluatie gemotiveerde leerkrachten, die er hun schouders mee willen onderzetten. Dus dat is nu eigenlijk wel wat… klinkt nu wel somber hé. Maar… ik geloof niet echt in dat resultaat. Ik denk ook dat ik van te voren kan voorspellen hoe dat het gaat uitdraaien.'*

Na de zelfevaluatie wordt communicatie over de zelfevaluatiedoelen door de verschillende schoolleiders als een aandachtspunt voor toekomstige zelfevaluaties aangezien.

> *SL: 'Wel ik denk dat we zeker in de beginfase, veel meer uitleg moeten geven over het systeem. Dus over IZES, dat is nu allemaal heel kort gebeurd. Allemaal via papierwerk en nota's enz. IZES werd wel toegelicht maar ik denk dat we dat… dat we daar veel meer tijd voor moeten gebruiken om zo'n systeem, of het nu IZES is of een ander, dat we veel duidelijker moeten motiveren waarom doen we dat.'*

5.1.3 Doeltreffende communicatie: omtrent het uitwisselen van ideeën, motieven en informatie

Zoals uit de drie cases kan blijken, is communicatie doorheen de zelfevaluatie van cruciaal belang. We maken hieronder een onderscheid tussen het communiceren over de doe-

len en het verloop van de zelfevaluatie enerzijds en communiceren over de resultaten anderzijds. Omdat we vaststellen dat de communicatie-initiatieven die men in de bestudeerde scholen op poten zet doorgaans niet doeltreffend zijn, proberen we ook te achterhalen waaraan dit kan worden toegeschreven.

a. Het communiceren over de doelen en het verloop van de zelfevaluatie

De meest opvallende vaststelling inzake communicatie is dat leerkrachten die verder afstaan van de zelfevaluatie vaak zeer beperkt of in het geheel niet op de hoogte zijn over het verloop van de zelfevaluatie. De bestudeerde zelfevaluaties blijken geen gespreksonderwerp te zijn tussen de teamleden. Vaak gebeurt dit ondanks de bestaande pogingen om hen te informeren. Er wordt in de cases immers wel degelijk over de zelfevaluatie gecommuniceerd. Het gaat dan eerder om informeren dan om bespreken. Informeren gebeurt voornamelijk via de personeelsvergadering, door middel van een brief of via de schooleigen informatiekanalen voor de leerkrachten (online platform of wekelijks informatiekrant). De scholen zetten dus geen speciale activiteiten op om de zelfevaluatie in hun school aan te kondigen en te verantwoorden. De bevraagde schoolleiders zijn bij de start van de zelfevaluatie overigens van mening dat het niet nodig is daar veel aandacht aan te besteden.

> *SL: 'Bij mij is dat zo van, oké, zeg dat gewoon tegen die mensen. Er is een enquête, daarvoor is het. Zijn er nog vragen en als er geen vragen zijn, is dat voor mijn part goed. Ik denk niet, dat er eigenlijk meer uitleg bij nodig is. Dat vind ik dan tijdverlies.'*

> *SL: 'Men is hier wel wat gewoon aan bevragingen. Maar nu gebeurt het via een officieel instrument en heb ik zoiets van dan wordt het een beetje opgeblazen. Dat gevoel heb ik. Want de leerkrachten zeggen ook van: "ja, goed". Sommigen zeggen ook: 'Moet daar allemaal zo'n tralalala van gemaakt worden? Moet al die uitleg?' Die stellen zich daar wel een beetje vragen bij.'*

De schoolleider van dit laatste citaat zet zich af tegen de suggestie van de ondersteuningsdiensten om wel degelijk doelbewust naar de teamleden te communiceren over doelen en inhoud van de zelfevaluatie. Na de zelfevaluatie zien de bevraagde schoolleiders dit anders. Dan geven ze aan dat ze, door bij de start beter te communiceren, meer uit de zelfevaluatie zouden kunnen hebben gehaald. Dat geldt ook voor de cases met zelfevaluatie-ervaring. Een eerdere zelfevaluatie ontslaat de coördinatoren van een zelfevaluatie dus niet van de communicatietaak die voor hen weggelegd is. Het gevaar bestaat immers dat men zaken vanzelfsprekend gaat vinden die dat in wezen niet zijn.

> *SL: 'Ik denk dat we er vorige keer misschien wel iets beter over gecommuniceerd hebben. Ik ben er zelf nu meer vanuit gegaan, van ja, kijk het is de 2de keer, we moeten niet terug uitleggen, van waar dat we gekomen zijn en wat dat we juist willen. Uiteindelijk hebben we dat vorige keer ook allemaal gedaan hé. Euhm, dus ik dacht dat het wel duidelijk zou zijn. Dat is misschien verkeerd gedacht, misschien hadden we der toch terug de grote communicatie rond moeten doen.'*

In de bestudeerde scholen geven de schoolleiders aan de volgende keer (nog) beter te willen communiceren bij de start van de zelfevaluatie. De uitdaging bestaat er volgens de respondenten in te communiceren wat teamleden willen en moeten weten om de zelfevaluatie tot een goed einde te brengen. Dat moet echter gebeuren zonder te vervallen in een ondoorzichtige stroom van informatie.

b. HET COMMUNICEREN OVER ZELFEVALUATIERESULTATEN

Over de resultaten van de zelfevaluatie wordt in de bestudeerde scholen doelbewust gecommuniceerd. De zelfevaluatierapporten worden bijvoorbeeld voor alle teamleden gekopieerd of op een elektronisch platform gezet. In de cases maken leerkrachten die niet bij de coördinatie van de zelfevaluatie betrokken zijn, echter geen gebruik van de mogelijkheid om het rapport op die manier in te kijken en eventueel te becommentariëren.

> *SL: 'Ik heb daar niks over gehoord. Ik heb dat laten kopiëren, zoveel bladzijden op 80 exemplaren, ik heb daar niks over gehoord. Ik vraag me ook echt af, hoeveel mensen dat daadwerkelijk gelezen hebben.'*

Het aanvoelen in de verschillende scholen is dat men als het ware verplicht wordt het zelfevaluatierapport te gaan herschrijven om de resultaten naar het team toe te brengen. Het is immers voor iedereen duidelijk dat het cijfermatige en volledige zelfevaluatierapport als dusdanig daartoe niet geschikt is. Het verslag dat de scholen van hun ondersteuners ontvingen, voldeed niet volledig aan de verwachtingen van schoolleiders. Precies omdat het volgens niet geschikt is om de communicatie met het team aan te gaan.

> *SL: 'Wij moeten dat in een bepaalde vorm gieten zodat we dat kunnen brengen. Als we dat dus in die vorm van dertig bladzijden gaan brengen, niemand gaat dat lezen. Niemand. Zo ken ik mijn mensen, die gaan dat niet doen. Dus op die manier gaat dat heel belangrijk zijn, dus hoe gaan we dat brengen … op een interessante manier, zelfs misschien op een speelse manier, we hebben dat vroeger ooit nog gedaan.'*

Zelfevaluatieresultaten worden om die reden ook op andere manieren gepresenteerd. Dat is volgens de respondenten nodig als men wil dat de resultaten door de teamleden ernstig worden genomen. In één school lichtte de adjunct-directur de resultaten toe op een personeelsvergadering. Verder werden er in die school geen initiatieven genomen. Op een andere school werden de resultaten van de zelfevaluatie op een pedagogische studiedag door een netgebonden begeleider toegelicht. Men was er in de school van overtuigd dat het eigenlijk beter zou zijn dat iemand van de school zelf de resultaten van de zelfevaluatie zou presenteren. *SL: 'Iets dat door eigen collega's gebracht wordt, dat zie je ook op pedagogische studiedagen, dat scoort altijd. Ik vind het belangrijk dat dat scoort hé. Ge hebt der uiteindelijk veel energie ingestoken.'* Toch werd er door tijdsgebrek toch voor gekozen om de presentatie door de begeleiders te laten verzorgen. De manier waarop dat gebeurde sloeg duidelijk niet aan bij het korps. Dat is zowel toe te schrijven aan de timing als aan de inhoud van de presentatie. Teamleden willen immers niet alleen de resultaten maar ook de conclusies horen.

> *TL: 'Ik had daar echt niets aan dat die begeleider daar stond cijfertjes af te lezen. En, euhm, wat die dan voorlas dat was bij mijn weten alleen maar wat wij*

toen in dat verhelderingsgesprek gehad hadden. (...) Dus ik had verwacht dat ze dan zouden zeggen, ja, jullie vinden dit. Maar dat ze daar nog een niveau zouden aanknopen van maar dat is omdat jullie een BSO-school zijn. Of dat is omdat hier voornamelijk meisjes zitten. Dus ik dacht dat zij dat nog gingen interpreteren en echt toepassen op de situatie. Maar dat gebeurde niet.'

c. *WAAROM DE COMMUNICATIE VAAK NIET DOELTREFFEND IS*

De interviews leren dat het niet volstaat om een perfect communicatieplan te hebben. *SL: 'Zolang de 'wil' er bij sommige teamleden niet is, zal je hen toch niet kunnen bereiken.'* Het ontbreken van de bereidheid om zich te informeren is wellicht echter niet de enige reden waarom teamleden zich niet informeren. Wij veronderstellen dat de communicatie in de bestudeerde scholen zich onvoldoende richt op wat de teamleden willen vernemen. Men richt zich immers voornamelijk op de praktische afspraken die voor leerkrachten relevant zijn. De coördinatoren van de zelfevaluatie communiceren wat volgens hen nodig is om de zelfevaluatie tot een goed einde te brengen. Men doet dat niet vanuit het perspectief van de grote groep teamleden, die bijvoorbeeld nog moet overtuigd worden van het belang van de zelfevaluatie. Men houdt als het ware te weinig rekening met de behoeften van de ontvangers van de informatie. Men vindt bovendien ook dat er eerst al iets op poten moet staan alvorens te informeren. De handelingsgerichtheid in de bestudeerde scholen steekt ook hier dus de kop op. Informeren met het oog op bijvoorbeeld gezamenlijke doelgerichtheid of participatie in de besluitvorming komt in de cases dus weinig voor. Aan de herkomst van het initiatief of de redenen waarom de zelfevaluatie wordt uitgevoerd, werd weinig of geen aandacht besteed. Het is niet ondenkbaar dat de gecommuniceerde informatie precies doordat ze van de doelen van het gebeuren losgekoppeld wordt door de teamleden vervolgens als onvoldoende relevant ervaren wordt. Dit kan verklaren waarom leerkrachten soms letterlijk niets kunnen vertellen over de zelfevaluatie die in hun school plaatsvindt, ondanks de zekerheid dat er naar hen toe over gecommuniceerd werd.

Sommige leerkrachten die in het geheel niet op de hoogte zijn van het verloop van de zelfevaluatie lijken er een enigszins dubbelzinnige houding op na te houden. Ze vinden enerzijds dat de informatie die ze krijgen vanzelfsprekend is en daardoor in zekere mate verloren moeite. Het luidt dan dat ze door een teveel aan informatie niet (meer) de moeite doen om het te doorgronden. De reacties op de communicatie over de zelfevaluatie op personeelsvergaderingen zijn illustratief in dat opzicht. *TL: 'Het gebeurt dat we achteraf zeggen: 'Van oké, we hebben hier nu wel gezeten, een half uur of een uur. Maar wat hebben we nu eigenlijk gehoord? Omdat hetgeen dat daar gezegd wordt, kan op 10 minuten ook gezegd worden, dat echt relevant is voor u als leerkracht eigenlijk.'* Anderzijds is men achteraf – na de zelfevaluatie - wel snel om aan te geven dat men vindt dat men onvoldoende geïnformeerd werd. Deze teamleden leggen de verantwoordelijkheid van hun beperkte achtergrondinformatie duidelijk buiten zichzelf. Anderen geven aan dat ze daar ook zelf een verantwoordelijkheid in dragen.

TL: 'Ik denk dat de mensen die echt in die stuurgroep zitten, daar wel bewust mee bezig zijn. Maar dat leerkrachten die daar niks mee te maken hebben, dat

> *ook niet echt weten of tenzij met deeltjes… Ik denk niet dat ik anders ben dan mijn collega's daarin. Maar ik denk, je krijgt de kans wel. Want de verslagen staan op internet, ze hangen uit in de leraarskamer. Dus als ge leest, weet je het wel hé.'*

De volledigheid gebiedt ons ook om aan te geven dat niet alle leerkrachten een beperkte communicatie als een probleem ervaren. Niet communiceren werd door één van de respondenten ook een als een te overwegen strategie verdedigd.

> *TL: 'Ik denk dat leraars nogal kritisch zijn. Ik heb dat al meermaals ondervonden. Dat het misschien beter is van het niet te doen. Dus, want anders gaan ze der te veel over nadenken en als ge te veel informatie geeft, dan krijgt ge nogal rap kritiek en zo rap negatief denken. En het gewoon op hun laten afkomen, ze weten niet wat het is – ik ook niet dus – ik denk dat dat de beste tactiek is.'*

d. COMMUNICATIE IN SCHOLEN IS VEELBESPROKEN

We geven nog mee dat communicatie duidelijk de meest besproken drager is van 'zelfevaluatie als beleidsdaad' (zie ook Tabel 84). De respondenten geven aan dat een gepaste communicatiestrategie cruciaal is om tot kwaliteitsvolle zelfevaluaties te komen. Dat komt omdat de term 'communicatie' gebruikt wordt om meerdere problemen te omschrijven. Zonder passende communicatie hebben ook de andere dragers van 'zelfevaluatie als beleidsdaad' volgens de respondenten immers weinig kans op slagen.

> *TL: 'Goh, der heerst hier zo op school wel een sfeer van, pfff, ze beslissen toch alles. Allez, dat er van bovenaf heel veel beslist wordt, zonder dat er gecommuniceerd wordt met de leerkrachten. Maar ik denk dat dat niet zo is, maar dat mensen dat gevoel hebben, omdat er alleen maar een verslag hangt en niemand dat eigenlijk leest.'*

5.1.4 Ondersteunende relaties en samenwerking: omtrent collegialiteit, vaardigheden en de beschikbaarheid van middelen

Vormgeven aan een zelfevaluatie vergt een voldoende grote draagkracht van scholen. Om die draagkracht te ontwikkelen streven de bestudeerde scholen naar interne samenwerking en ondersteuning. Daarenboven doen ze doelbewust een beroep op externe ondersteuning. Hieronder wordt ingegaan op de bevindingen omtrent beide vormen van ondersteuning.

a. INTERNE ONDERSTEUNING

De volgende drie thematieken blijken in de context van interne ondersteuning het meest van belang: vertrouwen en collegialiteit, kennis en vaardigheden, en de beschikbaarheid van tijd.

- *(Een gevoel van) vertrouwen en collegialiteit*

Uit de gevalstudies onthouden we in de eerste plaats dat er in de school sprake moet zijn van een voldoende 'veilige' context om de dialoog die met een zelfevaluatie gepaard gaat op gang te brengen/houden. Uit de volgende citaten blijkt dat leerkrachten het in de bestudeerde scholen niet vanzelfsprekend vinden om zich openlijk, vrijuit en kwetsbaar op te stellen.

> *TL: 'Ik denk dat dat heel moeilijk is als collega. Want uiteindelijk sta je daar ook wel als werknemer hé. En ze zullen dat nu misschien niet doen, maar ja de kans zit er wel in, ze hebben wel de macht om u te zoeken. Ik denk nu niet dat dat gebeurt, maar je bent toch altijd… je hebt toch altijd een zekere reserve naar uwe werkgever toe of naar uwe baas toe. En een werknemer, die gaat zeggen tegen z'n baas: 'Hélà manneke, ge doet dat hier niet goed' Allez ja, dat is niet zo simpel hé.'*

> *TL: 'Goh, dat is wel heel moeilijk denk ik (De vraag ging over knelpunten duidelijk maken aan de schoolleiding). Want niemand durft… ik zou dat niet durven zeggen. Wat ik nu allemaal vertel tegen u (de interviewer), dat zou ik tegen de directeur absoluut niet durven vertellen. Ik weet niet of dat er wel iemand is die dat durft, hé.'*

Hoewel uit deze citaten blijkt dat er wellicht niet altijd een échte grond toe is, is er een terughoudendheid bij de bevraagde leerkrachten om de dialoog met schoolleiding aan te gaan. Louter het feit dat het om schoolleiders gaat, blijkt voor een groep teamleden voldoende om een rem te zetten op de mate waarin zij zich bloot durven geven. Toch geven de bevraagde teamleden aan dat ze zich ondersteund voelen. De geïnterviewde leerkrachten blijken in hun school allemaal wel een aantal personen te hebben van wie ze zelf zeggen dat ze zich er goed door ondersteund voelen, zowel professioneel als persoonlijk. Dit zijn doorgaans personen uit de eigen vakwerkgroep. Leerkrachten kiezen hun vertrouwenspersoon naar eigen zeggen liefst zelf. Je kan dat niet verplichten, luidt het. Soms klikt het en soms niet. Dat blijkt bijvoorbeeld bij de begeleiding van nieuwe leerkrachten. Wanneer er in de context van een zelfevaluatie moet samengewerkt worden met teamleden die niet tot de zelfgekozen kring van vertrouwelingen behoren, blijken openheid en samenwerking echter veel minder vanzelfsprekend.

> *TL: 'Wij zitten wel met z'n allen samen maar eigenlijk vormen we een beetje 2 groepen. Als de verhelderingsgesprekken… ge voelt hoe die spanning. Dat blijft toch altijd een beetje een gespannen sfeer. We gaan ook echt nooit ons gedacht zeggen eigenlijk. Zoiets ga je der moeilijk uitkrijgen. Dat merkt ge ook gewoon op alle vergaderingen. Dan komen die spanningen eigenlijk of die staan in de weg… Sommige dingen worden verzwegen om geen ruzie te maken.'*

> *SL: 'Een grote hoop is bereid kritisch te zijn. Maar ik denk dat als ge dat in groep gaat moeten bespreken, dat dat moeilijker ligt. Als je die mensen apart, er loopt iets mis en je bespreekt dat met die persoon, dan gaat dat wel. Maar als je dat dan bijvoorbeeld in groep, binnen de vakgroep bijvoorbeeld, gaat bekijken dan zal men moeilijk over het eigen falen durven spreken.'*

Onder meer deze belemmeringen van collegialiteit naar de grote groep collega's dragen er toe bij dat men zelfevaluaties nogal eens als een verplichte activiteit ervaart. De bevraagde teamleden werken er aan mee omdat het van hen verwacht wordt, en minder om dat men ervan overtuigd is dat gezamenlijke reflectie tot waardevolle resultaten kan leiden.

- *Kennis en vaardigheden*

Doordat de bevraagde scholen een beroep hebben gedaan op externe ondersteuning werden hun eigen kennis en vaardigheden minder op de proef gesteld. Zij zijn dankzij de ondersteuning niet tegen problemen aangelopen waarmee ze anders wellicht wel te kampen zouden hebben gehad. Een belangrijk deel van het denk- en doewerk was immers al gebeurd of kon worden uitbesteed. De bestudeerde scholen hebben bijvoorbeeld zelf geen instrument ontwikkeld, volgden in grote mate de instructies zoals deze door de ondersteuners vooropgesteld werden en hoefden de gegevens niet zelf te analyseren of te rapporteren. Precies doordat de scholen voor zulke elementen een beroep deden op ondersteuners blijkt de 'interne ondersteuning en samenwerking' een minder sterk thema in de bestudeerde zelfevaluaties. Een sterk uitgebouwde externe ondersteuning lijkt dus minder eisen te stellen aan de interne ondersteuning, en vice versa.

Ondanks de beperkte evidentie voor het ontbreken van de nodige vaardigheden moeten we concluderen dat het zelfstandig uitvoeren van een zelfevaluatie kennis en vaardigheden vergt waarover schoolleiders naar eigen zeggen niet noodzakelijk beschikken. Het gaat bijvoorbeeld om het kiezen of ontwikkelen van zelfevaluatie-instrumenten, om het sensibiliseren van het team, om de gegevensanalyse of om het verzekeren van de doorstroming van de bevindingen naar de werkvloer. Sommige zaken blijken ook minder vanzelfsprekend dan men zelf spontaan zou denken.

> *SL: 'Gelijk die verhelderinggesprekken - daar waren wij dus nooit opgekomen. Ik zou ook geen instrument kunnen gaan samen stellen, om hetzelfde te verwachten. Nee, die expertise hebben we niet.'*

> *SL: 'En de interpretatie achteraf dus eigenlijk daar… dat zou ik niet kunnen. Ik denk dat niet. Ik denk dat wij misschien conclusies zouden trekken die totaal verkeerd zijn, kan zijn. Ik heb echt zoiets… ik denk niet dat er iemand aan zou durven beginnen hier.'*

Zulke ervaringen maken dat de bevraagde schoolleiders op zoek gaan naar externe ondersteuning. De verantwoordelijken voor de zelfevaluatie weten zich daardoor ondersteund door externen. Interne ondersteuning is in de bestudeerde scholen daardoor meer van toepassing op de andere teamleden. Voor hen is bijvoorbeeld het invullen van een zelfevaluatie-instrument niet vanzelfsprekend. Sommige leerkrachten kampen met een beperkte ICT-geletterdheid. De bestudeerde scholen anticipeerden op deze behoefte aan ondersteuning bij het invullen van de vragenlijst.

> *SL: 'Omdat we het volledig met de computer gingen doen waren er een aantal die zegden van: 'Awel, dat zie ik niet zitten.' Dus we hebben dan ook een groep samengesteld die de mensen kon begeleiden. Van hoe moet je het aanpakken?*

Hoe moet je het doen? Dat werd eerst allemaal uitgelegd en daarna kon men de begeleiders om hulp vragen.'

De scholen voorzien dus een helpende hand indien nodig. Het invullen van de vragenlijst blijkt evenwel een taak te zijn die niet tot noemenswaardige problemen leidt. Dat kan echter ook betekenen dat de interne ondersteuning functioneel bleek. Leerkrachten zelf geven bijvoorbeeld aan het te waarderen dat er voor hen een aanspreekpunt voorzien wordt voor als er problemen zijn. Zij voelden zich ondersteund tijdens de eigenlijke bevraging.

Ook met het gebruik van een stuurgroep wordt onder meer gestreefd naar het realiseren van interne ondersteuning en samenwerking. Problemen die men ervaart kunnen er uitgeklaard worden en verschillende leden van stuurgroepen in de bestudeerde scholen geven aan dat van zulk een structuur een motiverend effect uitgaat. Alleen al het bestaan van de stuurgroep blijkt voor dynamiek te zorgen. Het kan volgens de betrokkenen zorgen voor een gezonde druk die helpt om stappen vooruit te zetten.

- *De beschikbaarheid van tijd*

Schoolleiders geven herhaaldelijk aan dat ze naar hun aanvoelen over te weinig tijd beschikken om het zelfevaluatieproces optimaal te kunnen opvolgen. Ook teamleden stellen dit vast.

SL: ' Eerlijk gezegd, op dit moment, (zucht), weten wij, … dan praat ik dus vanuit de directie he, weten wij niet meer waar we aan moeten beginnen. En ik vrees dat wij dus ook niet gaan weten van, kunnen wij dit nog opvangen of niet? Op dit moment is het iets van… alle hens aan dek om de rest draaiende te houden, en dan in juli, gaan wij nog bekijken, als iedereen hier weg is, van, en hoe gaat het hier nu verder he, met die zelfevaluatie.'

TL: 'De directeur heeft te veel werk eigenlijk om zich daarop te concentreren. Hij zou er eigenlijk meer tijd voor maken. De tijd ontbreekt. Dus hij is nu volop bezig met dit onder andere. Hij is met te veel zaken bezig en heeft te weinig tijd hé.'

Het gebeurt dat schoolleiders met de beste intenties aan een zelfevaluatieproces beginnen maar gaandeweg vaststellen dat ze onvoldoende tijd hebben om te investeren in dat proces. Dat kan te wijten zijn aan onvoorziene omstandigheden – zoals een onverwachte doorlichting - die het verloop van een zelfevaluatie bemoeilijken. Maar vaak is het zo dat de zelfevaluatie een proces is waarvan men het gevoel heeft dat het niet prioritair is. Men geeft aan er aan te werken als de kans zich voordoet. De druk op beleidsmakers in scholen maakt dat de opvolging van het gebeuren naar de achtergrond verdwijnt. Het gebrek aan tijd leidt in de cases tot een gebrek aan prioriteit. Om dit op te vangen wordt nogal eens gesuggereerd dat het wenselijk zou zijn de verantwoordelijkheden niet door één persoon te laten dragen. Daarnaast wordt ook aangehaald dat het instappen in een bestaand ondersteuningstraject ook voor een gezonde druk kan zorgen.

Scholen hebben naar eigen zeggen te weinig tijd beschikbaar om het zelfevaluatieproces goed op te volgen. Bovendien kunnen ze niet flexibel omspringen met de beschikbare

tijd. Ook dat heeft implicaties. We stellen immers vast dat ook de timing van sleutelelementen van het zelfevaluatieproces bijdraagt tot een terughoudendheid bij de teamleden. In de ogen van zowel betrokken als niet betrokken personeelsleden blijkt die timing vaak betrokkenheid, openheid en diepgang in de weg te staan. We geven een aantal voorbeelden.

> Wat de timing en opzet van de gegevensverzameling betreft. *TL: 'Ja, dat was wel verplicht ja. Want ik, bij mij kwam dat eigenlijk heel slecht uit, ik ben donderdag vrij in de namiddag en ik ben daar speciaal voor moeten komen om kwart voor 4. En daar was ook geen andere weg mogelijk. Ik heb dan van collega's gehoord, dat die speciaal moesten van stage terugkomen. Dat was ook wel zo van elkaar af zitten. Echt precies een examen. Ik doe dat natuurlijk liever thuis dan dat ge…'*
>
> Wat de timing van verhelderingsgesprekken betreft. *SL: 'Die verhelderingsgesprekken die gaan dan door die vrijdag voor, voor de 1ste mei, dan hebben wij hier opendeurdag. Dus dat is ook weer zo'n moment, dat mensen andere dingen aan hun hoofd hebben. Dat is misschien ook weer niet zo'n goed gekozen moment. Ik ben er een beetje bang voor.(...) Dat maakt ook denk ik, dat mensen sneller wat negatief worden. Want ze vinden dat slecht georganiseerd van de school.'*
>
> Een leerkracht zegt over deze timing: *TL: 'Iedereen vertrekt daar van het feit, van we gaan dat het hier eens rap bonjouren hé. We gaan maken, dat dat zo rap mogelijk gedaan is en ik denk dat iedereen dan wel dingen verzwegen heeft, omdat zo vlug mogelijk laten door te gaan zo. Van ja, vlugger naar het volgende. Ik denk dat wel. Op een vrijdagavond.'*
>
> Wat de timing van het presenteren van de resultaten betreft. *TL: 'In de voormiddag moesten we dan wat administratie komen doen, wat opruimen en dan wat vakvergaderingen. En 's middags, dan is de gewoonte dat wij met de collega's iets gaan eten. Dus dat was eigenlijk heel fijn zo. Ontspannen vakantiesfeer. En dan moesten we de refter in, om 13h en dat heeft toch wel… Allez, ik weet niet, dat heeft toch wel anderhalf uur geduurd dacht ik. Ja, iedereen zat daar zo. Het was eigenlijk vakantie. En iedereen zijn kinderen hadden al vakantie. Dan heb je echt geen zin om naar zo'n saaie uiteenzetting te…'*

Teamleden worden blijkbaar in de verschillende scholen telkens op – in hun ogen - ongelukkige momenten ingeschakeld in het zelfevaluatieproces. Dat past niet in het plaatje van 'ondersteuning tijdens de zelfevaluatie'. Het is niet denkbeeldig dat men mede daardoor aan het gebeuren een negatieve gevoelswaarde koppelt. Hoewel de bevraagde schoolleiders dit inzien, leeft het gevoel dat dit vaak de enige momenten zijn waarop het praktisch mogelijk is om leerkrachten voor zelfevaluatie-activiteiten samen te brengen. Ook leerkrachten erkennen dit.

b. EXTERNE ONDERSTEUNING

Het uitvoeren van zelfevaluaties vergt specifieke kennis en vaardigheden. De door ons bestudeerde scholen genoten telkens ondersteuning. De meesten respondenten zijn van mening dat ondersteuning best ook buiten de school gezocht wordt. Het betrekken van

externen was dan ook een bewuste keuze in de bestudeerde scholen. Doorheen de beschrijving van de 'zelfevaluatie als onderzoeksdaad' en 'zelfevaluatie als beleidsdaad' wordt de concrete rol en impact van externe ondersteuners toegelicht. Daaruit blijkt dat de externe ondersteuning betrekking heeft op verschillende aspecten van de zelfevaluatie: het kiezen van instrumenten, het ontwikkelen en opzetten van een bevraging, aanvullende vormen van gegevensverzameling voorzien (zoals verhelderinggesprekken), het interpreteren van de zelfevaluatieresultaten en het vormgeven aan de communicatie omtrent de zelfevaluatieresultaten. In deze paragraaf beperken we ons verder tot de motieven om ondersteuning in te schakelen. De drie belangrijkste motieven kunnen samengevat worden als het binnenhalen van expertise, het betrekken van een onafhankelijke waarnemer en het bewaken van de opvolging van het zelfevaluatieproces.

Bij de start van de zelfevaluatie blijken de scholen in de gevalstudie het moeilijk te vinden om de eigen behoefte aan ondersteuning goed in te schatten. Zij geven aan niet goed te weten wat men moet verwachten, en welke kansen en valkuilen een zelfevaluatieproces met zich meebrengt. Die onduidelijkheid maakt dat scholen aangeven dat zelfevaluaties een deskundigheid vereisen waarover zij naar eigen zeggen waarschijnlijk nog niet beschikken. De voornaamste reden om een beroep te doen op externe ondersteuning schuilt precies in dat aanvoelen. Ondersteuning moet in sterke mate een antwoord kunnen geven op de vragen die leven bij de coördinatoren van de zelfevaluatie in de school. Scholen ervaren dat ondersteuners beter onderlegd zijn in het ontwikkelen van vragenlijsten en het opzetten van bevragingen. Een zekere vorm van specialisatie is volgens de respondenten nodig en kan in huis gehaald worden door een beroep te doen op ondersteuners. De respondenten hebben vertrouwen in en vertrouwen op de expertise van de ondersteuners. Ze ondergaan om die reden ook vaak de ondersteuning waarop ze een beroep doen.

> *SL: 'Wij ondergaan het een stuk, want voor ons is dat nieuw. De persoon die ons begeleid die zegt gewoon van dat moet er gebeuren en dat moet er gebeuren en wij volgen die stappen gewoon wat er moet gebeuren. Eigenlijk niet meer.'*

Als het zelfevaluatieproces eenmaal achter de rug is, blijkt één en ander echter veel concreter te zijn geworden. Men geeft dan ook aan meer eigen verantwoordelijk te durven nemen een volgende keer. Precies door het uitvoeren van een zelfevaluatie groeien de eigen competenties en is men van oordeel een volgende keer meer activiteiten zelf te kunnen aansturen. Dat maakt externe ondersteuning echter niet overbodig. Er blijven immers activiteiten waarvoor scholen graag een beroep doen op externen. Vaak worden het interpretatieproces en het komen tot actiepunten in deze context genoemd. De respondenten zijn vragende partij voor bijkomende ondersteuning hieromtrent. Op dit moment maken deze fasen immers geen noemenswaardig deel uit van de ondersteuning in het kader van de zelfevaluatie. Schoolleiders aanvaarden enerzijds wel dat ondersteuners zich op de achtergrond houden bij het selecteren van actiepunten. Anderzijds geven sommige respondenten aan precies ook in deze fase ondersteuning nodig te hebben. Die ondersteuning hoeft dan wel niet noodzakelijk van de begeleiders van de zelfevaluatie te komen. De specialisten moeten dan in andere zaken onderlegd zijn. De ondersteuners zelf gaan er vanuit dat de verdere begeleiding - na het presenteren van de zelfevaluatie-

resultaten - opgenomen zal worden door de reguliere begeleiders. In de feiten blijkt dit evenwel niet het geval te zijn. Er is in de cases een grote behoefte aan meer ondersteuning nadat de zelfevaluatieresultaten op tafel liggen.

> *SL: 'Wat de ondersteuners van zelfevaluaties nu doen, hetzelfde dat de doorlichting doet, is eigenlijk een meting uitvoering en dan zeggen, dat zijn de resultaten. Trekt uwe plan. En de enige ondersteuning die we kunnen gebruiken is die van de koepel, die mensen doen enorm hun best, maar zijn eigenlijk te weinig bemand om als er inderdaad grote pijnpunten zijn om die te veranderen.'*

Externe ondersteuning is echter meer dan het binnenhalen van externe expertise. Ook andere overwegingen spelen een rol in de keuze om een beroep te doen op externen. De aanwezigheid van begeleiding kan bijvoorbeeld voor een gezonde druk zorgen. De timing van het zelfevaluatieproces ligt grotendeels vast en de externe ondersteuners herinneren de school aan tussentijdse mijlpalen die men dient te respecteren.

> *SL: 'Alleen al het feit dat er een externe bij betrokken is, maakt dat ge dat iets of wat systematischer gaat doen. Ge moet wel data afspreken en ge moet u daar wel aanhouden, want er komt iemand van Antwerpen of Brussel. Dat is eigenlijk wel, ja.'*

Externe ondersteuning is volgens de respondenten bovendien ook een strategie om de objectiviteit te bewaken. Zo wordt bijvoorbeeld het feit dat verhelderingsgesprekken door een extern iemand geleid worden toegejuicht. Leerkrachten zien in externen een vorm van garantie dat hun inbreng ernstig genomen zal worden. Externen kunnen aspecten aankaarten, bijvoorbeeld naar de schoolleiding toe, die door leerkrachten niet in de verf gezet kunnen worden. Men vindt het bovendien waardevol dat externen de school eens vanuit een andere invalshoek bekijken.

> *TLZEV: 'Dat zijn dus mensen, die onafhankelijker kunnen oordelen. Dus die zijn helemaal niet bevooroordeeld, dus euhm, ik denk dat dat heel goed is. Dus, die kunnen meningen ook naast mekaar zetten, die kunnen meningen ook tegenover mekaar zetten. En mensen daar doen over praten.'*

> *TL: 'En een werknemer, die gaat zeggen tegen zijn baas: 'Hélà manneke, ge doet dat hier niet goed.' Allez ja, dat is niet zo simpel hé. Dus misschien is het wel goed dat daar iemand extern dat eventjes komt verduidelijken. Ja. Ik denk dat wel beter zou zijn, als die dat mondeling komen doen en dat die dan ook voor begeleiding zorgen.'*

De gevalstudies leren dat de behoefte aan ondersteuning in verschillende scholen sterk varieert. Er zijn objectieve en subjectieve verschillen in de mate waarin scholen over de nodige competenties beschikken. Voor bepaalde activiteiten heeft men soms de juiste mensen in huis, voor andere niet. Ook de vroegere ervaringen met zelfevaluaties spelen een rol. Terwijl de praktische opzet van een gegevensverzameling een eerste keer met veel vragen gepaard gaat, is dit bijvoorbeeld een activiteit die men in volgende zelfevaluaties sterker in eigen handen durft te nemen. De behoefte aan ondersteuning hangt verder ook af van de resultaten van de zelfevaluatie. Een school met goede zelfevaluatieresultaten kampt immers niet met dezelfde problemen als een school(leider) die zichzelf

geconfronteerd ziet met ernstige werkpunten. De invulling van ondersteuning gebeurt volgens de respondenten dus best zoveel mogelijk op maat van de school.

5.1.5 Gedeeld leiderschap: omtrent besluitvorming en betrokkenheid

In de bestudeerde zelfevaluaties speelde de schoolleiding duidelijk een centrale rol. Het tot stand komen van het initiatief voor de zelfevaluatie is een activiteit waarbij in geen enkele school leerkrachten betrokken werden. In het verdere verloop van de zelfevaluatie spelen leerkrachten wel een rol, evenwel niet in het minst omdat de externe begeleiders daarop aandrongen. De ondersteuners benadrukken dat gezamenlijke besluitvorming een breed draagvlak kan creëren voor de zelfevaluatie. De bevraagde schoolleiders bleken in eerste instantie niet volledig overtuigd van het belang van een dergelijke participatieve aanpak. Zij vinden het feit dat er een zelfevaluatie uitgevoerd wordt op zich al getuigen van participatieve beleidsvorming. De raadgevingen van ondersteuners om de besluitvorming te delen, worden soms bij aanvang zelfs wat weggelachen. Achteraf is men dan vaak wel overtuigd van het belang van inspraak:

> *SL: 'Als die stuurgroep er bijvoorbeeld niet geweest was, denk ik dat wij dus minder geïnformeerd waren. Omdat wij gingen er vanuit, eigenlijk is dat (de zelfevaluatie) vanzelfsprekend. Maar voor de leerkrachten is dat dus niet vanzelfsprekend. En daarom is het dan toch weer goed, dat ze ons gezegd hebben van: 'Betrek daar leerkrachten bij.'*

De besluitvorming over het verloop van de zelfevaluatie komt in de verschillende scholen op verschillende manieren tot stand. In één school werd die verantwoordelijkheid volledig gedragen door de adjunct-directeur. Het is misschien geen toeval dat de zelfevaluatie in deze school uiteindelijk niet heeft plaatsgevonden. In de twee andere scholen werd gebruik gemaakt van een stuurgroep. We staan hieronder eerst stil bij de werking van zulke stuurgroepen. Daarna richten we ons op de betrokkenheid van de andere teamleden.

a. DE STUURGROEP ALS STRATEGIE VOOR BETROKKENHEID

Omdat directe inspraak van alle teamleden in het zelfevaluatieproces niet werkbaar is, wilden twee scholen door middel van een stuurgroep toch een vorm van vertegenwoordiging realiseren. Een stuurgroep met leerkrachten geeft aan de keuzes die men maakt immers een zekere legitimiteit. Het komt voor dat een bestaande stuurgroep belast wordt met het coördineren van de zelfevaluatie. Een andere mogelijkheid is dat de stuurgroep speciaal voor de zelfevaluatie in het leven wordt geroepen. We lichten de vaststellingen bij beide werkwijzen even toe.

Wanneer een stuurgroep speciaal voor de zelfevaluatie in het leven geroepen wordt, moet de school op zoek naar leden die deze taak op zich willen nemen. We stelden vast dat het niet altijd vanzelfsprekend is om die te vinden. Wat zelfevaluaties betreft reageren leerkrachten blijkbaar niet enthousiast op inspraakmogelijkheden. Hun houding is in zekere mate inconsequent. Enerzijds zegt men het belangrijk te vinden om bij het beleid

betrokken te worden, anderzijds is men vaak niet bereid daarvoor de nodige inspanningen te leveren. In de bestudeerde scholen blijken teamleden niet gedreven om in een zelfevaluatiestuurgroep in te stappen. In de school die speciaal een stuurgroep in het leven wou roepen, leverde een eerste open oproep daartoe geen enkel resultaat op. De schoolleiding is dan zelf actief gaan ronselen. Sommige leden van de stuurgroep weten niet echt waarom de schoolleiding uiteindelijk bij hen terecht kwam. De leden van de stuurgroep zijn dan ook niet noodzakelijk sterk geïnteresseerd in de beleidsthema's die in de stuurgroep op tafel komen. Het opnemen van critici van de zelfevaluatie in de stuurgroep werd niet overwogen. Dit alles heeft uiteraard ook implicaties naar de werking van een stuurgroep toe. Het potentieel dat volgens de respondenten uitgaat van een stuurgroep stemt immers niet overeen met de feitelijke verwezenlijkingen. De voor de zelfevaluatie opgerichte stuurgroep werd voornamelijk betrokken bij praktische zaken zoals het opstellen van een timing en bij het organiseren van verhelderingsgesprekken. De eerste bespreking van de resultaten vond weliswaar plaats in de stuurgroep maar verder ging de rol van de stuurgroep niet. De stuurgroep is in totaal driemaal samen gekomen en hun impact is volgens de betrokkenen tijdens en na de zelfevaluatie beperkt gebleven. Dat is deels de verantwoordelijkheid van de schoolleiding, deels een effect van de ondersteuning en deels een gevolg van de beperkte inzet van de leden van de stuurgroep. Dat blijkt uit de volgende citaten.

> *SL: 'De stuurgroep heeft eigenlijk gewoon als taak gekregen… wij gaan een zelfevaluatie doen, jullie gaan mee helpen, dat geheel uitwerken. Maar eigenlijk was daar niet veel uitwerken aan, want uiteindelijk moesten we gewoon, een dag beslissen, wanneer gaan we die zelfevaluatie afnemen. En dan nu krijgen ze het resultaat voorgeschoteld, waar ze eigenlijk niet rechtstreeks iets mee mogen doen.'*

> *SL: 'De stuurgroep eigenlijk ... daar heb ik mij een beetje vragen bij. Of die stuurgroep echt nodig was. Uiteindelijk, je hebt een externe, het VSKO komt hier. Die mensen komen hier uitleg geven. Ze begeleiden ons en zeggen :'Awel, die fase moet je doorlopen, die fase moet je doorlopen, die fase moet je doorlopen' Maar dan heb je eigenlijk de stuurgroep, niet echt meer nodig. (...) Er zijn ook een aantal leden die zich de vraag stellen 'Wat moeten wij hier eigenlijk doen?' Want alles ligt vast.'*

> *TLZEV: 'Ja en in september dacht ik van 'oh, nee dat komt er weer bij' maar ik dacht, ik ga niks zeggen. Ik zal wel iets horen en… Ik heb er gewoon over gezwegen en het is niet van gekomen. Allez, er is nooit een uitnodiging gekomen of nooit een moment geweest van mensen we gaan terug samenkomen. Ik heb daar stilletjes over gezwegen.'*

In een andere school werd de coördinatie van de zelfevaluatie waargenomen door een bestaande werkgroep. Deze stuurgroep behartigde reeds alle aspecten die betrekking hebben op kwaliteitszorg. Het gaat daarbij om mensen die hier ooit spontaan zijn ingestapt of omwille van hun functie automatisch lid zijn. Toch is dat geen garantie voor een actieve werking. De leden van de stuurgroep hebben in deze school weliswaar als eersten het gehele zelfevaluatierapport ontvangen, maar vervolgens werd er gedurende 8 maanden geen overleg over georganiseerd.

b. BETROKKENHEID VAN DE ANDERE TEAMLEDEN

Hoewel een stuurgroep als doel heeft verschillende geledingen in het zelfevaluatieproces te betrekken, garandeert een stuurgroep niet dat individuele teamleden zich betrokken voelen. Ook teamleden pleiten nogal eens voor een grotere betrokkenheid. Betrokkenheid is volgens hen van bepalend belang om te begrijpen hoe teamleden de resultaten van een zelfevaluatie onthalen.

> *TL: 'Nu is onze inbreng eigenlijk gewoon die enquête invullen, eens komen luisteren naar het verslag, en dat was het eigenlijk he. Dus een beetje … (twijfelt) … het had iets meer kunnen zijn. Ik denk dat het effect daardoor ook minder gaat zijn. We zijn te weinig betrokken. Ik denk allez, volgend jaar zullen we een paar werkpunten hebben, maar ik ben een beetje bang … dat het niet echt au sérieux genomen worden. We zijn er niet echt mee bezig met heel die evaluatie.'*

Ook de bevraagde schoolleiders voelen dit aan. Zij geven aan dat een breed draagvlak creëren een voorwaarde is om met de zelfevaluatieresultaten aan het werk te kunnen gaan. Zij beseffen dat ze aan de leerkrachten ook een rol moeten toekennen in bijvoorbeeld het komen tot actiepunten op basis van de resultaten van de zelfevaluatie. Teamleden lijken in dit opzicht uiteen te vallen in twee groepen. Sommige teamleden willen in de besluitvorming betrokken worden. Deze groep ziet zichzelf immers als betrokken partij bij de thema's van de zelfevaluatie op schoolniveau. De tweede groep teamleden ziet er geen graten in dat het schoolbeleid door de schoolleiding (of door anderen) uitgestippeld wordt. Betrokkenheid en de nood aan inspraak worden dan ook zeer subjectief beleefd. Dat maakt het niet eenvoudig om een best passende strategie te kiezen. De gevalstudies bevestigen de wenselijkheid van inspraak wanneer teamleden zich betrokken voelen en wanneer ze het gevoel hebben inhoudelijk iets te kunnen bijdragen. Een moeilijkheid hierbij is echter dat sommige teamleden zoals reeds aangehaald tegenstrijdige geluiden laten horen. Voor sommigen zal het altijd 'te weinig' en 'te laat' zijn.

> *SL: 'Ik heb zelf bij die zelfevaluatie iets van 'zo is het wel gemakkelijk hé'. Ge vult een vragenlijst in, ge gaat in dat gesprek u wat zitten te ventileren. Euhm, en daar stopt het. Het moment dat ge dan, eigenlijk in actie moet gaan schieten, zeg je van allez ja nu wil ik niet meer. Dan laat ik dat over aan dat groepje daar en blijf ik verder gewoon aan de zijlijn staan en commentaar geven. Die oppositie. Die, euhm kritiek geven en euhm, maar niet op een constructieve manier. En op het moment dat ze de kans krijgen om echt te participeren en iets te doen, dan stoppen.'*

Het is dan ook niet eenvoudig om als schoolleiding een gevoel van gedeeld leiderschap te creëren. Toch moet de beperkte waardering van teamleden voor de inspraakinitiatieven deels ook aan de sturende aanpak van de schoolleiding toegeschreven worden.

> *SL: 'Ja, ik kan dat overleg zelf sturen hé. (lacht) Evalueren en rapporteren zal er zeker inzitten.'*

> *SL: 'Dat krijg ik al wel eens te horen, van dat wij alles beslissen. Men zegt dan van:'Awel, waarom moeten wij daar aan meewerken? Gij beslist toch.' Dan hebben we gezegd, er moet iets anders komen, maar wat er komt dat komt van jul-*

> *lie uit. Dus op die manier hebben ze soms wel eens de neiging van te zeggen van: "Jullie beslissen dat toch!". Dan is onze wil er wel maar toch blijft men denken dat wij het zullen bepalen.'*

Hieruit leren we dat gedeeld leiderschap maar werkt als het van bij de start en oprecht nagestreefd wordt. In het laatste citaat beslist de schoolleiding eerst dat er iets anders moet komen om daarna participatie na te streven. Gezien de bestudeerde zelfevaluaties allen door de schoolleiding geïnitieerd werden is een gelijkaardige reactie bij de teamleden tijdens het proces niet ondenkbaar. Hoewel de schoolleiding tijdens het zelfevaluatieproces dan oprecht streeft naar inspraak komt er toch kritiek omdat de teamleden zelf niet om de zelfevaluatie gevraagd hebben. Gedeeld leiderschap kan dus geen eenmalige gebeurtenis zijn.

We willen ten slotte nog illustreren dat er verscheidene redenen zijn die teamleden ervan kunnen weerhouden om op een participatieve manier mee te stappen in een zelfevaluatieproces. Die redenen gaan duidelijk verder dan het zelfevaluatieproces als dusdanig. Uit het volgende blijkt bijvoorbeeld dat zowel vroegere ervaringen als het statuut van teamleden een rol kunnen spelen.

> *TL: 'De directie staat wel open voor u, voor een gesprek. Dat wel, ge kunt er als ge met iets zit, maar ik zeg het… wat dat bij ons specifiek is (bij het ondersteunend personeel), wij doen dat minder en minder, omdat wij toch meestal bot vangen, we krijgen toch niet ons gelijk. Daar komt het op neer. En dus we maken daar minder gebruik van eigenlijk. Hoewel de directie er wel voor open staat eigenlijk. En daardoor, zo die zaken, die diep geworteld zitten dus… die worden niet meer besproken, omdat het toch… er is toch geen goede oplossing voor of er is toch geen aandacht voor. En dan zwijg je daar ook maar over tijdens de zelfevaluatie.'*

> *TL: 'Nee, maar ik denk dat ook een beetje is vanuit mijn positie ook. Gewoon door het feit dat ik maar, tijdelijk ben maar voor de moment. Dat ik mij voor de rest een beetje afzijdig houdt, langst de lijn een beetje kijken. Volgend jaar kan ik gedaan hebben hé, bij wijze van spreken. Dus hé… ge gaat zeker niks doen, om de directie te storen, zal ik maar zeggen.'*

5.1.6 Geïntegreerd beleid: omtrent verankering in bestaande systemen en denkkaders

De bestudeerde zelfevaluaties zijn niet strikt geïntegreerd of strikt gesegmenteerd. In alle cases zijn zowel indicaties te vinden voor de ene als voor de andere richting. Een algemene vaststelling is evenwel dat de eigenlijke opzet van de zelfevaluaties nogal eens als een eenmalige activiteit ervaren wordt. Vooral teamleden die niet bij de coördinatie van de zelfevaluatie betrokken worden, beschouwen het proces als op zichzelf staand. Zij zien vaak de link met andere initiatieven in de school niet. De bevraagde schoolleiders daarentegen zien wel in op welke manier de zelfevaluatie in het ruimere schoolbeleid inpast. Zij zien de zelfevaluatie als onderdeel van het integrale kwaliteitszorgsysteem dat de school hanteert. Ze geven evenwel aan dat de onderlinge linken nog kunnen aangescherpt worden. Zelfevaluatie zou volgens verschillende respondenten nog meer een

continu en geïntegreerd proces moeten worden. De gevalstudies geven hiertoe een aantal aanknopingspunten. Deze aanknopingspunten sluiten zowel aan bij zaken die volgens de respondenten goed lopen als bij zaken die mank lopen. We zetten ze hieronder op een rij.

We stellen in eerste instantie vast dat de scholen thema's evalueren die het naar eigen zeggen mogelijk maken om de samenhang tussen verschillende aspecten van het schoolbeleid te documenteren. Dat is uiteraard voornamelijk een verdienste van het instrument waarmee scholen werkten. Volgens de respondenten richten deze instrumenten zich op thema's die hen in staat stellen om een ruim en geïntegreerd beeld van de school te verkrijgen. De gecombineerde aandacht voor het pedagogische, onderwijskundige en organisatorische wordt dus gewaardeerd.

De respondenten geven aan dat het niet onverstandig is om de zelfevaluatie door een reeds bestaande werkgroep te laten coördineren. Dat geldt voor de praktische organisatie maar ook voor het komen tot actiepunten. Het geïntegreerde karakter van de zelfevaluaties komt voornamelijk tot uiting in de fase waarin met actiepunten aan de slag moet worden gegaan. Zo stellen we bijvoorbeeld vast dat men er in de verschillende scholen naar streeft om de actiepunten door bestaande werkgroepen te laten uitwerken. Aansluiten bij bestaande structuren heeft duidelijk voordelen volgens de respondenten. Door het gebruik van bestaande werkgroepen kan men bijvoorbeeld het overzicht bewaren en bewaken dat er een heldere lijn zit in de initiatieven die men neemt. Geïntegreerd beleid betekent immers ook dat men keuzes maakt die passen in een ruimer kader.

> *TL: 'Punten eruit halen en die door bestaande comités laten uit werken... Ik denk dat dat beter is, dan echt weer zo'n comité daar voor op te richten om al die punten, te realiseren.'*
>
> *SL: 'Ik vind het al wel belangrijk, euhm dat we geen, dat dat niet iets is dat er bovenop komt. Als het dingen zouden blijken te zijn, die door een bestaande groep binnen de school, kunnen opgenomen worden, dat die groep dat dan toegewezen krijgt.'*
>
> *TL: 'Ze moeten er ook vooral voor zorgen dat er structuur blijft hé. Dat ze niet eens daar een maatregeltje en daar een maatregeltje… dat dat een groter plan is. En ook niet te veel in één keer willen. Dat laten groeien en bezinken.'*

De link met de pedagogisch-onderwijskundige visie van de school is er in de bestudeerde scholen, maar blijft impliciet. Respondenten geven bijvoorbeeld aan of bevindingen al dan niet aansluiten bij wat zijzelf of hun school 'belangrijk' vinden. In geen enkele school werden de resultaten van de zelfevaluatie echter doelbewust naast de eigen visie of naast bestaande strategische of operationele doelen gelegd. De zelfevaluaties worden ook niet ondersteund door een theoretisch raamwerk. Ondanks het feit dat ondersteuners zulk een kader willen aanbieden zien sommige schoolleiders daar het nut niet van in.

> *SL: 'Het zou veel beter zijn, dat ge dan daar ook een keer uitlegt hoe een zelfevaluatie, hoe dat proces verloopt, hoe in de realiteit, wat daar allemaal gebeurt, dus dat je meer praktijkgericht werkt. Maar dat ligt misschien ook weer aan ons he, wij moeten, die theoretische blabla hoeven wij hier niet.'*

Tenslotte wijzen we nog op het belang van het afstemmen of integreren van verschillende zelfevaluatiesystemen in een school. Doordat zelfevaluaties op schoolniveau opgezet worden voor de hele school, bestaat – zo blijkt uit de cases - bijvoorbeeld het gevaar dat ze interfereren met initiatieven die door afzonderlijke afdelingen genomen worden.

> *TLZEV: 'Maar wij hadden eigenlijk ook zelf al een evaluatiedossiertje opgemaakt van documenten en vergaderingen en van dingen van de school. Maar dan gewoon met open vragen, dus een ander systeem. Dus het ongelukkige was eigenlijk dat één, de leerkrachten op dat moment twee bundels kregen en dat er bij de bundel van IZES eigenlijk vrij weinig informatie was gegeven.'*

> *TLZEV: 'Eens men echt beslist of de directie beslist van ja, we gaan dat doen, dan zou ik er voor pleiten om naar één geheel te gaan. En geen versnipperde, want dan hebben de mensen het gevoel van evalueren en nog eens evalueren en dat werkt niet. Dus dan zal ik daar zeker voor pleiten om te vragen van: 'kijk, bouw dat daar mee in dat we dat in één keer kunnen doen.'*

5.1.7 Responsief vermogen: omtrent de input van externe impulsen

Van responsief vermogen in strikte zin blijkt in de bestudeerde zelfevaluaties niet echt sprake. De zelfevaluaties worden volledig vorm gegeven binnen de eigen schoolcontext. Inspelen op vragen en verwachtingen uit de omgeving van de school is in deze zelfevaluaties niet echt aan de orde. Het doel van de scholen was de eigen werking te documenteren door de eigen personeelsleden middels een schriftelijke vragenlijst te bevragen. Er werden daarbij geen initiatieven genomen om ook mogelijke buitenstaanders bij de zelfevaluatie te betrekken. Schoolleiders én leerkrachten zijn het idee om andere doelgroepen bij een zelfevaluatie te betrekken echter niet ongenegen. Het gaat dan zowel om het betrekken van leerlingen, van ouders, van andere scholen als van zogenaamde kritische vrienden. Daarnaast wordt ook gepleit voor uitwisselingen met andere TSO/BSO-scholen die zelfevaluatieprocessen doorlopen.

> *TL: 'Ja, misschien zit daar meer inspiratie, als je het zo eens vanuit het jongerenstandpunt bekijkt. Dan komen er misschien betere oplossingen. Nu ja, als dat echt gaat over interne politiek … daar kunnen ze niet zo veel over weten. Maar als het echt gaat over omgaan met leerlingen dan denk ik dat daar heel zinvolle dingen kunnen uitkomen.'*

> *SL: 'Misschien zou het ook eens fijn zijn om van een buitenstaander te horen te krijgen van: 'Goh, zo zie ik uw school en dat zie ik mislopen.' Misschien zou dat wel eens plezant zijn. Misschien zou het een andere kijk geven hé. Dat denk ik wel, dat dat misschien eens nuttig kan zijn.'*

> *SL: 'Het is de moeite dat je de ervaring van anderen ook hebt. Van hoe zijn zij daar mee omgegaan? Hoe zijn zij bijvoorbeeld met weerstanden om gegaan? Daar leer je dan toch ook al een stukje uit. En dat is iets dat vanuit uw eigen beperktheid of dat ge zegt van: 'daar weet ik weinig van.' Ik heb zelf niet de wijsheid in pacht. Vandaar ook dat de nascholing, de gesprekken met collega's uit andere sectoren, wel belangrijk zijn.'*

Concrete stappen om zicht te krijgen op de verwachtingen en het oordeel van externe stakeholders werden tijdens de bestudeerde zelfevaluaties niet gezet. Men geeft aan dat dit toekomstmuziek is. Scholen rekenen op dit moment op hun begeleiders om te vernemen wat er in andere scholen gebeurt. Zij zijn hun link met de buitenwereld die het minder nodig maakt om te netwerken. SL: '*Daarvoor heb ik, vind ik, de begeleiders. Die hebben dus al die scholen gedaan. Let op voor die valkuilen, impulsen geven en zo...'*.

Een spoor dat wel wijst op enig responsief vermogen is de link van zelfevaluatie met de doorlichtingen van de onderwijsinspectie. Schoolleiders geven aan dat er geen directe link is, maar ontkennen niet dat zij door het uitvoeren van zelfevaluaties inspelen op de eisen die de bestuurlijke omgeving aan hen stelt. SL: '*Dat wordt soms al wel eens gezegd zo van…allez dan kunnen we dat toch tegen de doorlichting zeggen, dat we met zelfevaluatie bezig zijn.'* Bovendien zou ook het uitvoeren van een zelfevaluatie als dusdanig reeds beschouwd kunnen worden als een indicatie van responsiviteit. *TLZEV: 'De school is een stukje, is een deel van de maatschappij. En maatschappij staat niet stil, dus een school mag ook niet stil staan.'* Beide voorbeelden impliceren echter niet dat de school ook tijdens de zelfevaluatie responsief tewerk gaat.

Het feit dat de bestudeerde scholen hun zelfevaluaties niet sterk responsief ingevuld hebben, betekent niet dat het geen responsieve scholen zijn. De initiatieven die scholen linken aan hun sociaal-culturele en economische omgeving zijn talrijk. Die houding verdwijnt evenwel in de context van kwaliteitszorg. Zelfevaluatie houdt immers ook in dat men zich kwetsbaar opstelt en dit doen scholen blijkbaar het liefst zonder de betrokkenheid van bekende externen (ouders, naburige scholen, stageplaatsen, en dergelijke). De betrokkenheid van een externe in een vertrouwelijke en onbedreigende relatie (zoals in een ondersteuningscontext) lijkt dan weer minder voor problemen te zorgen. Toch is ook dat niet altijd vanzelfsprekend.

> *TL: 'Gewoon, kunnen we dat niet onder ons oplossen, zoals we dat altijd gedaan hebben. Dat lukt toch ook. Dat geeft toch altijd zo'n beetje: Waarom hebben we iemand van buitenaf nodig? En dat is met dit denk ik ook wel zo een beetje zo. Al wat van buitenaf komt is zo'n beetje … wrevel is nu het juiste woord niet, maar toch een afstand zo van, is dat nu echt nodig dat er iemand van buitenaf komt om ons te zeggen hoe het moet. '*

5.1.8 Innovatief vermogen: omtrent de zelfevaluatie als innovatie

De gevalstudies hebben slechts weinig empirische evidentie opgeleverd over innoverende benaderingen en technieken bij het uitvoeren van zelfevaluaties. Bijvoorbeeld wat betreft het communiceren over de zelfevaluatie, het sensibiliseren van het personeel, het presenteren van resultaten of het komen tot actiepunten werd er in de bestudeerde scholen telkens over eerder veilige en daardoor ook weinig creatieve paden gewandeld. Het lijkt erop dat het uitvoeren van zelfevaluaties op zich reeds een vorm van vernieuwing is in de bestudeerde scholen. Het is voor de schoolleiders en de andere teamleden een nieuwe methodiek om aan kwaliteitszorg te doen. Het proces roept nog vele vragen op,

kent nog vele onbekende elementen en het is voor de betrokkenen – ook in de scholen die al eerder een zelfevaluatie uitvoerden - niet geheel duidelijk waartoe het zal leiden. Wellicht mede daardoor beperkten de bestudeerde zelfevaluaties zich bijvoorbeeld hoofdzakelijk tot het gebruik van schriftelijke vragenlijsten bij de teamleden. Naast de verhelderingsgesprekken – een suggestie van de ondersteuners - hebben er geen bijkomende vormen van gegevensverzameling plaatsgevonden. Nochtans werd en wordt er in andere contexten in de scholen wel met andere benaderingen gewerkt. Daarbij denken we aan het gebruik van interviews, zelf ontwikkelde vragenlijsten, evaluatiegesprekken en dergelijke. Mogelijks worden de doelen die scholen wilden bereiken met het uitvoeren van de zelfevaluatie in de toekomst ook op andere manieren nagestreefd. De bestudeerde zelfevaluaties volgden echter redelijk strikt de voorschriften van de ondersteuners.

Hoewel er in de interviews weinig sprake bleek van een innovatieve aanpak van de zelfevaluatie, werd indirect toch veel vernomen over het innoverend vermogen van scholen. Uit onderzoek (Geijsel, 2001) weten we dat vernieuwingen meer kans maken wanneer ze de steun genieten van diegenen die de vernieuwing in de praktijk moeten realiseren en wanneer ze aansluiten bij de bestaande basisvisie van de school. Onze verwachting is dan ook dat scholen met een sterk innoverend vermogen bij het uitvoeren van een innovatie (zoals een zelfevaluatie) hoog zullen scoren op dragers als 'gezamenlijke doelgerichtheid', 'ondersteuning en samenwerking' en 'reflectief vermogen'. Wanneer een zelfevaluatie in wezen zelf een innovatie is, blijkt hoe een school hier vorm aan geeft dus in sterke mate uit de andere dragers van 'zelfevaluatie als beleidsdaad'. Dat maakt het 'innoverend vermogen' van scholen in deze context eerder tot een invalshoek dan tot een drager die bijkomende informatie oplevert.

5.2 Zelfevaluatie als onderzoeksdaad

De resultaten met betrekking tot 'zelfevaluatie als onderzoeksdaad' worden gestructureerd op basis van drie opeenvolgende stappen die scholen doorlopen bij het vorm geven aan een zelfevaluatie. Deze zijn het vastleggen van de inhoud van de zelfevaluatie, het uitwerken van de gegevensverzameling en het interpreteren van de resultaten.

5.2.1 De keuze van een instrument en het vastleggen van de inhoud

Op zich komen er ontzettend veel aspecten van het functioneren van scholen in aanmerking om aan een zelfevaluatie te onderwerpen. Scholen behoren dan ook keuzes te maken. In de bestudeerde scholen ging men in grote mate voorbij aan deze keuzes en had men weinig inbreng in de inhoud van de zelfevaluatie. Het vastleggen van de inhoud van de zelfevaluatie bleek immers sterk samen te hangen met de instrumentkeuze.

> *SL: 'Op zich niks hé. Want uiteindelijk, wij gaan dat instrument op zich gewoon gebruiken zoals het ons aangeboden werd. Dus wij hebben daar niet kunnen zeggen: 'Laat ons dat bekijken of dat bekijken'. Gewoon dat officieel instrument gebruiken eigenlijk.'*

> *Interviewer: Waarover verzamelt het zelfevaluatie-instrument gegevens? SL: 'Dat … durf ik nu niet zeggen. Want eigenlijk hebben we dat dossier, allez, hebben we alles nog niet doorgenomen. Dus ik weet niet, waar we allemaal informatie van hebben. Dus dat weet ik eigenlijk nog allemaal niet.'*

De inhoud van de zelfevaluaties werd in de bevraagde scholen dus grotendeels bepaald door de keuze van het te gebruiken instrument. Zij kozen een instrument en legden zich vervolgens neer bij de inhoudelijke keuzes die de ontwikkelaars van het instrument maakten. De keuze van inhouden was in de cases dus eerder aanbod- dan vraaggestuurd. Schoolleiders geven zelf aan dat men eigenlijk meer doordacht tewerk zouden moeten gaan bij het kiezen van een instrument. Momenteel leggen scholen op dat punt echter een grote verantwoordelijkheid bij hun ondersteuningsdienst.

> *SL: 'Allez, misschien moeten we de tijd daarvoor pakken… Dat hebben wij dus niet gedaan. Waarschijnlijk wel, dat lijkt me logisch… Want of dit een goed instrument is, dat kunnen wij niet zeggen. Wij gaan gewoon verder op hetgeen, dat vanuit Brussel komt. Zij zeggen: 'Awel, dit instrument is ter beschikking.' Dan gaan wij er van uit, dat zij het beste instrument ter beschikking stellen aan ons. Want als wij daar ook nog eens al onze energie moeten gaan insteken, dan vrees ik dat hier te horen krijgen van: 'Laat het dan, ze.' We gaan er vanuit dat als Brussel dit aanbiedt, dan moet dat het beste zijn. Want anders moeten ze zoiets niet aanbieden.'*

Men geeft wel aan in de toekomst meer bewust en meer participatief tot een inhoudelijke invulling van de zelfevaluatie te willen komen.

> *SL: 'Ik zou volgende keer aan de werkgroepen gaan vragen welke onderdelen gaan we terug bevragen, welke zijn absoluut niet relevant voor ons. Niet relevant voor dit moment dan. En op die manier…dat geeft natuurlijk ook voor hen zelf meer motivatie om daar aan mee te werken.'*

a. CRITERIA BIJ DE INSTRUMENTKEUZE

Er is in geen enkele bevraagde school evidentie voor systematische reflecties op de inhoud van verschillende instrumenten. Dat neemt niet weg dat scholen wel degelijk bepaalde motieven hebben bij het kiezen van een instrument. Het gaat dan eerder om impliciete criteria die ze in hun instrument bevestigd zien.

> *SL: 'Want als we op dat vlak eerlijk moeten zij … DISO is ons aangeboden en we hebben niet op zoek gegaan naar andere instrumenten. Omdat wij te horen kregen, dat andere scholen ook gebruikte, hebben wij gezegd van: 'Eigenlijk is dat de moeite waard om eens een keer te kijken.'*

> *SL: 'Wij hadden 4 jaar geleden een zelfevaluatie gedaan met DISO. En bij ons is dat eigenlijk wel meegevallen, want wij zijn daar heel content over, over heel het proces en de begeleiding. Euhm, dus zijn daar nu terug contacten gelegd met ja, het VSKO, … omdat die contacten er dan toch waren, hebben wij ingepikt om die 2de ronde te doen. Dus het lag een beetje voor de hand om, nu dan terug daarmee te werken.'*

> *SL: 'En wat de doorslag gegeven heeft was, dat het gratis was euh, dat er een deel opleiding aan vast zat, voor een aantal mensen.'*

Uit deze citaten blijkt onder meer het belang van vroegere ervaringen met het instrument, van de ervaringen van andere scholen, van de kostprijs en de verwachting dat de keuze voor een instrument gepaard gaat met een vorm van ondersteuning. Er zijn verder ook criteria die men wel belangrijk vindt, maar die bij de keuze van het instrument geen rol gespeeld hebben. Zo geeft men aan dat het verstandig kan zijn om telkens met hetzelfde instrument te werken en dat men de eigen resultaten ook wil vergelijken met deze van andere scholen. Een criterium dat pleit tegen bestaande instrumenten is de afstemming op de eigenheid van de school. Een nadeel van bestaande instrumenten is immers dat ze algemeen zijn en daardoor niet goed lijken aan te sluiten bij de specifieke context van de school.

> *SL: 'Als je een bevraging laat doen door een externe moet je zien dat die externe zijn instrument kan verfijnen naar de eigen schoolsituatie. De vragen naar bedrijfsgerichtheid waren bijvoorbeeld zo geformuleerd dat de verpleegkundigen dat naast zich neer hebben gelegd en gezegd ja, maar bedrijven dat hebben wij niet en als reactie hebben die daar nee op geantwoord. Terwijl geen enkel school zo goed bezig is die bedrijfscontacten te leggen.'*

b. EEN EIGEN (INBRENG IN HET) INSTRUMENT?

Het belang van maatwerk zet de bevraagde scholen aan het denken om zelf instrumenten te ontwikkelen of om waar mogelijk bestaande instrumenten aan te passen. Zelf een volledig instrument ontwikkelen, vinden de bevraagde schoolleiders echter een te omslachtig werk. Er steekt te veel tijd en energie in het maken van vragen. Bovendien heeft het werken met een bestaand instrument het voordeel dat het door het team als meer objectief beschouwd wordt.

> *SL: 'Dat is een enorm werk, daar heb ik zoiets van 'Is het dan nog wel nuttig?' Want je steekt daar zoveel tijd en energie in om die vragen te maken? Dus het afnemen is niks, maar het is eigenlijk de zaak opstarten, dat veel moeilijker gaat zijn denk ik.'*

> *SL: 'We gaan geen eigen instrument ontwikkelen, hé. Die knowhow hebben we niet en die tijd hebben we niet om daar in te steken.'*

> *SL: 'Stel u voor, dat ik nu gewoon zelf als directielid iets in mekaar zou steken. Een vragenlijst, ja, nee, matig en onvoldoende en hup. Dan denk ik dat ik wel heel veel tegenkanting zou krijgen. Euhm, want ja, iets dat van de directie komt dat is sowieso, al een beetje verdachter, denk ik, dan iets dat van buitenaf komt. Zeker als ge kunt zeggen dat het iets is, dat op meerdere plaatsen gebruikt wordt.'*

In één school met zelfevaluatie-ervaring werd de bestaande vragenlijst wel aangepast aan de eigen behoeften. Zij kregen deze kans van de ondersteuners omdat ze een aantal jaren voordien reeds met het standaardinstrument gewerkt hadden. Er werd na overleg in de stuurgroep een beperkt aantal vragen aan de standaardvragenlijst toegevoegd. Dat heeft volgens de schoolleider van die school het effect gehad dat de vragenlijst meer als een eigen instrument beschouwd werd.

> *SL: 'Als je school de kans krijgt van daar een beetje een eigen accent in te leggen in zo'n vragenlijst dat geeft wel een gevoel van betrokkenheid op dat ding hé. En dat is wel goed.'*

Zelf vragen ontwikkelen blijkt echter geen vanzelfsprekende onderneming. De betrokkenen vinden het moeilijk om zelf vragen te formuleren die in een vragenlijst opgenomen zouden kunnen worden. Het ontwikkelingsproces vergt vaardigheden waarvan men vindt dat men er onvoldoende over beschikt. Bovendien vreest men dat de bevraging door met eigen vragen te werken minder wetenschappelijk gaat overkomen.

5.2.2 Gegevensverzameling

De bestudeerde zelfevaluaties omvatten twee vormen van gegevensverzameling. In de eerste plaats werd er gewerkt met een schriftelijke vragenlijst. In de scholen die met DISO werkten werden de resultaten van deze schriftelijke vragenronde besproken tijdens een verhelderingsgesprek. Voor beide vormen van gegevensverzameling staan we stil bij de argumenten die het gebruik volgens de respondenten rechtvaardigen en bij aandachtspunten om een betrouwbare werkwijze te garanderen.

a. HET GEBRUIK VAN SCHRIFTELIJKE VRAGENLIJSTEN

De bevraagde scholen houden rekening met de efficiëntie van de vormen van gegevensverzameling waar zij voor kiezen. Met een vragenlijst kan je op een snelle manier veel mensen bevragen. De respondenten zijn het er evenwel mee eens dat het ook waardevol kan zijn om door middel van gesprekken gegevens te verzamelen. Men vindt het echter te omslachtig en tijdrovend om een volledig team door middel van interviews te bevragen. Een geïnstrumentaliseerde zelfevaluatie geniet onder meer om deze reden vaak de voorkeur. Schoolleiders hebben echter ook om andere redenen een voorkeur voor schriftelijke instrumenten. Ze hebben duidelijk vertrouwen in de kwaliteit van de instrumenten die hen ter beschikking worden gesteld. Men verwacht op die manier een betrouwbaar en representatief beeld van de school te bekomen.

> *SL: 'Het voordeel van een instrument is natuurlijk, dat ge der zeker van bent, dat ge ieder z'n mening hebt. Ik vind dat wetenschappelijk onderbouwde instrumenten in ons ogen een garantie geven. Dat ge een beetje een betrouwbare meting hebt hé. Dat ge kunt zeggen, dat dat ook wel de mening van het personeel vertegenwoordigt.'*

De ervaringen in de gevalstudies leren dat er bij het gebruik van vragenlijsten evenwel een aantal mogelijke problemen op de loer liggen die de betrouwbaarheid en representativiteit van de gegevensverzameling in het gedrang kunnen brengen. Zo zouden de vragen in de zelfevaluatie-instrumenten niet altijd duidelijk zijn voor alle teamleden. Dat is bijvoorbeeld het geval wanneer niet iedereen op de hoogte is van het feit dat je meerdere antwoorden kan geven bij bepaalde vragen of wanneer onduidelijk is vanuit welk perspectief men de vragenlijst moet invullen. Een betrouwbare gegevensverzameling vergt dan ook dat men aan het correct invullen van een vragenlijst voldoende aandacht besteedt. Ook de anonimiteit dient gewaarborgd te worden.

> *TL: 'Dat was dan, dat standpunt, ge moest dus de enquête invullen, vanuit het standpunt van… ja, vanuit een hele groep of een hele school eigenlijk. Van wat of hoe staat de school tegenover iets, terwijl dat ik dat heel moeilijk vind, want uw mening of u manier van handelen, is niet altijd gelijk aan die van de school. Dat vond ik, wel een beetje ambetant en een beetje tegenstrijdig ook.'*
>
> *SL: 'Het gaf een veilig gevoel voor de mensen, van awel het is niet op naam. Dat vonden ze dan wel weer goed hé. De grote hoop zou daar niets uit maken denk ik. Maar er zouden er wel een aantal bij zijn die zeggen, van nu vul ik het niet in. Dus dat kan ik mij wel voorstellen.'*

Soms kiest men ervoor om de vragenlijst door alle leerkrachten op een gezamenlijk en verplicht moment te laten invullen. Waar dat gebeurde werd daar door de ondersteuners op aangedrongen. Een groot aantal respondenten is echter van mening dat men het invullen minder als een opdracht zou ervaren wanneer men zelf zou kunnen kiezen wanneer men het doet. Dat zou een impact kunnen hebben op de resultaten. Net als het tijdstip van de afname overigens.

b. HET GEBRUIK VAN VERHELDERINGSGESPREKKEN

Verhelderingsgesprekken maken standaard deel uit van het DISO-zelfevaluatietraject. Ze zijn een initiatief van de ondersteuners en worden door hen geleid. Verhelderingsgesprekken vinden plaats in groepen die schoolleiders of teamleden van eenzelfde afdeling of graad samenbrengen. Men krijgt er de kans om een kwalitatieve toelichting te geven bij de resultaten van de schriftelijke bevraging. Dat heeft volgens de respondenten een aantal voordelen. Door het gesprek met mensen aan te gaan, komt men ruimere en andersoortige informatie op het spoor. De respondenten kijken doorgaans zeer tevreden terug op de verhelderingsgesprekken. De gesprekken zorgen voor nuances en diepgang. Men vindt de variatie in meningen en argumenten die in een verhelderingsgesprek naar voren komt zeer leerrijk.

> *TL: 'Mondeling kan je toch wel altijd beter de dingen uitleggen of echt wel heel specifiek vertellen. Terwijl dat op computer toch meer eenzijdig is. Je ziet daar een zin staan. Je kunt die eigenlijk nog een beetje interpreteren zoals je zelf wil. Dat kan je echt wel helemaal verduidelijken in die verhelderinggesprekken. In die zin vond ik dat dan wel zinvol.'*
>
> *Lk: 'In die verduidelijkingsgesprekken heb je bijvoorbeeld een bepaalde collega die zegt, ge krijgt hier toch zo weinig een dank u wel. Of iemand die zegt, wij schrijven wel verslagen van onze werkgroep, maar die worden verticaal geklasseerd, zonder dat er daar iets mee gebeurt. En dat frustreert mij. Dat kan je niet kwijt in een vragenlijst.'*

De verhelderingsgesprekken maken duidelijk dat een betrouwbare gegevensverzameling niet hoeft te betekenen dat iedereen het eens is. Sommigen respondenten lijken onenigheid in een vragenlijst spontaan als een probleem te ervaren. Ze schrijven dit dan nogal eens toe aan tekorten van de gegevensverzameling. Verhelderingsgesprekken maken teamleden en schoolleiders bewust van het bestaan van verschillende zienswijzen in hun school. Daar schuilt echter ook een gevaar in. Er leeft bij sommige respondenten bijvoor-

beeld een gevoel dat de 'harde roepers' hun stempel (te hard) drukken op het verslag van de verhelderingsgesprekken. Er wordt daarom ook aangehaald dat verhelderinggesprekken maar een meerwaarde hebben wanneer ze kwaliteitsvol verlopen. Zoniet kunnen zij een vertekend beeld geven. Men geeft aan dat men de representativiteit van de bevindingen goed moet bewaken. De vaardigheden van de ondersteuners die de gesprekken leiden zijn in die zin zeer belangrijk. Daarnaast wordt er ook op gewezen dat de verhelderingsgesprekken best door een extern iemand begeleid worden. Anders dreigt er eveneens een vertekend beeld te ontstaan.

Tenslotte nog een praktische aanvulling. Als men op basis van de resultaten van verhelderingsgesprekken uitspraken wil doen over verschillende onderdelen van de school, dan blijkt uit de cases dat men goed moet bewaken dat de teamleden de vragenlijst vanuit de juiste invalshoek invullen. Fouten in die zin maken uiteraard dat de verwachtingen van teamleden ten aanzien van de zelfevaluatie bijgesteld worden.

> *Lk: 'Bij de verhelderingsgesprekken begreep ik eigenlijk niet goed waarom ik in groep 'handel' zat. En euhm, ik had verwacht dat ik in 'personenzorg' ging zitten en met als gevolg ook, dat ik heel die enquête opgelost met mijn ervaringen naar personenzorg toe. En bleek dat ik de papieren handel ingevuld had. Ik was ook niet de enige die in de verkeerde groep zat. Allez, als ze mij vragen euhm… wordt er veel zelfstandig gewerkt? Dan is dat voor een 2de graad anders dan voor een 3de graad, mijn antwoord. (...) Maar ik vond die enquête dan eigenlijk een klucht geweest, want… wij schrokken echt van die resultaten, omdat we… allez omdat dat echt niet klopte… maar dan eventjes verder gedacht, ja maar ja, in de 3de graad is dat wel zo. Dus dat was al heel belangrijk.'*

5.2.3 Interpretatie van de gegevens

De gegevensverzamelingen leiden tot een verslag waarin resultaten gepresenteerd worden. Deze resultaten zijn in eerste instantie niet meer dan een beschrijving van de cijfermatige antwoorden van de respondenten. Op basis van deze resultaten dienen scholen tot een interpretatie te komen. In deze paragraaf zetten we eerst de activiteiten op een rij die men in scholen uitvoert om tot een correcte interpretatie te komen. Daarna zoeken we naar elementen die kunnen verklaren waarom het interpretatieproces blijkbaar zo moeilijk van de grond komt.

a. HET INTERPRETATIEPROCES ALS DUSDANIG

In de bestudeerde scholen is het interpretatieproces een activiteit waarbij ofwel enkel de schoolleiding betrokken wordt ofwel gewerkt wordt via een bestaande stuurgroep. Het is echter niet zo dat stuurgroepen per definitie in het interpretatieproces betrokken worden. Een participatieve aanpak van de eerdere zelfevaluatiefasen wordt niet automatisch doorgezet in het interpretatieproces. In de school die speciaal voor de zelfevaluatie een stuurgroep oprichtte, was het uiteindelijk de adjunct-directeur die het interpretatieproces volledig voor zijn rekening nam. Dat gebeurde ongeveer een half jaar nadat het definitieve DISO-rapport in de school ter beschikking was. Het resultaat van zijn interpreta-

tieproces werd afgetoetst bij andere leden van het beleidsteam, maar niet bij leden van de stuurgroep. Ook in de school waar de zelfevaluatie vroegtijdig werd stopgezet werd bij de vorige zelfevaluatie op deze manier tewerk gegaan.

In de school waar de zelfevaluatie gecoördineerd werd door een bestaande werkgroep plande men de interpretatie in de schoot van deze stuurgroep te laten plaatsvinden. Tot op het einde van de onderzoeksperiode was die stap echter nog niet gezet. Dat betekent dat het zelfevaluatierapport na de presentatie van de resultaten reeds 8 maanden wachtte op verdere acties. Op de suggestie van de ondersteuners om de presentatie voor het hele schoolteam (8 maanden eerder) te laten volgen door een collectief interpretatiemoment werd niet ingegaan. Dit zou de rol van de werkgroep immers ondermijnen. Men plande de werkgroep in de maand volgend op deze van de laatste interviewronde opnieuw samen te brengen. De planning was nog steeds dat de werkgroep voor de interpretatie en vervolgens voor het formuleren van prioriteiten en actiepunten zou zorgen. Daarna zou dit alles doorgeschoven worden naar andere werkgroepen in de school. De onderzoeksperiode liet niet toe dit proces verder op te volgen en te staven in welke mate deze plannen gerealiseerd werden.

b. WAAROM HET INTERPRETATIEPROCES ZO MOEIZAAM OP GANG KOMT

In de cases blijkt er telkens veel tijd te gaan over het zetten van de stap van de zelfevaluatieresultaten naar de interpretatie. De respondenten verwachten nochtans dat zelfevaluaties tot betere resultaten zouden leiden indien het proces sneller afgehandeld zou worden. De vertraging in het interpretatieproces wordt dan ook als een probleem ervaren.

> *TL: 'Op den duur gaat dat effect een beetje verloren, want je moet al goed gaan nadenken wat heb ik toen verteld enzo. Dus als dat niet blijft regelmatig herhaald worden, op den duur gaat ge dat volledig verloren moeite zijn. Denk ik.'*
>
> *SL: 'Er is te veel tijd overheen gegaan. Voor velen is dat zo al ver weg, en dan heb ik zoiets van: binnen dat jaar moet je toch proberen dat te brengen, want anders heeft het weinig zin.'*

Ondanks het belang dat men in de gevolgde scholen hecht aan een snelle verwerking wordt de zelfevaluatie na het ontvangen van het rapport blijkbaar nogal eens op de lange baan geschoven. De volgende vaststellingen suggereren dat verschillende elementen hierin een rol kunnen spelen.

- *Men onderschrijft de resultaten niet*

De interpretatie van de resultaten van een zelfevaluatie is een proces waarbij invalshoeken een rol spelen. Dat maakt het moeilijk om tot een eenduidige interpretatie van de resultaten te komen. De bevraagde schoolleiders onderschrijven bovendien niet altijd wat door andere personeelsleden in een vragenlijst naar voren wordt geschoven. Invalshoeken kunnen dus in een gespannen verhouding staan. De personen die verantwoordelijkheid dragen in het interpretatieproces geven bijvoorbeeld aan dat de antwoorden op een

schriftelijke vragenlijst soms een beeld van het schoolbeleid weergeven dat zij niet delen. En toch moeten ze met de resultaten aan de slag.

> *SL: 'Ik denk van goh, mensen, er zijn andere prioriteiten hier op school. Ik denk dat jullie dat ook weten, maar dat zie ik hier (in het zelfevaluatieverslag) niet in staan. Maar dat is eigenlijk niet aan de orde omdat hun verhaal is. Dingen die fout lopen, alles wat ge aan de kant van de directie kunt schuiven, dat wordt daarin vernoemd. Maar dingen die dat te maken hebben met samenwerking. Dat loopt allemaal goed. Terwijl vanuit mijn standpunt.'*

> *SL: 'Ge kunt niet anders dan er iets mee doen denk ik. Allez als ge dat verslag krijgt, dan vloekt ge… maar dat doe je dan met uw collega directeur, dat laat je niet aan de leerkrachten zien hé. Dat spreekt voor zich. Dan begin je te vergaderen en dan begin je constructief met dat verslag te werken. Dat hoort ook wel zo, denk ik. Ook al denkt ge misschien soms wel van zeg, dat is wel… in mijn beleving onterecht hé. Maar het is wel een feit dat, het gros van de leerkrachten dat wel zal gezegd hebben, dus ge moet daar toch iets mee doen. Als dat in hun aanvoelen verkeerd zit, dan moet je daar iets mee doen hé.'*

Ook de bevraagde teamleden voelen iets gelijkaardigs aan. Sommigen geven aan dat je er bij de interpretatie van de resultaten van een schriftelijke bevraging rekening moet mee houden dat een vragenlijst geen instrument is om nuances aan te brengen. De antwoorden in een schriftelijke vragenlijst kunnen onbedoeld te hard en ongenuanceerd overkomen.

> *TL: 'Ik heb altijd de neiging van bijvoorbeeld te nuanceren, maar… dat ging niet hé. Dus je moest ergens een kruisje aanduiden hé. Dan komt dat altijd wel straffer over. Er waren zo voorbeelden, het ging over participatie van leerkrachten of je betrokken wordt bij bepaalde beslissingen, dan zeg ik nee. Maar dan heb ik zoiets van moet dat wel? Maar ja, ik moet nee antwoorden, want ik wordt ook niet betrokken…. He, dus… ik had zo… voor mij kwam precies op mijn antwoorden de directie er niet echt als open en communicatief uit - ze zijn dat ook wel niet - maar ja… ik probeer dat dan altijd goed te praten, maar op die enquête gaat dat niet hé.'*

Het is mogelijk dat zelfevaluatieresultaten minder prioritair aangepakt worden omdat schoolleiders zich er overdreven hard door aangepakt voelen of zich er niet perfect in kunnen vinden (Schildkamp, 2007). Het vergt dan moed om openlijk over de resultaten te communiceren. Men moet zich in zulk een geval immers kwetsbaar durven opstellen. Een mindere score op een bepaald thema is voor schoolleiders overigens niet altijd een signaal om op dat thema verder te gaan werken. Dat is bijvoorbeeld het geval wanneer men vaststelt dat men minder goed scoort op thema's die de voorbije jaren nochtans het onderwerp uitgemaakt hebben van het gevoerde schoolbeleid. Als onterecht ervaren resultaten zaaien bij sommigen twijfel over de waarde van de gegevensverzameling.

> *SL: 'Maar uiteindelijk de conclusies die dat daar nu uitkomen, zijn grosso modo dezelfde als 4 jaar geleden. Terwijl dat toch aan bepaalde dingen daarvan heel hard gewerkt is. Niet alleen door directie, maar door veel mensen. En als ge dat dan gewoon terug ziet opduiken, dan denk ik van ja… Volgens mij zijn dat de thema's die toch sowieso zullen moeten aan bod komen. Communicatie! Ah voi-*

> *là, dat is natuurlijk een superhit hé. (...) Dan denk ik, dat wij over 10 jaar bij een zelfevaluatie nog altijd communicatie… omdat dat nooit niet genoeg kan zijn. Dus allez ja, in dat op zicht, ik geloof wel in die resultaten hoor. Maar ik weet niet of dat dat nu echt zo bruikbare dingen euhm… of echte pijnpunten boven legt. Dat weet ik eigenlijk niet.'*

Sommigen respondenten lijken de resultaten overigens niet volledig te onderschrijven omdat ze vaststellen dat er veel variatie is in de antwoorden van verschillende teamleden. Het instrument geeft dan geen eenduidig beeld en de spreiding in antwoorden maakt het volgens sommigen moeilijk om een oordeel uit te spreken over de resultaten.

> *SL: 'Als scores vind je dan eerder wel, eerder niet, dus alles van die schaal. In het uiteindelijke resultaat, vind ik nog altijd geen eenduidig beeld. Dan wordt er in dat rapport geformuleerd: In de 2de graad handel euhm, is men iets meer van mening, dat er ook naar het proces gekeken wordt dan in de 3de graad personenzorg. Ja, wat heb ik daar nu aan, weet je. Ja, eigenlijk niks.'*

- *Het rapport draagt niet bij tot een vlotte interpretatie*

De respondenten hebben het vaak moeilijk met het onderscheid tussen het beschrijven en het beoordelen van de onderzochte thema's. Terwijl de zelfevaluatierapporten een beschrijving van de bestaande situatie willen geven, verwachten de respondenten dat men in het rapport al een stap verder zou gaan. Een rapport dat enkel de resultaten op een rij zet, voegt niet veel toe in hun mening. Zij blijven dan ook soms wat teleurgesteld achter met het zelfevaluatierapport. Sommigen hebben het gevoel dat de ondersteuners het moeilijke werk hebben laten liggen, waardoor de motivatie om werk te maken van interpretatie afneemt. Voor anderen geldt dit niet.

> *TLZEV: 'Dat was allemaal zo. Allez, als ik dat las, was zo van oké, maar wat is de conclusie daaruit? Want dat zijn zo allemaal wat samenraapsels van een kleine minderheid zegt dit. De grote meerderheid van de 3de graad zegt dat. De leraars van praktijkvakken die zeggen dat. Maar dat is zo, allez ja, zo 3 bladzijden lang over één bepaald gegeven. Maar dan vind ik, wat is dan de conclusie? Of wat haal je er uit, wat is dan het besluit? Dus, ik vond, dat ge daar echt niets aan had. (...) En dan trekt uwe plan hé, zie nu zelf maar welke conclusies dat je wilt trekken.'*

> *SL: 'Maar ik verwacht ook niet dat zij dat zouden verklaren, want dat zou ook straf zijn, dat zij de verklaring kunnen geven, terwijl ze nooit op school zijn, dus die mensen. Ik vind … het voornaamste, dat zij een eerlijk verslag naar ons geven is van: kijk, dit vinden de mensen op die vraag, en ze bedoelen ook effectief dit.'*

- *Een doorlichting maakt de zelfevaluatie minder prioritair*

De stap naar het interpretatieproces (en bij uitbreiding naar het vooropstellen van actiepunten) wordt maar gezet wanneer men de bevindingen van de zelfevaluatie voldoende belangrijk vindt. Men moet een prioriteit kunnen en willen maken van de bevindingen. De dwingendheid van het dagelijkse werk maakt dit moeilijk volgens de bevraagde schoollei-

ders. Ook de combinatie van een zelfevaluatie met een schooldoorlichting brengt deze voorwaarde in het gedrang.

> *SL: ' Dus je hebt hier… ja … twee informatiebronnen krijg je nu binnen, en dat van de doorlichting, ja, dat moet prioritair, daar hebben we geen keuze aan. En dan zit je hier met een hoop informatie, waar je heel goede dingen kan uithalen, van die zelfevaluatie, waar je eigenlijk een beetje op een zijspoor zit. En dat vind ik dan spijtig. (...) Een aantal zaken gaan we kwijtspelen. Persoonlijk, ik denk dat heel veel mensen zich gaan richten op: de doorlichting vindt dat in die afdeling, dat in die afdeling en dat in die afdeling, he, en dat ze dat ter harte gaan nemen. En als we dan gaan zeggen ja maar kijk mensen, vanuit de zelfevaluatie … gaan we die en die dingen doen. Dan gaan veel mensen zeggen: ja maar, dat is wel het minst belangrijke nu he.'*

- *De ondersteuning valt weg*

Onze vaststelling is tenslotte ook dat het wegvallen van ondersteuning niet vreemd is aan de vertraging in het zelfevaluatieproces. De actieve inbreng van ondersteuners eindigt bij het opleveren en eventueel presenteren van het zelfevaluatierapport. Daardoor stopt op dat moment het voorgeschreven traject en staan de scholen volledig zelf in voor de keuzes bij en vormgeving van het verdere verloop van de zelfevaluatie. Bovendien verdwijnt de externe druk om op gezette tijdstippen werk te maken van het zelfevaluatieproces. Een deel van de studiedagen zou volgens een aantal respondenten best volgen als de school het verslag al in handen heeft. Pas dan kan men concreet werken rond interpretatie en het komen tot actiepunten.

5.2.4 Komen tot gepaste actiepunten

Gezien de interpretatie van de zelfevaluatieresultaten reeds lang op zich liet wachten, hoeft het niet te verbazen dat ook het proces van het 'komen tot actiepunten' deels plaatsvond buiten de door ons onderzochte periode. Toch willen we een aantal bevindingen over het komen tot actiepunten presenteren. Verschillende teamleden geven immers adviezen over hoe het formuleren van actiepunten naar aanleiding van een zelfevaluatie best zou kunnen verlopen. Hun adviezen hebben betrekking op de aard van de actiepunten en de manier waarop deze tot stand komen.

> *TL: 'Ik zou er dus een paar concrete punten uit halen, een paar werkpunten. En dan… dit schooljaar 2 werkpunten. Niet te veel in één keer. Een paar dingen, stap voor stap.'*

> *TL: 'Ik denk dat dat het belangrijkste is eigenlijk… Niet te vaag. Echt een concreet werkpunt eruit pikken, één of twee. Ze moeten dat uitwerken in concrete stappen, om dat te verbeteren, en niet de zomaar aan iedereen de vrijheid laten om daar aan te werken.'*

> *TL: 'Tja,… ik zou het echt spijtig vinden… van iedereen te gaan overbelasten met zo van die dingen die echt van die heel grootse zaken zijn, bijna niet te realiseren. Maar van die kleine puntjes, dat vind ik leuk. Ik bedoel, ik denk ook dat de andere mensen daar iets kunnen aan hebben en dat die dan niet direct*

> *gaan zeggen van zeg… de boom in met heel dat zelfevaluatie gedoe, dus weg er mee. Als je dan zegt van, dat en dat was allemaal positief en er zijn een aantal werkpuntjes die we gaan verdelen over de verschillende werkgroepen. En zeggen van: als je er zelf al iets aan kunt doen, aarzel dan maar niet. Maar zoiets zou ik leuk vinden, en dat komt niet over als… weer zoiets.'*
>
> *TL: 'Het moeten punten zijn waar iedereen kan achterstaan dan. Hé. Want anders heeft het ook… als iets van bovenop wordt opgelegd zo… dat werkt toch niet. Het moet van de groep zelf uitkomen. Dus euhm, dat iedereen er een beetje aan wil meewerken. Anders… '*

Uit deze citaten onthouden we dat gepaste actiepunten best concreet uitgewerkt worden. Daarbij moeten volgens de teamleden keuzes gemaakt worden. Enkel een beperkt aantal actiepunten vinden ze werkbaar en daardoor haalbaar. Actiepunten werken volgens de respondenten bovendien het best als ze incrementeel aangepakt worden. De actiepunten moeten richtinggevend zijn voor het team en de realisatie ervan moet kunnen worden opgevolgd. Men wil dan ook dat de actiepunten duidelijk aangeven wat er op welke moment verwacht wordt. Tenslotte geeft men aan dat de teamleden de actiepunten moeten aanvaarden. Op dit laatste aspect gaan we in het vervolg van deze paragraaf dieper in. We gaan na of de schoolleiding waarborgen inbouwt om tot voor de teamleden acceptabele actiepunten te komen.

De schoolleiding speelt in de cases een cruciale rol in het komen tot actiepunten. Ook wanneer de actiepunten in een stuurgroep geformuleerd worden, kunnen zij sterk hun stempel op dat proces drukken. Elementen die volgens hen prioritair zijn, worden bijvoorbeeld sterker in de verf gezet. De stap naar actiepunten wordt in bepaalde gevallen blijkbaar sneller gezet wanneer de resultaten in het plaatje van de beleidsmensen passen. Bovendien aarzelen schoolleiders niet om (sterk) sturend tewerk te gaan.

> *SL: 'Wanneer ik daar zelf niet van overtuigd ben, dan zal ik op dat vlak in een vergadering ook niet enthousiasmerend zijn. Zal ik mensen daar niet in proberen mee te krijgen. Klopt. Ja.'*
>
> *SL: 'Ik heb er al een aantal dingen voor mezelf uitgehaald en in die zin zal ik dat ook wel wat sturen. Dat zal ik ook echt wel proberen te doen. Ik denk dat we dan gaan komen tot 3 à 4 punten waar dat iedereen van zegt: 'Ja, dat is nu wel het belangrijkste.'*

Wanneer men op zoek gaat naar oplossingen voor problemen die door teamleden aangekaart worden, dan schuilt er een gevaar in een te grote sturing door de schoolleiding. Schoolleiders schuiven weliswaar met de beste bedoelingen actiepunten naar voren, maar deze zijn in de ogen van de teamleden niet noodzakelijk de (beste) oplossing voor het probleem dat zich stelt.

> *TL: 'Bijvoorbeeld na die 1ste enquête werden de vergaderingen wat opgedreven. En kregen wij wat meer informatie over hoe spenderen we hier ons geld en hoe worden de uren verdeeld. We kregen dan de cijfers van de kosten van de verwarming en van de kosten van de boeken. Maar dat waren dan de dingen, waar wij zo van zegden van ja, nu mogen we dat weten, maar dat is eigenlijk niet wat dat we willen weten. Dus… we hadden dan… allez, ik persoonlijk, had*

dan zo de indruk van ze gaan nu wel meer open zijn in hun communicatie, maar op de verkeerde aspecten.'

Teamleden in het proces betrekken kan zulke problemen volgens de respondenten helpen vermijden. Dat kan bijvoorbeeld door het inschakelen van bestaande werkgroepen waarin leerkrachten een rol spelen. En door goed te communiceren. Teamleden aanvaarden beslissingen van een centraal beleidsniveau doorgaans op voorwaarde dat deze goed uitgelegd worden. Communicatie is volgens de respondenten dan ook 'de motor' om op basis van zelfevaluatieresultaten tot gepaste actiepunten te komen. De link tussen de zelfevaluatie en de beleidskeuzes wordt best sterk in de verf gezet.

SL: 'De communicatie moet goed zijn. Dan weten mensen ook waarom dat ze die vragenlijst hebben ingevuld en wat dat er dan mee gedaan is. Want dat is wel iets heel frustrerend. Een vragenlijst invullen en 60 slides zien. En dan nog altijd niet doorhebben, van gebeurt hier nu iets mee. Maar ik weet wel van de vorige keer dat dat niet zo eenvoudig is. We hebben dan een zelfevaluatie gedaan, met een aantal duidelijke punten. En wat is daarvan gekomen? Dat is een proces van een paar jaar. We hebben dat de 1ste keer wel duidelijk gedaan hoor. Bijvoorbeeld op de pedagogische studiedag, bij de inleiding daarvan – als dat ging over didactische ondersteuning en werkvormen – eigenlijk altijd de koppeling gelegd met 'in het zelfevaluatieverslag van DISO'. En dat is iets wat dat in de evaluatie van de pedagogische studiedag ook wel altijd op prijs gesteld werd, dat zo duidelijk het hele verhaal terug eventjes werd verteld. Dat is eigenlijk een aandachtspunt ja… het zou meer moeten.'

We geven tenslotte nog mee dat de aard van de informatie in het zelfevaluatieverslag ook een rol kan spelen in het komen tot actiepunten. Sommige zelfevaluatie-instrumenten (zoals IZES) vragen de respondenten immers niet alleen om te beoordelen of de items in de vragenlijst al dan niet van toepassing zijn op hun school. De respondenten worden bijkomend gevraagd of ze van oordeel zijn dat er in de school verandering wenselijk is wat die items betreft. Het beschikken over gegevens omtrent de wens tot verandering kan de keuze van actiepunten volgens een aantal respondenten eenvoudiger maken. Het geeft meteen een indicatie over welke veranderingen op een breed draagvlak kunnen rekenen.

5.3 Resultaten van de zelfevaluatie

In wezen werd de onderzoeksperiode te vroeg afgesloten om een volledig beeld te geven van de resultaten van de bestudeerde zelfevaluaties. Ondanks deze beperking kunnen we uit de gevalstudies een aantal waardevolle lessen trekken omtrent de resultaten van de zelfevaluaties. We focussen in het volgende eerst op het onderscheid tussen directe en indirecte resultaten. Daarna brengen we in kaart welke resultaten de bevraagde schoolleiders en teamleden rapporteren. Vervolgens staan we stil bij de vraag hoe we de beperkte resultaten door de ogen van teamleden kunnen verklaren.

5.3.1 Een onderscheid tussen directe en indirecte resultaten

Er is heel wat evidentie voor het belang van de resultaten van zelfevaluatieprocessen. De respondenten geven aan dat een zelfevaluatie in de eerste plaats een middel is dat indirect tot de realisatie van een aantal wenselijke geachte verwachtingen moet bijdragen. Met indirect bedoelen we via het gebruik van de zelfevaluatiegegevens. Een zelfevaluatie is dan maar kwaliteitsvol als met de gegevens iets gedaan wordt.

> *SL: 'Ge kunt een bevraging doen, maar uiteindelijk doe je dat ook maar in functie van acties die dat er uit gaan volgen, en… Euhm, als die achterwege blijven, ja dan, werpt dat toch wel wat ander licht op heel dat zelfevaluatieproces.'*
>
> *TL: 'Niet zomaar, van vult dat eventjes in en we gaan daar eens over vergaderen. En dan steken we dat in kast, tot dat de doorlichting komt. Dat daar echt wel gevolg aan gegeven wordt en dat de mensen dat ook willen serieus nemen.'*

De respondenten benadrukken het belang van het streven naar resultaten, en dus het indirecte effect van zelfevaluatie, zeer sterk. Er is daarnaast evenwel ook evidentie voor een direct effect. Door het gebruik van een schriftelijk vragenlijst en verhelderingsgesprekken creëert men volgens sommigen bijvoorbeeld meteen een vorm van inspraak. Los van het gebruik van de resultaten is de gegevensverzameling in deze logica op zich reeds een instrument om teamleden hun stem te laten horen. Zelfs als het zelfevaluatierapport verder ongebruikt blijft, is het zelfevaluatieproces waardevol geweest. Naast dit – zo zal in de volgende paragraaf blijken - bediscussieerbare directe 'inspraakeffect' zijn er in de cases nog andere directe effecten. Daarbij denken we bijvoorbeeld aan het feit dat het invullen van een vragenlijst mensen aan het denken kan zetten en dat verhelderingsgesprekken op zich reeds leerrijke ervaringen kunnen zijn.

> *Lk2: 'Want ik was daar van geschrokken, dat ik daar zo tegenover stond. Ik dacht dat dat minder erg was. Want ik ben hier gelukkig op school en ik werk hier graag. Ik zal niet zeggen dat er mij hier iets stoort, maar door die vragenlijst ben ik wel beginnen nadenken. Dat stoort mij toch eigenlijk wel, ja…'*
>
> *TLZEV (over de verhelderingsgesprekken): 'Dus ieder heeft z'n eigen mening gegeven, er is over gepraat. Er is dus een duidelijk, gemeenschappelijk standpunt op dat moment, naar voor gekomen. Dus dat is… dat is heel goed vind ik.'*

In de volgende paragraaf focussen we verder op de indirecte effecten van zelfevaluaties. Het gros van de verwachtingen en resultaten van de bestudeerde zelfevaluaties heeft immers betrekking op hoe de school met de zelfevaluatiegegevens omgaat.

5.3.2 Indirecte resultaten in de bestudeerde casussen

In de bestudeerde cases valt bovenal op dat de inschatting van de resultaten door teamleden grondig verschilt van de inschatting door schoolleiders. Deze laatste spreken zich systematisch positiever uit over de effecten van de zelfevaluaties.

a. RESULTATEN DOOR DE OGEN VAN SCHOOLLEIDERS

De bevraagde schoolleiders zijn van mening dat de zelfevaluatie in hun school tot waardevolle resultaten geleid hebben of zullen leiden. De manier waarop ze aangeven gebruik te maken van de zelfevaluatieresultaten varieert evenwel. We vinden evidentie voor een instrumenteel, conceptueel, symbolisch en inspraakgericht gebruik.

- *Instrumenteel: zelfevaluatie als basis voor beleidsbeslissingen*

Het instrumenteel gebruik van de zelfevaluatieresultaten maakt dat men de zaken anders aanpakt dan voorheen. In het instrumenteel gebruik gaat het in de eerste plaats om het komen tot actiepunten. Schoolleiders geven aan zowel te komen tot actiepunten voor zichzelf als tot actiepunten voor de school als geheel. De actiepunten die de bevraagde schoolleiders voor zichzelf vooropstellen blijven voor anderen vaak onzichtbaar. De schoolleider past in dat geval bijvoorbeeld de eigen aanpak en leiderschapsstijl aan zonder andere teamleden daarvan op de hoogte te brengen. De actiepunten voor de school als geheel hebben bijvoorbeeld betrekking op de verticale communicatie, de werking van werkgroepen, de opvolging van het schoolreglement, de aandacht voor onderwijskundige thema's zoals 'leren leren' en dergelijke. Actiepunten daaromtrent vertalen zich bijvoorbeeld in het opzetten van pedagogische studiedagen, in aangepaste afspraken, in vernieuwde werkwijzen, in het oprichten van werkgroepen of in een aanpassing van het schoolwerkplan.

- *Conceptueel: zelfevaluatie als beïnvloeding van het denken van schoolleiders*

We spreken van conceptuele resultaten wanneer het gebruik van de zelfevaluatie enkel een impact heeft op het denken van mensen. Het gaat voornamelijk om het bijstellen of bevestigen van het beeld dat men heeft van de school en om het detecteren van ongekende problemen. De bevraagde schoolleiders omschrijven het feit dat men een representatief beeld krijgt van hoe anderen denken over de bevraagde thema's als een waardevol resultaat van de zelfevaluatie. De zelfevaluatieresultaten kunnen vervolgens bijdragen tot een vernieuwd of een bevestigd beeld van wat er leeft in de school. Conceptuele resultaten betekenen dus niet dat men het beeld dat men heeft van de school simpelweg aanpast aan het beeld dat in het zelfevaluatierapport naar voren komt. Dat beeld is volgens de bevraagde schoolleiders immers ook maar een perceptie en is niet noodzakelijk objectief. Schoolleiders geven ook aan dat de zelfevaluatieresultaten de aandacht weten te richten op zaken die voorheen een blinde vlek waren. Daarnaast kunnen zelfevaluaties ook de intuïties omtrent bepaalde problemen bevestigen. In de dagelijkse praktijk doen schoolleiders immers voortdurend vaststellingen die maken dat men aanvoelt dat er aan bepaalde zaken iets moet gebeuren of die maken dat men vaststelt bij bepaalde zaken te weinig stil te staan. De informatie die men door het uitvoeren van een zelfevaluatie kan verzamelen, leidt volgens de bevraagde schoolleiders ook op die manier tot nieuwe inzichten.

- *Symbolisch: zelfevaluatie als verantwoording van bestaande standpunten en veronderstellingen*

Onder symbolisch gebruik verstaan we dat de resultaten eerder een argument zijn in het beleid dan een actiepunt als dusdanig. Dat is bijvoorbeeld het geval wanneer schoolleiders hun vernieuwde inzichten niet voor zichzelf houden maar ook gaan uitdragen naar de rest van het team. Wanneer de zelfevaluatie zou pleiten voor verandering dan vinden de bevraagde schoolleiders in het zelfevaluatierapport bijvoorbeeld een argument om de teamleden te wijzen op hun verantwoordelijkheden. Precies omdat de teamleden zelf de pijnpunten naar voren schuiven, verwacht men van hen ook een bijdrage in het formuleren van oplossingen. Men vindt in de zelfevaluatieresultaten soms evenwel ook een argument om de bestaande werking te bestendigen. Het resultaat van de zelfevaluatie is dan een formele legitimatie van het beleid dat gevoerd wordt.

> *SL: 'Zo een procedure is goed omdat ge het allemaal wat formeler maakt. Maar de punten die dat daaruit naar voorkomen. Eigenlijk had ik die ook wel op voorhand kunnen formuleren denk ik. Dus, maar het is goed, dat het allemaal wat formeler gemaakt wordt.'*

Wanneer de zelfevaluatieresultaten aangeven dat de zaken goed draaien, dan ziet men daarin een bevestiging van de waarde van gemaakte keuzes en de gehanteerde strategieën. De stap naar een instrumenteel gebruik is, zo blijkt uit deze cases, dan minder aan de orde.

> *SL: 'Want echt slechte scores zaten er nog niet veel in hé. En dan bedoel ik slechte scores, waar dat ge van zegt: 'Hier moeten we dringend werk van maken'. Want al de resultaten waren boven het gemiddelde enz. het was dus allemaal wreed goed. Maar euhm, voor de rest is daar niet zo heel veel mee gebeurd.'*

- *Inspraakgericht: zelfevaluatie als strategisch hulpmiddel in een participatief schoolbeleid*

We vermeldden hoger reeds dat sommige respondenten spreken van een direct inspraakeffect. Anderen onderschrijven de mogelijkheden van zelfevaluaties als strategie om inspraak mogelijk te maken, maar koppelen dit aan de voorwaarde van indirecte effecten. Daarmee bedoelen ze dat zelfevaluaties door teamleden alleen als een gewaardeerde vorm van inspraak gepercipieerd worden indien er eveneens sprake is van instrumentele, conceptuele en/of symbolische resultaten. Enkel als de zelfevaluatie effectief tot één van bovenstaande resultaten bijdraagt, wordt het proces volgens deze respondenten door teamleden zelf gepercipieerd als een strategisch instrument om tot inspraak in het schoolbeleid te komen.

> *SL: 'Want, over zelfevaluatie kun je… eigenlijk is dat een heel schoon verhaal hé. Van een vragenlijst, en iedereen mag er op antwoorden en dan doen we nog gesprekken over. Uiteindelijk die theorie is heel mooi, ik denk dat… dat bij ons op school, dat ook heel mooi is. Maar als ge dan gaat kijken naar en welke acties zijn daar nu uit gevolgd… dat is toch wel een mager beestje hoor.'*

SL: Als mensen het gevoel hebben dat er iets gedaan wordt met hetgeen dat zij gezegd hebben... Als ze het gevoel hebben van goh, heel die zelfevaluatie dat is toch niet gewoon voor de show, maar daar wordt ook wel echt iets mee gedaan, dan vind ik dat het goed is. Gewoon dat ze merken dat er iets mee gedaan wordt en dat echt ook merken. Als ze dan nog zouden zeggen, we merken dat er iets mee gedaan wordt… dan zou het, dan is het fantastisch…'

b. DE RESULTATEN DOOR DE OGEN VAN TEAMLEDEN

Bij de bevraagde teamleden weerklinkt een heel ander geluid. Zij geven vaak aan na de presentatie van de zelfevaluatieresultaten niets meer over het zelfevaluatieproces te hebben vernomen. Ondanks het feit dat er in de school naar aanleiding van de zelfevaluatie initiatieven genomen zijn of nog gaan genomen worden, is hun oordeel (eerder) negatief. Zij vinden dat de gegevensverzamelingen tot onvoldoende resultaten geleid hebben en dat er kansen gemist werden. Uit de interviews blijkt overigens dat teamleden reeds tijdens het proces van de zelfevaluatie er sterk aan twijfelen of de zelfevaluatie de verwachtingen zal kunnen inlossen. Dat blijkt uit de volgende citaten. Merk bij het lezen ervan op dat het niet ondenkbaar is dat de houding van teamleden positiever zou zijn, moest de laatste interviewronde een paar maanden later hebben plaatsgevonden.

TL: 'Ja, het is weer een gemiste kans eigenlijk hé. Dat gebeurt dus regelmatig in de school. Er wordt aan iets begonnen, dat is ook typisch voor misschien alle scholen. Een initiatief genomen, dat wordt niet afgewerkt. Er lopen misschien 5 of 10 zaken tegelijk, maar er wordt nooit iets afgewerkt hé. Dat werkt ook op den duur een beetje negatief, want dan zeggen leerkrachten zeg gast, dat is weer iets nieuws…'

TL: 'Tot nu toe heb ik er mijn twijfels over… ik heb nog niks concreet, we zijn al een jaar verder, het is nog niks … allez ja … ik weet toch van niks, dus ik heb … ik ben niet echt tevreden over. Ik had er meer van verwacht.'

TLZEV: 'Ik zou echt niet weten wat. Voor mij persoonlijk denk ik. Of voor collega's uit dezelfde groep van algemeen vakken. Ik zou echt niet weten wie ervan gezegd heeft van 'Hmm, dat weet ik nu of dat vind ik wel goed'. Niemand. Dus echt, niks. Dus ik weet niet, misschien dat er iemand daar wel iets heeft uitgehaald maar ik heb dat nog niet gehoord, dat iemand daar wel iets aan gehad heeft.'

De bevraagde teamleden spreken zich in de laatste interviewronde dus behoorlijk negatief uit over de resultaten van de zelfevaluatie. Het proces sleept te lang aan, het wordt onvoldoende gecoördineerd en opgevolgd, de resultaten blijven uit of de communicatie over de resultaten voldoet niet. De oorzaak van deze tekorten wordt door hen bij de schoolleiding gelegd. Er is ook geen evidentie in de cases voor eigen initiatieven door teamleden om zelf lessen te trekken uit het zelfevaluatieverslag. Zij sluiten evenwel niet uit dat anderen in de school met de zelfevaluatieresultaten aan het werk zijn gegaan. Opmerkelijk, tenslotte, is dat deze beperkte betrokkenheid op de zelfevaluatieresultaten soms ook tot gevolg kan hebben dat men beleidskeuzes onterecht linkt aan de zelfevaluatie. Eén van de geïnterviewde teamleden beschouwde een aantal initiatieven die in haar school genomen werden immers onterecht als een resultaat van de zelfevaluatie.

c. VERKLARINGEN VOOR DE BEPERKTE RESULTATEN IN DE OGEN VAN TEAMLEDEN

We stelden vast dat teamleden de resultaten van zelfevaluaties doorgaans erg gering in schatten. Zelfs wanneer schoolleiders aangeven dat de zelfevaluatie tot waardevolle resultaten leidde. De bevraagde schoolleiders pleiten zelf voor duidelijke en door het team gedragen zelfevaluatieresultaten. Ze zijn van mening dat het belangrijk is leerkrachten te tonen wat het resultaat is van een zelfevaluatie. Dat is blijkbaar niet eenvoudig. Ondanks de bekommernissen van hun schoolleiders leggen de bevraagde teamleden de link tussen actiepunten en de zelfevaluatie immers zeer moeilijk. Hieronder zetten we een aantal mogelijke verklaringen hiervoor op een rij.

- *De resultaten hebben vaak voornamelijk betrekking op het schoolbeleid*

Zelfevaluaties hebben vaak implicaties op domeinen die (op het eerste zicht) verder afstaan van de dagelijkse activiteiten van teamleden. Dat vergroot de afstand tussen de zelfevaluatie en de dagelijkse praktijk van teamleden. De inhoudelijk aard van de resultaten van de zelfevaluatie kan in die zin mee verklaren waarom teamleden in vergelijking met schoolleiders de indruk hebben dat de resultaten beperkt zijn. Met de inhoudelijk aard bedoelen we dat zelfevaluatieresultaten niet altijd zichtbaar zijn voor de teamleden. Bijvoorbeeld: een school heeft naar aanleiding van een eerdere zelfevaluatie een structuur uitgebouwd die verschillende werkgroepen integreert. Hoewel het voor de schoolleiding duidelijk is dat dit een antwoord was op de vaststellingen uit de zelfevaluatie, maken teamleden in de school deze link niet.

- *De resultaten worden vaak niet geëxpliciteerd*

Naast de inhoud speelt ook de aanpak van scholen echter een rol. De link tussen de actiepunten naar aanleiding van de zelfevaluatie en de eigenlijke zelfevaluatie wordt doorgaans niet geëxpliciteerd. Sommige schoolleiders zijn er niet van overtuigd dat het nodig is doelbewust te communiceren over de actiepunten die uit de zelfevaluatie voortvloeien. Anderen zijn dat wel. Om de resultaten van de zelfevaluatie tot leven te brengen en levendig te houden, proberen zij om waar mogelijk expliciet de link te leggen met de activiteiten die in de school ondernomen worden.

> *SL: 'Nee, de mensen gaan dat niet zien. Het is aan ons nu. Wij sommen dat een keer op, van kijk mensen dat zijn de resultaten dat we bereikt hebben. Maar wij gaan daar niet zeggen, alwel dat kwam van DISO van jullie uit.'*

> *SL: 'En ik vind het ook belangrijk met die resultaten, dat we er ook iets mee doen. Vandaar dat in de werkgroepen en op studiedagen, ook gerefereerd wordt naar IZES. In die zin dat de intro van die studiedag was: "kijk mensen, we hebben vastgesteld via het onderzoek IZES, herinner je u vorig jaar, vorig studiejaar. Van dat, dat, dat…" Dat zij ook beseffen van waar komt dat.'*

De onduidelijkheid bij de bevraagde teamleden over de resultaten van de zelfevaluatie wordt nogal eens toegeschreven aan het verstrijken van teveel tijd tussen de gegevensverzameling en verdere activiteiten naar aanleiding van de zelfevaluatie. Toch hoeft dat niet volgens iedereen een probleem te zijn. Tenminste als men bewaakt dat het team

dankzij de zelfevaluatie het gevoel heeft ernstig te worden genomen. De manier waarop men de link in de verf zet, is dan ook belangrijker dan de snelheid waarmee dit gebeurt.

> *SL: 'Bijvoorbeeld bij de pedagogische studiedag in de jaren na de vorige zelfevaluatie. Als dat ging over didactische ondersteuning en werkvormen - werd eigenlijk altijd de koppeling gelegd met 'in het zelfevaluatieverslag van DISO'. En dat is iets wat dat in de evaluatie van de pedagogische studiedag ook wel altijd op prijs gesteld werd, dat zo duidelijk het hele verhaal terug eventjes werd verteld.'*

Een aantal schoolleiders verwachten hiervan een positief effect. Men mag echter niet uit het oog verliezen dat deze verwachting waarschijnlijk wel eens alleen terecht zou kunnen zijn wanneer teamleden een positief gevoel hebben bij heel het zelfevaluatiegebeuren.

> *TL: 'Goh, ik vind dat niet echt nodig. Allez, ik vind als die… als het verbetert verbetert het hé. en misschien dat ge wel door dat woord (het teamlid bedoelt 'DISO-zelfevaluatie'), mensen er tegen… allez, tegen die maatregelen gaat opzetten. Omdat ze weten 'DISO'... Sommigen zeggen wel 'Ah dat is dat spel waar dat we moeten voor blijven zijn al een paar keer' of 'daar zijn ze weer met hun DISO'. Allez ja. dus bij sommige brengt dat ook wel negatieve ervaringen mee hé.'*

- *Eerdere ervaringen van teamleden zorgen voor 'realistische' verwachtingen.*

Of teamleden het gevoel hebben dat zelfevaluaties al dan niet tot waardevolle resultaten geleid hebben, hangt tenslotte ook af van vroegere ervaringen. Het gaat dan zowel om ervaringen met betrekking tot veranderingsmanagement in het algemeen als om ervaringen met betrekking tot zelfevaluatie in het bijzonder.

> *TLZEV: 'En dat is gevaarlijk voor elk initiatief dat ge in de toekomst neemt. Maar dat is mijn visie. Ik denk als ge een onderzoek doet en je bekomt resultaten, dan moet je met die resultaten iets doen, want anders hypothekeer je uw volgend onderzoek.'*

Op dit moment is de term 'realistisch' blijkbaar vaak een eufemisme voor 'beperkt'. Onze respondenten verwachten niet dat de zaken door deze zelfevaluatie plots anders zullen gaan. Dat maakt het expliciteren van de resultaten van zelfevaluatie tegelijkertijd zo belangrijk en zo moeilijk in scholen waar er geen sprake is van een innovatieve en reflectieve cultuur. Mensen zijn blijkbaar geneigd te zien wat ze verwachten te zien.

6 CONCLUSIES EN THEORETISCHE REFLECTIES

In de laatste paragraaf van dit deel komen we tot een aantal conclusies en geven we stof tot discussie. In eerste instantie gebruiken we de onderzoeksvragen van de gevalstudie daarbij als structurerend kader (zie paragraaf 6.1). De vooropgestelde hypothesen worden met andere woorden aan de empirische evidentie getoetst. In de daarop volgende paragraaf (6.2) komen tot een aantal theoretische reflecties. We stellen daar de vraag

welke aanpassingen en aanvullingen bij het theoretisch kader aangewezen lijken op basis van de bevindingen uit de gevalstudies.

6.1 Een antwoord op de onderzoeksvragen

De drie onderzoeksvragen die wij bij de start van deze gevalstudie vooropstelden komen in het volgende elk aan bod. We staan eerst stil bij de vraag aan de hand van welke indicatoren men de kwaliteit van zelfevaluaties kan operationaliseren. Daarna richten we ons op het verloop van zelfevaluaties. We gaan na in welke mate de bestudeerde zelfevaluaties tegemoet komen aan de dragers van 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad'. We staan daarbij stil bij de link tussen het verloop van zelfevaluaties en de resultaten van zelfevaluaties. We ronden af met conclusies omtrent de behoefte aan en impact van ondersteuning.

We willen benadrukken dat de volgende conclusies gebaseerd zijn op ervaringen inzake zelfevaluatie in drie scholen voor technisch en beroepssecundair onderwijs. De vaststellingen en verklaringen moeten in die context gesitueerd worden. Verdere generalisering is op basis van onze onderzoeksactiviteiten niet mogelijk.

6.1.1 Aan de hand van welke indicatoren kan men de kwaliteit van een zelfevaluatie operationaliseren?

Centraal in het kwaliteitsconcept dat we in deze studie hanteren, staat het concept 'verwachtingen'. In het theoretisch kader (en later ook in de kwantitatieve analyses) vonden we aanknopingspunten om drie verschillende types van verwachtingen te hanteren. Wanneer zelfevaluaties aan deze verwachtingen tegemoet komen worden ze omschreven als kwaliteitsvol. De gevalstudies leveren voornamelijk bewijskracht op voor een respondenteigen kwaliteitsindicator en de schooloverstijgende indicator 'komen tot passende actiepunten'. We stelden ook vast dat de verwachtingen van schoolleiders anders liggen dan deze van teamleden. Dat brengt ons bij een aantal conceptuele en operationele conclusies. De conceptuele conclusies richten zich in eerste instantie op de (verschillende) verwachtingen van schoolleiders en teamleden. Het gaat hier om de vraag wat te verstaan onder het concept 'kwaliteitsvolle zelfevaluatie'. In het operationele luik stellen we ons vervolgens de vraag of en hoe dit concept gemeten kan worden.

a. CONCEPTUELE CONCLUSIES: VOLDOEN AAN WELKE VERWACHTINGEN?

De gevalstudies tonen aan dat de respondenten verscheidene verwachtingen koesteren bij het initiëren van een zelfevaluatie (cf. hypothese 1). Precies deze verscheidenheid zette ons er eerder reeds toe aan om te spreken van een respondenteigen kwaliteitsindicator. De kwaliteit van zelfevaluaties kan door verschillende betrokkenen in een school zeer verschillend gepercipieerd worden. De oorzaak hiervan moet onder meer gezocht worden in de verschillende verwachtingen. Voornamelijk de verschillen tussen teamleden en schoolleiders hebben in die zin conceptuele implicaties.

- *Teamleden: De respondenteigen verwachtingen zijn gelijk aan 'resulteren in concrete actiepunten'*

Teamleden werden in de bestudeerde scholen niet betrokken bij het vooropstellen van zelfevaluatiedoelen. Het hoeft ons dan ook niet te verwonderen dat zij het vaak moeilijk hebben om de verwachtingen onder woorden te brengen en dat ze de verwachtingen vaak weinig concreet invullen. Precies hierdoor blijkt er evenwel grote overeenstemming in hun verwachtingen. Teamleden verwachten van zelfevaluaties immers dat ze resulteren in gepaste actiepunten. Ze verwachten van een zelfevaluatie dat deze resulteert in specifieke, meetbare, acceptabele, realistische en tijdsgebonden actiepunten. Ook hier duikt het SMART-principe dus op. Op welke beleidsdomeinen deze actiepunten betrekking behoren te hebben of welke concrete vertaling deze actiepunten behoren te krijgen, laten teamleden in de bestudeerde scholen in het midden. De respondenteigen benadering van kwaliteit valt voor de bevraagde teamleden dan ook samen met de verwachting dat de zelfevaluatie 'resulteert in gepaste actiepunten'.

- *Schoolleiders: De respondenteigen verwachtingen zijn ruimer dan 'resulteren in concrete actiepunten'*

Hoewel de verwachtingen ook door de bevraagde schoolleiders niet doelbewust geëxpliciteerd worden, zijn de functies van een zelfevaluatie in hun ogen wel duidelijker omschreven. Schoolleiders verwijzen in hoofdzaak naar de spiegelfunctie en de beleidsvoorbereidende functie van zelfevaluaties. Zelfevaluaties moeten volgens de bevraagde schoolleiders een objectief beeld geven van de school waarop men vervolgens de beleidsvoering kan enten. Enkel tussen de lijnen wordt aangegeven dat zelfevaluaties ook uitgevoerd worden met het oog op verantwoordingsgerichte functies, zoals anticiperen op verwachtingen van een doorlichting. In vergelijking met de spiegel- en de beleidsvoorbereidende functie is de verantwoordingsfunctie evenwel van beperkt belang bij het initiëren van een zelfevaluatie. De eerste twee functies zijn volgens schoolleiders de werkelijke reden om een zelfevaluatie uit te voeren; de verantwoordingsfunctie is bij wijze van spreken mooi meegenomen.

In lijn met deze functies stellen de bevraagde schoolleiders ook concrete verwachtingen voorop. Net zoals bij hun teamleden gaat het dan in de eerste plaats om het komen tot gepaste actiepunten. Daarnaast willen zij het gesprek aangaan met hun team. Zelfevaluaties zijn in de ogen van schoolleiders een instrument om participatie in het schoolbeleid te realiseren. Andere concrete verwachtingen hebben betrekking op het responsabiliseren van teamleden en het zoeken naar oplossingen voor concrete problemen.

We concluderen voor de drie bestudeerde cases dat de verwachtingen van schoolleiders zowel ruimer als specifieker zijn dan deze van teamleden. Verder stellen we vast dat er bij de schoolleiders precies daardoor minder onderlinge overeenstemming is. Deze vaststelling bevestigt de hypothese over de variatie in de verwachtingen ten aanzien van zelfevaluaties (cf. Hypothese 1).

b. OPERATIONELE CONCLUSIES: KUNNEN WE DEZE VERWACHTINGEN METEN?

De conceptuele conclusies hebben belangrijke implicaties voor het operationele luik van meta-evaluatie. Het meten van de kwaliteit van zelfevaluaties blijkt immers een ontzettend complexe onderneming. Onze hypothese (cf. 8) over de mogelijkheden tot meta-evaluatie vergt dus nuancering. De gevalstudies tonen aan dat het zeer moeilijk is om volledig recht te doen aan de eigenheid van zelfevaluaties. Toch dringt de vraag naar meta-evaluatie zich op. Ons antwoord vertrekt vanuit een bewustzijn van de limieten in het beoordelen van zelfevaluaties en zoekt van daaruit naar werkbare kwaliteitsindicatoren. In de volgende paragrafen lichten we de twee operationele houdingen die men kan aannemen toe. Men kan enerzijds uitgaan van de houding dat het onmogelijk is het onvatbare te meten. Anderzijds kan men tevreden zijn met het meten van wat wel 'vatbaar' is. Een 'problematiserende' versus een 'pragmatische' aanpak dus.

- *De problematiserende houding: Meta-evaluatie als het 'onvatbare' meten*

De bevindingen uit de gevalstudies brengen een aantal vaststellingen aan het licht die uitspraken over de kwaliteit van zelfevaluatie bemoeilijken. Uitspraken doen over de kwaliteit van zelfevaluaties is een probleem:

1. *omdat het onduidelijk is wiens verwachtingen en oordeel als basis moet genomen worden.* Zelfevaluaties zijn in de cases in de eerste plaats een activiteit van de schoolleiding. Dat roept de vraag op of dit rechtvaardigt dat indicatoren voor de kwaliteit van zelfevaluaties voorbijgaan aan het oordeel van de teamleden. Indien de zelfevaluatie enkel opgezet wordt voor gebruik door de schoolleiding – en dit ook zo aan de teamleden gecommuniceerd wordt - is het wellicht correct om aan het oordeel van teamleden voorbij te gaan. Ons inziens kan een zelfevaluatie dus kwaliteitsvol zijn, zonder dat teamleden weet hebben van de resultaten. Zulke zelfevaluaties kunnen immers perfect aan de verwachtingen van schoolleiders én teamleden voldoen. Precies daarom is de respondent- of schooleigen benadering van belang. Of dit ook een wenselijke situatie is, laten we in onze zoektocht naar kwaliteitsindicatoren in het midden. We stelden evenwel vast dat schoolleiders doorgaans aangeven dat de zelfevaluatie een proces zou moeten zijn waarin ook de teamleden een actieve rol spelen. Kunnen kwaliteitsindicatoren gegeven deze verwachtingen voorbij aan de vraag of ook de verwachtingen van teamleden werden ingelost?
2. *omdat het onduidelijk is welke verwachtingen als basis moet genomen worden.* We concludeerden reeds dat de verwachtingen van de bevraagde teamleden samen te vatten zijn als het komen tot passende actiepunten. Voor hen zou dit een algemeen aanvaarde kwaliteitsindicator kunnen zijn. Bij schoolleiders liggen de verwachtingen ruimer en specifieker. Dat maakt het voor deze laatste groep moeilijk om a priori en schooloverstijgend te bepalen welke verwachtingen de basis zouden moeten vormen van het kwaliteitsvoordeel. Tenzij een zelfevaluatie enkel uitgevoerd zou worden naar aanleiding van externe verwachtingen. Is de conclusie dat we enkel in dat geval extern opgelegde verwachtingen als basis kunnen nemen terecht?

3. *omdat verwachtingen zelden geëxpliciteerd worden.* De bevraagde scholen stellen bij het uitvoeren van zelfevaluaties wel doelen en verwachtingen voorop maar deze worden niet systematisch geïnventariseerd en gedocumenteerd. Dat maakt het na afloop moeilijk om te achterhalen welke verwachtingen de basis van een kwaliteitsbeoordeling zouden moeten vormen. Kan de basis voor kwaliteitsbeoordeling dan enkel bestaan uit expliciete doelen of moeten ook onuitgesproken doelen een rol spelen? En wat met onuitspreekbare doelen?

4. *omdat verwachtingen kunnen evolueren doorheen het zelfevaluatieproces.* We hebben in de gevalstudies vastgesteld dat de verwachtingen die men bij de start vooropstelt niet noodzakelijk identiek zijn aan de verwachtingen bij het einde van het proces. Ook hier is de vraag naar de beoordelingsbasis problematisch. Is een zelfevaluatie die wel aan de verwachtingen bij het einde tegemoet komt maar niet aan de verwachtingen die men bij de start van het proces vooropstelde kwaliteitsvol of niet?

5. *omdat het al dan niet tegemoet komen aan de verwachtingen ook afhangt van de resultaten van de zelfevaluatie.* Zelfevaluatiegegevens waaruit blijkt dat scholen goed functioneren en die niet oproepen tot het opnemen actiepunten vergen een andere reactie van scholen dan zelfevaluatiegegevens die op manifeste tekorten wijzen (Visscher, 2002; Schildkamp, 2007). Is het dan correct om een zelfevaluatie die - in het eerste geval – niet tot gepaste actiepunten leidt per definitie als niet-kwaliteitsvol te bestempelen?

Deze bedenkingen zetten ons aan tot voorzichtigheid. Er zijn duidelijk grenzen wanneer men uitspraken wil doen over de kwaliteit van zelfevaluaties die scholen in het kader van schoolontwikkeling uitvoeren. De uitdaging bestaat erin om binnen deze grenzen te komen tot een pragmatische manier van werken. Deze uitdaging gaan we aan in de volgende paragraaf.

➢ *De pragmatische houding: Meta-evaluatie als het 'vatbare' meten*

Pragmatisch betekent gericht op nut en bruikbaarheid. Vermits zelfevaluaties een groeiende rol van belangrijkheid zullen spelen in het Vlaamse kwaliteitszorgsysteem, kunnen we niet voorbij aan de behoefte aan indicatoren om die kwaliteit in kaart te brengen. Dat dit met een aantal beperkingen - en bijgevolg ook terechte kritieken - gepaard zal gaan, is gegeven de complexiteit van het gebeuren onvermijdelijk. Voortbouwend op de lessen uit de gevalstudies schuiven wij de volgende raadgevingen naar voren.

1. *Maak een onderscheid tussen de kwaliteit door de ogen van teamleden en de kwaliteit door de ogen van schoolleiders.* Schoolleiders zijn momenteel op een andere manier bij het proces- en het resultaat van zelfevaluaties betrokken dan hun teamleden. In het uitspreken van een oordeel maakt men daarom best een onderscheid tussen het kwaliteitsoordeel van beide groepen. Het gaat dan om twee op zich staande en complementaire oordelen.

2. *Focus op de grootste gemene deler.* Het oordeel over de kwaliteit van zelfevaluaties hoeft bij schoolleiders en teamleden niet terug te gaan op dezelfde verwachtingen.

Het lijkt aangewezen om voor beide groepen op zoek te gaan naar de grootste gemene deler. Bij teamleden ligt dit voor de hand. Hun verwachtingen zijn dan dat de zelfevaluatie moet 'resulteren in gepaste actiepunten'. Voor hen is dat – uitgaande van de vaststellingen in deze gevalstudie - momenteel de enige indicator. Bij schoolleiders ligt dit moeilijker. Naast het 'resulteren in gepaste actiepunten' – die indicator kan dus ook bij schoolleiders gebruikt worden - zetten zij zeer meer verscheiden en individuele doelen voorop. Een concrete invulling geven aan deze verwachtingen is zeer moeilijk wanneer men over scholen heen uitspraken wil doen. De volgende twee indicatoren, beide met eigen mogelijkheden en beperkingen, kunnen in deze context een operationalisering zijn.

- *De abstracte indicator: Afstappen van de concrete verwachtingen.* Men zou de kwaliteitsbeoordeling door de schoolleider kunnen beperken tot uitspraken die niet naar concrete verwachtingen verwijzen. De vraag is dan in welke mate de verwachtingen volgens schoolleiders zijn ingelost, los van de inhoud van de verwachtingen in kwestie. Deze werkwijze speelt in op het respondenteigen karakter van verwachtingen. Wat voor één schoolleider een kwaliteitsvol resultaat is, hoeft dat voor een andere niet te zijn. Door de kwaliteitsbeoordeling op een abstract niveau te laten plaatsvinden, kunnen zelfevaluaties met zeer verschillende resultaten toch als even kwaliteitsvol beoordeeld worden. Of andersom: zelfevaluaties die tot dezelfde resultaten leidden kunnen door verschillende schoolleiders verschillend beoordeeld worden. De sterkte van deze indicator is dus meteen ook haar zwakte. Merk bovendien op dat het hebben van geen verwachtingen ook een verwachting is.

- *De concrete indicator: focussen op gedeelde concrete resultaten.* De andere benadering bestaat erin om toch te focussen op concrete verwachtingen van schoolleiders. Ook daarin kunnen we op zoek gaan naar de grootste gemene deler. Het aantal bevraagde schoolleiders is uiteraard te klein om uitspraken te doen over de veralgemeenbaarheid van de vooropgestelde verwachtingen in de cases. Toekomstig onderzoek moet zich onder meer ook hier op richten. Als mogelijke concrete verwachtingen houden wij hier momenteel evenwel vast aan de suggesties die in de Delphistudie naar voren kwamen. Deze worden immers ook in de gevalstudie vernoemd. We vullen deze indicator aan met verwachtingen die in de gevalstudies door schoolleiders in de verf werden gezet. Een zelfevaluatie is gegeven deze indicator kwaliteitsvol als deze:

  - een goed beeld geeft van het functioneren van de school;
  - een goed beeld geeft van wat er leeft in de school;
  - concrete problemen oplost;
  - teamleden stimuleert om verantwoordelijkheid op te nemen;
  - stimuleert tot het bespreken van de eigen opvattingen over goed onderwijs;
  - de wil om kritischer tegen de zaken aan te kijken bevordert;
  - in extra motivatie resulteert om kwaliteit te leveren;
  - de betrokkenheid van teamleden bij het schoolbeleid versterkt;
  - bijdraagt tot meer samenwerking in de school;
  - ertoe bijdraagt dat men in de school in sterkere mate dezelfde doelen gaat nastreven;

- de samenhang in de school versterkt.

Het doel van deze indicator is te expliciteren tot welke resultaten de zelfevaluatie wel en niet leidde. De verwachting is dus niet dat een zelfevaluatie bijdraagt tot al deze resultaten.

Voor schoolleiders kunnen we dus één abstracte en twee concrete indicatoren ('resulteren in actiepunten' en 'resulteren in concrete resultaten') voor de kwaliteit van de zelfevaluaties in hun school hanteren.

3. *Expliciteer de doelen van de zelfevaluatie*. Wanneer men uit is op het evalueren van de kwaliteit van zelfevaluaties dan doet men er goed aan om scholen vooraf te laten expliciteren welke doelen zij willen bereiken. Deze doelen kunnen dan het uitgangspunt vormen van een latere kwaliteitsbeoordeling.
4. *Betrek de context in de meta-evaluatie*. Om in kaart te brengen in welke mate zelfevaluaties de vooropgestelde verwachtingen realiseren, doet men er goed aan de zelfevaluatie op een holistische en gecontextualiseerde manier te benaderen. Er zijn verscheidene elementen die maken dat een bepaald zelfevaluatieresultaat in verschillende contexten tot verschillende beoordelingen zou moeten leiden. Daarbij denken we aan de vooropgestelde betrokkenheid van teamleden, de ingezette middelen, de eerdere ervaringen met reflectie en zelfevaluatie, en de aard van de resultaten van de zelfevaluatie. Om dergelijke verschillen in kaart te brengen is het aangewezen het gesprek met schoolleiders en teamleden aan te gaan. Inzicht in hun beleving en betekenisgeving omtrent het zelfevaluatieproces en de zelfevaluatieresultaten is noodzakelijk om uitspraken te doen over de kwaliteit van zelfevaluaties.

6.1.2 Hoe kwaliteitsvol zijn zelfevaluaties in de bestudeerde scholen?

De onderzochte periode in de gevalstudies bedroeg anderhalf jaar, te beginnen nadat in de scholen een beslissing werd genomen om een zelfevaluatie uit te voeren. Binnen die periode bleek het niet mogelijk het volledige verloop van de zelfevaluaties te bestuderen. Dat heeft belangrijke implicaties voor onze conclusies omtrent de kwaliteit van de bestudeerde zelfevaluaties. Het proces liep immers nog en de inschatting van de kwaliteit is dus tussentijds. Ondanks deze beperking proberen we in lijn met de vooropgestelde raadgevingen toch tot een aantal conclusies te komen. Daarbij wordt een onderscheid gemaakt tussen de kwaliteit door de ogen van teamleden en de kwaliteit door de ogen van schoolleiders. Daaruit zal blijken dat onze verwachtingen over verschillen in de perceptie van schoolleiders en teamleden omtrent de kwaliteit van zelfevaluaties bevestigd werden (cf. hypothese 5c).

a. KWALITEIT DOOR DE OGEN VAN TEAMLEDEN

We stelden hoger dat zelfevaluaties volgens de bevraagde teamleden moeten resulteren in gepaste actiepunten. De teamleden geven aan geen concrete aanwijzingen te hebben dat de zelfevaluatie in hun school tot actiepunten geleid heeft. Na de presentatie van de

zelfevaluatieresultaten werd het voor hen reeds maandenlang windstil omtrent de zelfevaluatie. We mogen hier echter niet voorbijgaan aan de ruimere context van het zelfevaluatiegebeuren. Na de laatste interviewronde kunnen er in de bestudeerde scholen immers nog activiteiten plaats hebben gevonden die tot een andere beoordeling leidden bij de teamleden. Het beeld dat wij hier schetsen is echter wel het beeld op meer dan één jaar nadat de teamleden de vragenlijst invulden. Wij concluderen dus dat de bevraagde teamleden zich in de laatste interviewronde behoorlijk negatief uitspreken over de resultaten van de zelfevaluatie. (Zelfs) hun beperkte verwachtingen werden onvoldoende gerealiseerd om in hun ogen van kwaliteitsvolle zelfevaluaties te kunnen spreken. Voor de teamleden is het oordeel over de kwaliteit van de zelfevaluatie negatiever dan verwacht (cf. hypothese 5a).

b. KWALITEIT DOOR DE OGEN VAN SCHOOLLEIDERS

We concluderen dat de bevraagde schoolleiders de zelfevaluatie die in hun school plaatsvond - zoals verwacht (cf. hypothese 5a) - wel als eerder kwaliteitsvol bestempelen. We rapporteerden evidentie voor een instrumenteel, conceptueel, symbolisch en inspraakgericht gebruik van de zelfevaluatieresultaten. Dat doet vermoeden dat de zelfevaluaties - minstens ten dele - tegemoet zijn gekomen aan de verwachtingen die vooropgesteld werden. Of dit het geval was gaan we na aan de hand van de drie indicatoren die we voor de kwaliteitsbeoordeling door de ogen van schoolleiders vooropstelden. Onze verwachting was dat het oordeel over de verschillende kwaliteitsindicatoren positief zou samenhangen (zie hypothese 5b). De volgende resultaten spreken dit alvast niet tegen.

Een abstracte indicator: Afstappen van de concrete verwachtingen. Wanneer de bevraagde schoolleiders in algemene termen spreken over het zelfevaluatieproces, blijken zij het een waardevolle onderneming te hebben gevonden. Zij geven aan dat de zelfevaluatie als dusdanig hun verwachtingen heeft ingelost. Louter uitgaande van deze indicator blijken de zelfevaluaties dus van een goede kwaliteit. We mogen daarbij niet vergeten dat achter dit oordeel complexe redeneringen kunnen schuilgaan. Ondanks de tekorten, de ongelukkige samenloop van omstandigheden, de zaken die men volgende keer anders zou aanpakken, zijn de bevraagde schoolleiders tevreden over de zelfevaluatie.

Een concrete indicator: resulteren in gepaste actiepunten. De bevraagde schoolleiders geven aan zowel te komen tot actiepunten voor zichzelf als tot actiepunten voor de school als geheel. We stelden bijvoorbeeld vast dat ze aangeven hun eigen leiderschapsstijl bij te stellen of dat ze nieuwe structuren voor een goede aansturing van het schoolbeleid op poten zetten. Ook voor leerkrachten volgen er in de ogen van schoolleiders duidelijke actiepunten uit de uitgevoerde zelfevaluaties. Schoolleiders zijn dus van mening dat de zelfevaluatie in die zin kwaliteitsvol is verlopen. De verwachtingen zijn in hun ogen wat dat betreft ingelost. Daarbij geven ze ook aan dat er meer mogelijk is. Mits nog meer aan een aantal door henzelf vooropgestelde aandachtspunten zou tegemoet worden gekomen (zie de onderzoeksvraag in paragraaf 6.1.3), zouden zelfevaluaties in nog meer gepaste actiepunten kunnen resulteren.

Een concrete indicator: 'resulteren in concrete resultaten'. Om uitspraken te doen over de derde indicator zouden we schoolleiders moeten uitnodigen om zich uit te spreken over het al dan niet gerealiseerd zijn van de concrete resultaten in kwestie. Die oefening kon – gezien de timing van zelfevaluatie zelfevaluaties - tijdens de eigenlijke gevalstudies nog niet systematisch uitgevoerd worden. Er is desalniettemin wel evidentie omtrent een aantal van de vooropgestelde resultaten. Die evidentie leidt echter niet tot eenduidige conclusies. De vraag is bijvoorbeeld of de zelfevaluatie volgens schoolleiders geleid heeft tot een goed beeld van het functioneren van de school. We rapporteerden indicaties afkomstig van schoolleiders omtrent mogelijke vertekeningen van het representatieve beeld door foutieve groepsindelingen, door een ongelukkige timing van de afname en door het overheersen van meningen met een luide stem in de verhelderingsgesprekken. De vraag is welke impact zulke vaststellingen hebben op de beoordeling door schoolleiders. Daarover werden in de interviews geen gegevens verzameld. Een gelijkaardige vaststelling geldt voor de andere concrete resultaten. Deze indicator zal dan ook voornamelijk in toekomstige meta-evaluaties aangewend moeten worden. Om op dit moment toch reeds een indicatie te geven van het oordeel van de bevraagde schoolleiders over de concrete resultaten, stelden we hen schriftelijk de vraag of de zelfevaluatie die in hun school plaatsvond volgens henzelf al dan niet tot de genoemde concrete resultaten leidde. Dit leidde tot de resultaten in Tabel 86[34].

Tabel 86: Resulteren in concrete resultaten door de ogen van schoolleiders

Ik vind dat de zelfevaluatie die in onze school plaatsvond:	Niet	Eerder niet	Eerder wel	Wel
...een goed beeld gaf van het functioneren van onze school			TIV/SOI	
...een goed beeld gaf van wat er leeft in onze school			SOI	TIV
...concrete problemen oploste		TIV/SOI		
...teamleden stimuleerde om verantwoordelijkheid op te nemen		TIV/SOI		
... het bespreken van de eigen opvattingen over goed onderwijs stimuleerde			TIV/SOI	
...de wil om kritischer tegen de zaken aan te kijken bevorderde			TIV/SOI	
...in extra motivatie resulteerde om kwaliteit te leveren		SOI	TIV	
...de betrokkenheid van teamleden bij het schoolbeleid versterkte		TIV	SOI	
...bijdroeg tot meer samenwerking in de school		TIV/SOI		
...bijdroeg tot het in sterkere mate nastreven van dezelfde doelen		TIV	SOI	
...de samenhang in de school versterkte			TIV/SOI	

6.1.3 Hoe kunnen de verschillen in de resultaten van zelfevaluaties verklaard worden?

In het verklaren van verschillen in de uitkomsten van zelfevaluaties verwijzen respondenten vaak naar aspecten die betrekking hebben op het verloop van de zelfevaluatie. De

[34] Het gaat hier om de perceptie van de pedagogisch directeurs van het Technisch Instituut Voldegem (TIV) en van het Sint-Odrada Instituut (SOI).

conclusies hieromtrent structureren we op basis van drie relevante thema's. Het belang van voldoen aan de 'dragers van zelfevaluatie als beleidsdaad', het belang van voldoen aan de dragers van 'zelfevaluatie als onderzoeksdaad' en het belang van ondersteuning. Binnen deze thema's zullen knelpunten en succesfactoren naar voren geschoven worden om een volledig beeld te krijgen van de elementen die een rol gespeeld hebben in de vormgeving en de resultaten van de zelfevaluaties in de bestudeerde scholen. Het gaat ons hierbij om het inventariseren van sterktes en zwaktes zonder te willen generaliseren wat hun frequentie van voorkomen betreft. Deze aanpak impliceert dat scholen die hoger scoren op de zelfevaluatiegebonden kenmerken ook volgens de respondenten tot kwaliteitsvollere zelfevaluatie zullen komen. Dat is conform onze 6de onderzoekshypothese. Doorheen het beschrijven van de conclusies zal verder blijken dat ook de gevalstudies onze onderzoekshypothese omtrent het realiseren van de dragers van 'zelfevaluatie als beleids- en onderzoeksdaad' in sterke mate ondersteunen. De bestudeerde scholen scoren inderdaad gematigd positief of eerder negatief op de verschillende dragers, maar het concrete plaatje varieert over de verschillende cases. Ook de perceptie van verschillende respondenten binnen de school blijkt te variëren (cf. hypothese 3a).

a. HET BELANG VAN VOLDOEN AAN DE DRAGERS VAN 'ZELFEVALUATIE ALS BELEIDSDAAD'

Voor alle dragers van 'zelfevaluatie als beleidsdaad' inventariseren we in wat volgt succesfactoren en knelpunten. De selectie van inhouden is ingegeven door de link met de resultaten van de zelfevaluatie. Succesfactoren zijn dan de elementen die volgens respondenten bijgedragen hebben tot de kwaliteit van de zelfevaluatie; knelpunten omvatten aspecten die – wanneer er aan tegemoet zou worden gekomen - tot kwaliteitsvollere zelfevaluatie zouden leiden. De dragers staan niet los van elkaar. We gebruiken het onderscheid om structuur aan te brengen in het holistische gebeuren dat het uitvoeren van zelfevaluaties is.

- *Bevindingen inzake reflectief vermogen*

De meest positieve vaststelling omtrent reflectief vermogen is dat de bevraagde schoolleiders overtuigd zijn van de leerkansen die reflectie en zelfevaluatie met zich meebrengen. De doelen die zij vooropstellen bij het initiëren van een zelfevaluatie blijken in hoofdzaak intrinsiek en schoolontwikkelingsgericht. De positieve houding ten aanzien van zelfevaluatie gaat evenwel niet op voor de teamleden in hun school. Daarmee worden onze hypotheses hieromtrent bevestigd (cf. hypothese 2). Schoolleiders nemen vaak zichzelf evenwel als referentiepunt en gaan er daardoor nogal eens foutief vanuit dat ook teamleden zelfevaluatie als een vanzelfsprekend iets ervaren. Zelfevaluaties worden door de bevraagde teamleden nog als een vreemd iets ervaren. Sommige teamleden geven aan dat het gebeuren 'bedreigend' overkomt. Ze reageren vervolgens met vlucht- of vechtgedrag (Swieringa & Wierdsma, 1990). We zullen verder concluderen dat dit deels toe te schrijven aan de vaststelling dat de bestudeerde zelfevaluaties slechts in beperkte mate tegemoet komen aan dragers zoals gedeeld leiderschap, gezamenlijke doelgerichtheid en communicatie.

Er is in de bestudeerde scholen een sterke handelingsgerichtheid. Respondenten geven zelf aan dat er in hun school te weinig tijd wordt gespendeerd aan reflectie. Tijdsgebrek wordt hierbij vaak als verontschuldiging ingeroepen. We stelden echter vast dat ook de houding ten aanzien van reflectie en gewoontevorming reflectie bemoeilijken. Ook de vaststelling dat eerdere zelfevaluaties niet tot voelbare veranderingen geleid hebben, tempert het enthousiasme van de teamleden. We concluderen dat van een cultuur van reflectie die zou bijdragen tot een positieve houding ten aanzien van zelfevaluatie bij de bevraagde teamleden weinig of geen sprake is (cf. hypothese 2).

Reflectief vermogen is zowel van belang op schoolniveau als op individueel respondentniveau (Vanhoof & Van Petegem, 2006b). Beide niveaus dienen verder te worden ontwikkeld om tot kwaliteitsvolle zelfevaluatie te komen. Dat blijkt een moeilijke opdracht omdat de vicieuze cirkel – met name de negatieve houding die leidt tot minder geslaagde zelfevaluatie die op haar beurt leidt tot een negatieve houding – moet worden doorbroken. Er is echter één en ander reeds in beweging gezet volgens de respondenten. De benoemde knelpunten zouden immers langzaamaan aan belang verliezen. Respondenten geven bovendien aan zelf wel open te staan voor reflectie en zelfevaluatie. Het zijn voornamelijk de anderen die er problemen mee hebben. Dat kan er op wijzen dat zelfevaluatie op individueel niveau niet echt een (groot) probleem (meer) is. Vormen van reflectie – zoals de hier bestudeerde zelfevaluaties - die een grotere openheid en kwetsbaarheid vergen, liggen volgens de respondenten evenwel vaak nog steeds moeilijk (Schildkamp, 2007).

- *Bevindingen inzake gezamenlijke doelgerichtheid*

Het initiatief om een zelfevaluatie uit te voeren kwam in de gevalstudies van de school zelf. Er zit meer achter dan de wil om aan externe verwachtingen tegemoet te komen. Het initiatief om zelfevaluaties uit te voeren komt telkens van de schoolleiding of zelfs enkel van een deel van de schoolleiding. De vraag naar schoolzelfevaluatie zou momenteel nog niet groeien vanuit de teamleden. Schoolleiders stellen verscheidene doelen voorop bij het starten van een zelfevaluatie, maar deze doelen blijven echter vaak impliciet. Men wil aan de hand van informatie over de eigen situatie zicht krijgen op het eigen functioneren en dat van collega's. Die inzichten heeft men immers nodig om er vervolgens de beleidsvoering op te enten.

Sommige initiatiefnemers geven aan zich bewust te zijn van het belang om anderen voor de zelfevaluatie warm te maken. De beperkte pogingen die daartoe ondernomen worden (zoals bespreking op de personeelsvergadering en aankondiging in bestaande communicatiekanalen) blijken echter niet effectief. In de bestudeerde scholen hebben respondenten die verder van de zelfevaluatie afstaan geen zicht op wat de aanleiding is om in hun school een zelfevaluatie uit te voeren. Dat heeft verschillende mogelijke oorzaken. Zo gaan sommige schoolleiders er van uit dat hun teamleden het vanzelfsprekend vinden dat er aan zelfevaluatie gedaan wordt. Deze schoolleiders vinden het daardoor overbodig aan de beoogde doelen aandacht te besteden. Uit de cases blijkt dat deze veronderstelling maar ten dele opgaat. Wanneer teamleden niet betrokken zijn bij de voorliggende zelfe-

valuatiethema's en het gevoel hebben er geen sterke inhoudelijke bijdrage toe te kunnen leveren, dan blijken zij inderdaad niet zwaar te tillen aan een beperkte gezamenlijke doelgerichtheid. Wanneer ze zich wel betrokken en onderlegd voelen, dan ervaren zij de opgelegde aanpak wel als een probleem. Als men van een zelfevaluatie verwacht dat deze resulteert in door teamleden gedragen actiepunten, dan streeft men volgens de respondenten dus best bij aanvang reeds naar gezamenlijke doelgerichtheid. Anders dreigt de zelfevaluatie vast te lopen in de valkuil die we vaststelden; ondanks de intenties en pogingen om de teamleden te betrekken bij het vooropstellen van doelen voor toekomstig beleid, blijken de actiepunten die uit de zelfevaluatie voortvloeien ervaren te worden als van bovenaf opgelegd.

- *Bevindingen inzake doeltreffende communicatie*

De bevraagde leerkrachten zijn doorgaans zeer beperkt of in het geheel niet op de hoogte van de doelen, het verloop en de resultaten van de zelfevaluatie. Dit ondanks de bestaande pogingen om hierover te communiceren. Het gaat dan om toelichtingen tijdens personeelsvergaderingen, het verspreiden van integrale rapporten, of de presentatie van de resultaten door een ondersteuner. Deze initiatieven blijken aan de communicatiedoelen voorbij te gaan en worden bovendien niet gesmaakt door de teamleden. Het grote aantal bemerkingen omtrent de communicatie in het kader van de zelfevaluatie geeft aan dat respondenten dit een cruciaal aspect vinden.

Het beperkte effect van de communicatiestrategieën heeft verschillende oorzaken. Schoolleiders (en sommige teamleden) geven aan dat de verantwoordelijkheid voor de manke communicatie ook bij de teamleden zelf ligt. Hun bereidheid om zich te informeren zou beperkt zijn en hun inspanningen niet in verhouding tot hun verwachtingen. Toch speelt ook de aanpak van de communicatie een rol. We stelden vast dat communicatie zich beperkt tot informeren over en niet komt tot interactieve besprekingen en discussies. Dit is grotendeels toe te schrijven aan de inschatting van de bevraagde schoolleiders dat meer diepgaande communicatie overbodig is. Zij gaan er nogal eens van uit dat de eigen teamleden geen fundamentele bedenkingen en vragen hebben bij het uitvoeren van een zelfevaluatie. Die veronderstelling gaat echter niet volledig op, zelfs niet in scholen die er reeds eerdere zelfevaluaties hebben opzitten. Na de zelfevaluatie geven schoolleiders aan dat het wenselijk is in de toekomst meer en anders te communiceren over heel het zelfevaluatiegebeuren. Onze vaststelling is verder dat de communicatie best ook vorm wordt gegeven vanuit het standpunt van de teamleden. Het moet gaan om meer dan informeren om een vlot praktisch verloop van de zelfevaluatie te verzekeren. Ook het inschakelen van teamleden in het communicatieproces zou kunnen overwogen worden. Dat geldt bijvoorbeeld bij het communiceren van de resultaten en de actiepunten waartoe de zelfevaluatie bijdroeg. Wat door teamleden aangebracht wordt, wordt door teamleden beter opgepikt (cf. de lessen uit de theorie omtrent kennisconversie) en anders onthaald dan wat door de schoolleiding gepresenteerd wordt.

Communicatie is verder ook van belang omdat het een cruciale voorwaarde is om aan andere dragers tegemoet te komen. Zo is het onmogelijk om bij teamleden een gevoel

van gezamenlijke doelgerichtheid of participatief leiderschap te creëren indien de initiatieven daartoe niet gepaard gaan met een passende communicatie.

- *Bevindingen inzake ondersteunende relaties en samenwerking*

De gevalstudies hebben aangetoond dat er in de school sprake moet zijn van een voldoende 'veilige' context om tot kwaliteitsvolle zelfevaluatie te komen. De bevraagde teamleden vinden het immers niet vanzelfsprekend om zich open en kwetsbaar op te stellen tegenover collega's en de schoolleiding. Momenteel is er terughoudendheid om een open dialoog aan te gaan. Tenminste wanneer dit dient te gebeuren in relatie tot personen die men zelf niet verkiest. De teamleden geven allen aan zich ondersteund te voelen door personen die door henzelf gekozen werden. De bestudeerde zelfevaluaties verruimen deze relaties echter naar personen waarvoor dit niet opgaat. Dat uit zich bijvoorbeeld in een defensieve houding tijdens verhelderingsgesprekken. Mede daardoor ervaart men deze vormen van zelfevaluatie als een opgelegde activiteit. We concludeerden dat het in de bestudeerde scholen ontbreekt aan vertouwen en collegialiteit om de mogelijkheden van zelfevaluatie tenvolle te benutten.

We stelden ook vast dat de geïnterviewde schoolleiders naar eigen zeggen niet volledig beschikken over de kennis en vaardigheden die het zelfstandig uitvoeren van een zelfevaluatie met zich meebrengt. Om daaraan tegemoet te komen werden in de bestudeerde scholen interne en externe vormen van ondersteuning opgezet. De externe ondersteuning richt zich voornamelijk op de coördinatoren van de zelfevaluatie. De interne ondersteuning richt zich eerder op de andere teamleden. Zij hadden voornamelijk ondersteuning nodig bij het invullen van het zelfevaluatie-instrument. Op dergelijke ondersteuningsbehoeften weten scholen op een passende manier in te spelen. De rol van teamleden was in de cases echter beperkt, zodat interne ondersteuning voor hen niet sterk aan de orde was. Interne ondersteuning van de coördinatoren werd door middel van informele contacten en het gebruik van een stuurgroep gerealiseerd. Doordat de bevraagde scholen een beroep hebben gedaan op externe ondersteuning werd de kennis en vaardigheden van coördinatoren evenwel niet sterk op de proef gesteld. De rol van de stuurgroep was tijdens het zelfevaluatieproces mede daarom eerder beperkt. Wanneer zelfevaluaties minder door externe begeleiders aangestuurd zouden worden, zouden de stuurgroepen wellicht aan belang winnen. De bevraagde schoolleiders voelden zich voldoende ondersteund doorheen de stappen die deel uitmaakten van het begeleidingstraject. Bij stappen die daarbuiten vallen – zoals het komen tot actiepunten – is dat niet meer het geval.

Verder blijkt ook het tekort aan beschikbare tijd een hinderpaal om een vlot verloop van zelfevaluaties te waarborgen. Ondanks de goede intenties bij de start van de zelfevaluatie geven de schoolleiders aan over te weinig tijd te beschikken om het zelfevaluatieproces optimaal te kunnen opvolgen. Het gebrek aan tijd leidt tot een gebrek aan prioriteit. Het zelfevaluatieproces lag tijdens cruciale fasen te zeer in de handen van één verantwoordelijke. Het delen van deze verantwoordelijkheden en het volgen van een vooraf opgestelde stappenplannen zou het proces beter op de agenda kunnen houden. Een bijkomend probleem om een draagvlak te creëren voor het uitvoeren van zelfevaluaties is

volgens een aantal respondenten een gebrek aan flexibiliteit in de tijdsbesteding van teamleden. Cruciale aspecten van het zelfevaluatie moesten in de cases noodgedwongen georganiseerd worden op momenten die betrokkenheid, openheid en diepgang in de weg staan.

- *Bevindingen inzake gedeeld leiderschap*

Schoolleiders gaan er nogal eens van uit dat het uitvoeren van een zelfevaluatie voor hun teamleden een vanzelfsprekende activiteit is. Zij nemen hun eigen houding daarbij als referentie. Om een draagvlak te creëren is een sterk participatieve aanpak gegeven deze veronderstelling niet nodig. De zelfevaluatie is volgens schoolleiders op zich reeds een vorm van inspraak. De initiatieven om teamleden in het zelfevaluatieproces te betrekken werden dan ook eerder ingegeven door de impulsen van externe ondersteuners. Het uitgangspunt van de ondersteuners is immers – zoals ook uit deze gevalstudies blijkt overigens - dat de genoemde veronderstelling van schoolleiders niet correct is. De mate waarin teamleden betrokken worden en de manier waarop dat gebeurt, verschilt over de bestudeerde scholen. Nergens werden teamleden echter betrokken bij de beslissing om een zelfevaluatie uit te voeren en bij de keuze van het zelfevaluatie-instrument. Daarmee lag ook het vastleggen van inhouden voornamelijk in handen van schoolleiders.

Om de inspraak van verschillende geledingen in het zelfevaluatieproces vorm te geven doet men nogal eens een beroep op een stuurgroep, al dan niet speciaal met het oog op de zelfevaluatie opgericht. Deze structuur heeft volgens de respondenten veel mogelijkheden die nog te weinig gebruikt worden. De rol van de stuurgroepen is in de cases immers beperkt gebleven. Voornamelijk praktische aspecten werden behandeld, maar ook deze waren gezien de externe ondersteuning beperkt. Tijdens de interviewperiode was er geen noemenswaardige betrokkenheid bij het interpreteren van de zelfevaluatieresultaten of bij het vooropstellen van actiepunten. Dat is zowel toe te schrijven aan het feit dat de schoolleiding de stuurgroepen niet samenriep als aan de afwachtende houding van de leden van de stuurgroepen. We stelden immers vast dat het feit dat men van zulk een stuurgroep deel uitmaakt, niet betekent dat men zich verantwoordelijk voelt voor het welslagen van de zelfevaluatie.

De betrokkenheid van teamleden die geen deel uitmaakten van de stuurgroep beperkte zich in de bestudeerde scholen tot het invullen van de vragenlijst. Tijdens en na de zelfevaluatie ervaren de bevraagde schoolleiders dit als een probleem. Zij geven aan dat ze aan de leerkrachten een rol moeten toekennen als men tot breed gedragen actiepunten wil komen. Teamleden zelf lijken in dit opzicht uiteen te vallen in twee groepen. Sommige teamleden willen in de besluitvorming betrokken worden, anderen zien er geen graten in dat het schoolbeleid door de schoolleiding (of door anderen) uitgestippeld wordt. Deze verschillen tussen teamleden maken het moeilijk om een passende strategie uit te tekenen. De gevalstudies bevestigen de relevantie van de criteria 'betrokkenheid' en 'expertise' om hierbij keuzes te maken (Hoy & Miskel, 2001). Dit impliceert evenwel dat men niet van alle teamleden inspraak moet verwachten. Teamleden moeten echter ook geres-

ponsabiliseerd worden volgens schoolleiders. Zelfevaluaties mogen zich volgens hen niet beperken tot het geven van commentaar van op de zijlijn.

De gevalstudies impliceren verder ook dat gedeeld leiderschap een structureel kenmerk moet zijn om teamleden te kunnen overtuigen. Wanneer de schoolleiding eerst autonoom beslist dat er een zelfevaluatie moet komen, welke thema's daarbij bevraagd zullen worden en welk instrument men gaat gebruiken wordt het blijkbaar zeer moeilijk om teamleden in latere fasen van de zelfevaluatie als nog tot betrokkenheid en inspraak te brengen.

- *Bevindingen inzake geïntegreerd beleid*

De bestudeerde zelfevaluaties dragen elk indicaties van een geïntegreerde aanpak in zich. Dat kan gaan van het werken in de schoot van een bestaande werkgroep, het koppelen van bevindingen van de zelfevaluatie aan deze van de doorlichting of het inschrijven van de actiepunten in het schoolwerkplan. Zelfevaluaties maken in de bestudeerde scholen echter nog geen systematisch deel uit kwaliteitszorgbeleid. Het gebeuren wordt door teamleden nog sterk als een eenmalige activiteit ervaren. Schoolleiders zien de link van de zelfevaluatie met andere activiteiten in het kader van de beleidsvoering veel duidelijker. Toch geven ook zij aan dat onderlinge verbanden nog beter afgelijnd en geëxpliciteerd zouden moeten worden. De zelfevaluaties kunnen volgens de respondenten inhoudelijk nog beter aansluiten bij de beleidsprioriteiten van de school, het trekken van conclusies uit de zelfevaluatieresultaten zou kunnen gebeuren door werkgroepen die dichter bij teamleden staan, en de link van actiepunten met de pedagogisch-onderwijskundige visie zou meer geëxpliciteerd kunnen worden. Schoolleiders zouden er verder baat bij kunnen hebben om meer te denken vanuit theoretische raamwerken. Ook een betere afstemming van verschillende vormen van gegevensverzameling en reflectie zou de kwaliteit van zelfevaluaties nog ten goede kunnen komen.

- *Bevindingen inzake responsief vermogen*

Of responsief vermogen de kwaliteit van zelfevaluatie ten goede komt, kunnen we op basis van de gevalstudies niet hard maken. Er is immers op dit moment blijkbaar nog maar weinig sprake van responsief vermogen in de bestudeerde zelfevaluaties. Er werden – op de ondersteuners na – geen initiatieven genomen om externen in de zelfevaluatie te betrekken. De bevraagde schoolleiders geven nochtans aan dat het betrekken van andere stakeholders (leerlingen, ouders, stageplaatsen, scholen en bedrijven waar leerlingen naar doorstromen) potentiële leerkansen in zich draagt. Momenteel verkozen de bestudeerde scholen in de context van kwaliteitszorg nog binnen de 'veilige' schoolmuren te werken. Zich kwetsbaar opstellen naar de buitenwereld is volgens de respondenten momenteel nog niet vanzelfsprekend. Het hoeft dan ook niet te verwonderen dat scholen momenteel het liefst werken aan een zelfevaluatie zonder de betrokkenheid van externen. De zelfevaluatie zou dan in het teken komen te staan van verantwoording in plaats van ontwikkeling. De stap naar responsieve zelfevaluaties – of bij uitbreiding collegiale visitatie – wordt in de bestudeerde scholen dus wel naar waarde geschat maar men kampt duidelijk nog met koudwatervrees.

➢ *Bevindingen inzake innovatief vermogen*

Het uitvoeren van zelfevaluaties bleek op zich reeds een vorm van vernieuwing in de bestudeerde scholen. We schreven daarom dat het 'innoverend vermogen' van scholen in deze context eerder een invalshoek is dan een drager die bijkomende informatie opleverde om verschillen in de kwaliteit van zelfevaluatie te voorspellen. Toch gaat het om een zeer belangrijke vaststelling. De verwachtingen die aan zelfevaluatie toegeschreven worden zijn immers maar realistisch wanneer het uitvoeren van zelfevaluaties een gangbare praktijk zal zijn in Vlaamse scholen. In de bestudeerde cases is dat nog niet het geval.

b. HET BELANG VAN VOLDOEN AAN DE DRAGERS VAN 'ZELFEVALUATIE ALS ONDERZOEKSDAAD'

Ook in het verloop van de 'zelfevaluatie als onderzoeksdaad' blijken we elementen te kunnen onderkennen die samenhangen met de kwaliteit van zelfevaluaties. Kwaliteitsvolle zelfevaluaties berusten met name op een doordachte selectie van zelfevaluatieinhouden, een systematische gegevensverzameling en een correcte interpretatie van de resultaten (cf. hypothese 6). In de volgende paragrafen komen we tot conclusies omtrent de succesfactoren en knelpunten inzake deze aspecten van 'zelfevaluatie als onderzoeksdaad'.

➢ *Bevindingen inzake inhoud en instrumentkeuze*

De bestudeerde scholen hadden weinig inbreng in het bepalen van de inhoud van de zelfevaluaties. De inhoudelijke selectie ging immers telkens samen met de instrumentkeuze. Men evalueert dus eerder wat er in instrumenten beschikbaar is, dan datgene waaromtrent men informatie behoeft. Volgens de bevraagde schoolleiders is er echter een grote overlap tussen beide zodat dit niet als een probleem ervaren wordt. Toch geeft men aan in de toekomst meer bewust en meer pro-actief tewerk te willen gaan. Ook een meer participatieve aanpak wordt daarbij bepleit.

Het afwegen van verschillende instrumenten voor zelfevaluatie is blijkbaar geen gewoonte alvorens scholen een keuze maken. Het instrument moet volgens de respondenten desalniettemin aan een aantal criteria voldoen. De motieven die de bestudeerde scholen tot een specifieke keuze brachten, zijn vroegere ervaringen met het instrument, de ervaringen van andere scholen, de kostprijs en de verwachting dat de keuze voor een instrument gepaard gaat met een vorm van ondersteuning. Ook de mogelijkheid tot een vergelijking met andere scholen spreekt sommige respondenten aan. Eenmaal vertrouwd met het gebruik van zelfevaluatie-instrumenten geeft men aan dat ook de afstemming op de eigen schoolcontext in de toekomst een bekommernis zal zijn. Zelf instrumenten ontwikkelen vindt men evenwel een te omslachtig werk. Het heeft bovendien ook nadelen. Een eigen instrument wordt als minder objectief beschouwd en zou naar de teamleden toe minder legitiem overkomen. Samen met de ondersteuners een aantal aanpassingen doorvoeren aan bestaande instrumenten is dan een werkbare middenweg. Men kan op die manier de voordelen van bestaande en zelfontwikkelde instrumenten combineren.

- *Bevindingen inzake de gegevensverzameling*

In de bestudeerde zelfevaluaties was er sprake van twee vormen van gegevensverzameling: een schriftelijke vragenlijst en verhelderingsgesprekken. De keuze voor schriftelijke vragenlijsten is in belangrijke mate ingegeven door overwegingen inzake efficiëntie. Vragenlijsten maken het volgens de respondenten mogelijk om op een snelle manier een betrouwbaar en representatief beeld van de school te bekomen. De respondenten hebben bovendien een groot vertrouwen in de kwaliteit van de ter beschikking gestelde instrumenten. Het gebruik van vragenlijsten gebeurde in de bestudeerde scholen in wezen zonder ernstige problemen. Verhelderingsgesprekken zijn een initiatief van de ondersteuners van de DISO-zelfevaluatie. De gesprekken zorgden volgens de respondenten voor nuances en diepgang, en worden dan ook als een meerwaarde beschouwd. Verhelderingsgesprekken maken teamleden en schoolleiders bewust van het bestaan van verschillende zienswijzen in hun school. Herhaaldelijk wordt echter benadrukt dat een goede begeleiding van zulke gesprekken een belangrijke voorwaarde is. Ook de relevantie van de groepsindeling dient te worden bewaakt. Zoniet kunnen verhelderingsgesprekken een vertekend beeld geven.

- *Bevindingen inzake de interpretatie van de gegevens*

Het interpretatieproces was in de bestudeerde scholen een activiteit die ofwel in handen lag van de schoolleiding ofwel in overleg binnen de zelfevaluatiestuurgroep plaats vond. Teamleden speelden hier geen noemenswaardige rol. Alle respondenten geven aan dat het interpretatieproces best snel volgt op de fase van gegevensverzameling. We hebben echter vastgesteld dat het een activiteit is die nogal eens op de lange baan geschoven wordt. Daardoor liepen de bestudeerde zelfevaluatieprocessen een maandenlange vertraging op. De betrokkenheid op het gebeuren daalde daardoor sterk. Kimball (2002) suggereert dat de terugkoppeling best zo snel mogelijk gebeurt. De ervaringen in de bestudeerde scholen suggereren dat er verschillende redenen zijn voor deze vertragingen. We stelden bijvoorbeeld vast dat schoolleiders niet goed weten welke conclusies ze moeten trekken wanneer er zeer verschillende bevindingen of in hun ogen onterechte bevindingen in het zelfevaluatierapport opduiken. Wanneer schoolleiders zich onrechtvaardig hard aangepakt voelen door de zelfevaluatieresultaten of zich er niet goed in kunnen vinden, worden de resultaten minder prioritair aangepakt. Deze bevindingen bevestigen het belang van de gepercipieerde accuraatheid en het aansluiten bij de behoeftes van de gebruikers (Teddlie, Kochan & Taylor, 2002; Visscher, 2002). Verder blijkt ook het zelfevaluatierapport als dusdanig vaak niet bij te dragen tot een vlotte interpretatie. Verschillende respondenten geven aan dat ze verwacht hadden dat het rapport een stap verder zou gaan dan het beschrijven van de vaststellingen. Die teleurstelling en het ontbreken van kant-en-klare aanknopingspunten voor de interpretatie in de eigen schoolcontext maakt dat men moeilijkheden ervaart om verdere stappen te zetten (Coe & Visscher, 2002). Een andere vaststelling is dat men een prioriteit moet kunnen en willen maken van de bevindingen. De onontkoombaarheid van het dagelijkse werk maakt dit moeilijk volgens schoolleiders. Ook het opduiken van meer dwingende beleidsactiviteiten, zoals het opvolgen van de bevindingen van een doorlichting, kan voor vertraging zorgen. Een laats-

te reden voor het moeizame verloop van het interpretatieproces is de vaststelling dat de ondersteuning ophoudt bij het opleveren en bespreken van het zelfevaluatierapport.

c. HET BELANG VAN ONDERSTEUNING

De scholen die in deze gevalstudies werden opgenomen, deden alle drie een beroep op externe ondersteuning. Over de vraag of zelfevaluaties kwaliteitsvoller verlopen in scholen met ondersteuning dan in scholen zonder ondersteuning kunnen we dus in strikte zin geen uitspraken doen (cf. hypothese 7). We kunnen wel in kaart brengen op welke manier de ondersteuning volgens de respondenten de kwaliteit van de zelfevaluatie in hun school ten goede is gekomen en of de respondenten aanknopingspunten zien om het rendement van ondersteuning nog te verhogen.

De bestudeerde scholen kiezen bewust voor ondersteuning. Zij die geen ervaringen hebben met zelfevaluatie in het algemeen en het werken met schriftelijk vragenlijsten in het bijzonder geven aan niet goed te weten wat ze moet verwachten. Er is zowel onduidelijkheid over de kansen als de valkuilen die een zelfevaluatieproces met zich mee kan brengen. Men kan zich vooraf moeilijk een beeld vormen van welke kennis, vaardigheden en attitudes nodig zijn om het proces tot een goed einde te brengen. Externe ondersteuners worden in dat geval betrokken omdat ze volgens de respondenten precies over deze expertise beschikken. Scholen hebben daarin vertrouwen. Wanneer de zelfevaluaties grotendeels achter de rug waren, gaven de schoolleiders aan in de toekomst meer taken in eigen handen te kunnen nemen. Toch zou men een beroep blijven doen op ondersteuners. Dat blijkt ook uit het feit ook de twee scholen met zelfevaluatie-ervaring opnieuw een beroep deden op ondersteuners. Naast het binnenhalen van externe expertise heeft ondersteuning immers nog andere voordelen. De aanwezigheid van begeleiding houdt bijvoorbeeld de druk op de ketel. De timing van het zelfevaluatieproces wordt daardoor beter bewaakt. Daarnaast is externe ondersteuning een garantie voor het bewaken van objectiviteit en onpartijdigheid. Wanneer dit in een veilige context gebeurt, kunnen leerkrachten meer open spreken over het schoolbeleid in het bijzijn van externen dan in het bijzijn van collega's of hun schoolleider(s). Externen kunnen bovendien bij de schoolleiding aspecten aankaarten, die door leerkrachten niet in de verf gezet kunnen worden. Ondersteuning kan de kwaliteit van zelfevaluatie volgens de respondenten dus op verschillende manieren ten goede komen (cf. hypothese 7).

In de bestudeerde scholen hadden de ondersteuners een grote impact op het verloop van de zelfevaluatie. Dat is gegeven het strikte schema van het ondersteuningsproces niet verwonderlijk. De impact uitte zich in de eerste plaats in organisatorische en praktische zaken. De ondersteuners doen suggesties omtrent hoe de zelfevaluatie best aangepakt wordt. Het vertrouwen in de ondersteuners maakt dat scholen zich sterk houden aan de voorstellen die door ondersteuners gedaan worden. De scholen ondergaan als het ware de ondersteuning waarop ze een beroep doen. Mede daardoor zijn ze niet tegen problemen aangelopen waarmee ze zonder ondersteuning wellicht wel te kampen zouden hebben gehad. We stelden ook vast dat er hierdoor slechts in beperkte mate aandacht was voor 'interne ondersteuning en samenwerking'. Een sterk uitgebouwde externe ondersteuning

lijkt minder eisen te stellen aan de interne ondersteuning, en vice versa. We stelden ook vast dat het ondersteunen van de coördinatoren van de zelfevaluatie niet impliceert dat ook teamleden zich ondersteund voelden. Toch bleken teamleden aan te geven geen problemen te hebben ervaren bij hun weliswaar beperkte rol in het zelfevaluatiegebeuren.

Het ondersteuningstraject eindigde in de bestudeerde scholen met het bespreken van het zelfevaluatierapport. Daarna waren de scholen weer op zichzelf aangewezen. De schoolleiders begrijpen waarom ondersteuners een strikt onderscheid maken tussen 'beschrijven' en 'beoordelen' maar zijn vragende partij voor bijkomende ondersteuning bij het interpreteren van de zelfevaluatieresultaten en bij het komen tot actiepunten. De gevalstudies tonen aan dat zelfevaluaties moeizaam verlopen vanaf het moment dat de ondersteuning wegvalt. Volgens de respondenten zou ook het zelfevaluatieverslag meer aanknopingspunten voor verdere acties moeten bevatten. De ondersteuning zou zich dus moeten verruimen. Dat betekent dat de ondersteuners van de zelfevaluatie scholen langer behoren op te volgen of dat verschillende activiteiten inzake begeleiding van scholen op systeemniveau beter op elkaar afgestemd dienen te worden.

De respondenten ervaren voornamelijk de contactmomenten met ondersteuners binnen de school als zeer waardevol. Het gaat dan bijvoorbeeld om intakegesprekken, het informeren van de stuurgroep over het waarom en hoe van de zelfevaluatie, het begeleiden van verhelderingsgesprekken en gesprekken over de resultaten van de zelfevaluatie. De nascholingssessies buiten de school zouden minder tot de kwaliteit van de zelfevaluatie bijgedragen hebben. Dat kan toe te schrijven zijn aan de handelingsgerichtheid van de bestudeerde technische en beroepsscholen. De theoretische achtergrond en de abstracte beschouwingen die men op plenaire nascholingssessies meekrijgt, beantwoorden niet aan een behoefte in de bestudeerde scholen. Een andere verklaring kan zijn dat de ondersteuning voornamelijk tijdens individuele contacten sterk op maat van de school wordt vormgegeven. Daarmee wordt bevestigd dat ondersteuning waardevoller is naarmate deze meer aangepast is aan de specifieke behoeften van scholen (Nevo, 1995; Schollaert, 2000). Dat blijkt bijvoorbeeld ook uit de vaststelling dat ondersteuning meer tot kwaliteitsvolle zelfevaluaties leidt wanneer de school en de ondersteuners (ook letterlijk) dezelfde taal spreken.

Ondersteuning is de kwaliteit van de zelfevaluaties dus voornamelijk ten goede gekomen omdat ze het zelfevaluatieproces in de verschillende scholen in goede banen leidde. Ze heeft scholen kennis, inzichten en vaardigheden bijgebracht. Dat is een verdienste maar in principe zou men nog meer kunnen verwachten. De ondersteuning heeft volgens de respondenten immers er niet toe bijgedragen dat scholen in de toekomst het geleerde ook spontaan in hun eigen situatie zullen gaan toepassen. De respondenten geven immers aan dat bijvoorbeeld de houding tegenover zelfevaluatie niet positiever geworden is bij teamleden. Als ondersteuning wil bijdragen tot een verdere verspreiding en een groter draagvlak voor zelfevaluatie als strategie voor kwaliteitszorg, zal er ook meer naar veranderingen op integratieniveau moeten gestreefd worden (Peeters e.a., 1996).

6.2 Theoretische reflecties naar aanleiding van de gevalstudies

Deze gevalstudie draagt bij tot de kennisbasis omtrent zelfevaluaties in Vlaamse scholen. Dat blijkt uit de bevestiging en nuancering van de vooropgestelde onderzoekshypothesen. Toch blijven er nog vele onbekenden en onzekerheden. We eindigen de conclusies van de gevalstudies daarom met een aantal theoretische reflecties. De voorbereidende studies resulteerden in een omvattend theoretisch kader. Zowel de kwantitatieve als de kwalitatieve analyses hebben evidentie opgeleverd die de geldigheid van dit model ondersteunen. De hypothesen die we op basis van het theoretisch kader vooropstelden werden ook in de gevalstudies in grote mate bevestigd, net als in de kwantitatieve analyses. De veronderstelde relaties tussen de schoolgebonden kenmerken, de zelfevaluatiegebonden kenmerken, de ondersteuning en de kwaliteit van zelfevaluaties sluiten dus in grote mate aan bij de onderzoeksbevindingen. Een noemenswaardig deel van de variantie in het verloop en in de kwaliteit van zelfevaluatie kan in de lijn der verwachtingen worden verklaard. De vaststelling dat er nog onverklaarde variantie is, wijst er evenwel op dat er in de realiteit nog factoren van belang zijn die niet in het theoretisch kader opgenomen zijn en die toch van invloed zijn (ervan uitgaande dat het niet alleen om meetfouten gaat). In de volgende alinea's doen we een aantal suggesties voor aanvullingen bij het theoretisch kader. Uitgaande van de bevindingen van de gevalstudies kan een bijkomende verklaringsgrond immers mogelijk gevonden worden in:

- *bijkomende informatie over organisatorische kenmerken van de school* waarin de zelfevaluatie plaats vindt. Daarbij denken we aan de vraag of in scholen al dan niet ruimte wordt gemaakt voor het creëren van een middenkader, of er een werkgroep is die doelbewust belast wordt met aspecten van kwaliteitszorg en de vraag hoe bevoegdheden verdeeld worden tussen verschillende leden van de schoolleiding. Ook de impact van de scholengemeenschap rekenen we hierbij.
- *bijkomende informatie over de aard van de specifieke zelfevaluatie-instrumenten*. Het gaat hier bijvoorbeeld om de vraag of de inhoud van het zelfevaluatierapport aansluit bij de behoeften van de gebruikers, of al dan niet gebruik werd gemaakt van verhelderingsgesprekken en om de vraag of de resultaten op een eenvoudige manier de stap naar het komen tot actiepunten mogelijk maken. Ook de vraag of het instrument zich inhoudelijk eerder richt op pedagogisch-onderwijskundige aspecten dan wel op aspecten van het beleidsvoerend vermogen van scholen zou van belang kunnen zijn. Dat pleit ervoor om in het theoretisch kader ook een impact te voorzien van de karakteristieken van het gebruikte zelfevaluatie-instrument (Visscher, 2002).
- *bijkomende informatie over de zelfevaluatieresultaten*. Het is ook niet ondenkbaar dat het beeld dat in de zelfevaluatieresultaten naar voren komt ook een invloed heeft op het gebruik ervan. Goede resultaten vergen een andere opvolging dan minder goede of slechte resultaten. In het kwantitatieve luik van deze studie werden van alle zelfevaluaties dezelfde uitkomsten verwacht, ongeacht hoe de school scoorde in het zelfevaluatieverslag. Daardoor werd mogelijks aan de complexiteit en het unieke van zelfevaluatieprocessen voorbij gegaan. Ook de vraag welk beeld er van de schoolleiding in het zelfevaluatierapport naar voren komt, zou wel eens een impact

kunnen hebben. Het is niet ondenkbaar dat de initiatieven die men in scholen neemt daar mede afhankelijk van zijn.

- *bijkomende informatie over de aard van de externe ondersteuning*. De ene invulling van externe ondersteuning is wellicht effectiever dan de andere. Om hoeveel contactmomenten gaat het? Waar begint (inclusief instrumentkeuze) en eindigt (inclusief implementatie van actiepunten) de ondersteuning? Hoe verhoudt de ondersteuning zich tot andere externe initiatieven en tot de interne professionalisering?
- *bijkomende informatie over de individuele kenmerken van de respondenten*. We hebben vastgesteld dat verschillende respondenten zelfevaluatieprocessen en zelfevaluatieresultaten anders benaderen. Het zou kunnen dat sommige respondenten alles wat er in hun school gebeurt sceptischer benaderen en dus negatiever ervaren dan anderen. Het is niet ondenkbaar dat dergelijke verschillen een rol spelen in het geval van zelfrapportage over de zelfevaluatie die in hun school plaats vond. Bij uitbreiding kan het interessant zijn om na te gaan of persoonlijkheidstrekken (bijvoorbeeld de zogenaamde big 5 (Goldberg, 1993)) van invloed zijn op de manier waarop individuen reflectie en zelfevaluatieprocessen ervaren. Een gelijkaardige vraag dringt zich op inzake verschillen in leerstijlen tussen respondenten. De specifieke aard van zelfevaluatieactiviteiten zou wel eens kunnen impliceren dat ze verschillend gepercipieerd en beoordeeld worden door mensen met verschillende leerstijlen (Vermunt, 1992; Kolb, 1993). Ook de motivatie van respondenten om zich te engageren in kwaliteitszorgprocessen kan bijkomende verklaringskracht bieden (Hofman, Dijkstra, de Boom & Hofman, 2004).
- *bijkomende informatie over aspecten van de rijkheid aan informatie van de beleidscontext in de school*. We stelden in de gevalstudies duidelijk vast dat de bevindingen van een zelfevaluatie op de achtergrond kunnen verdwijnen wanneer deze om aandacht moeten concurreren met andere bronnen van informatie. De vraag is dan hoe prioritair de zelfevaluatiegegevens zijn (Schildkamp, 2007). Er kan immers sprake zijn van een spanning tussen de onderzochte zelfevaluatieresultaten enerzijds en doorlichtingsverslagen, resultaten van een deelname aan (internationaal vergelijkend) onderzoek, resultaten van peilingsonderzoek of netgebonden centrale toetsen, of informatie uit relevante databanken anderzijds.

Toekomstig onderzoek zal zich onder meer moeten richten op het toetsen van de hypothesen die in deze theoretische reflecties vervat zitten.

7 AFRONDEND

In dit deel beschreven wij de aanpak en de resultaten van de gevalstudie van zelfevaluaties zoals deze in 3 secundaire scholen plaatsvonden. Op basis van deze bevindingen kwamen wij tot een antwoord op de onderzoeksvragen en tot suggesties voor aanvullingen bij het theoretisch kader. Doorheen de verticale en horizontale analyse van de bestudeerde zelfevaluatie werd bovendien ‘vlees en bloed’ gegeven aan het bestaande

theoretische kader. Daarmee werden de vooropgestelde doelen gerealiseerd. Rest ons nog om de evidentie uit de verschillende onderzoeksactiviteiten te integreren. Dit doen we in het afsluitende deel van dit proefschrift.

DEEL 9

SYNTHESE EN EINDBESCHOUWINGEN

1 ACHTERGROND, ONDERZOEKSVRAGEN EN METHODOLOGIE

Het uitgangspunt van deze studie was de vaststelling dat vormen van zelfevaluatie op schoolniveau in vele Vlaamse scholen gemeengoed aan het worden zijn. Uitgaande van de intenties van Vlaamse beleidsmakers kan overigens verwacht worden dat de rol van zelfevaluaties in de toekomst nog zal toenemen. In de beleidsbrief van de Vlaamse minister van Werk, Onderwijs en Vorming lezen we dat het stelsel van interne kwaliteitszorg verder moet worden uitgebouwd en versterkt, onder meer door het uitwerken van instrumenten voor zelfevaluatie (Vandenbroucke, 2004, 50). De verwachtingen inzake zelfevaluatie zijn dan ook in sterke mate positief. De empirische evidentie voor deze positieve verwachtingen bleek evenwel beperkt. De beschikbare evidentie wijst erop dat het voor scholen niet evident is om in de eigen vertrouwde context het systematisch en cyclische proces van de zelfevaluatie te initiëren en realiseren (Devos e.a., 1999; Van Petegem e.a., 2005b; Schildkamp, 2007). Er is daarnaast onduidelijkheid omtrent de basis waarop beoordelingen over de kwaliteit van zelfevaluatie (cf. meta-evaluatie) kunnen worden uitgesproken. Dat uit zich bijvoorbeeld in de vraag welke verwachtingen men met het uitvoeren van een zelfevaluatie dient te realiseren om deze als kwaliteitsvol te bestempelen (Harrington & Harrington, 1994). Verscheidene aspecten van het uitvoeren van zelfevaluaties verdienen dan ook de aandacht van onderzoekers. In wezen komt het er op neer dat (de determinanten van) zowel het proces als de resultaten van een zelfevaluatie onder de loep dienen te worden genomen. In de eerste plaats wilden we met het uitvoeren van deze studie dan ook bijdragen tot de kennisbasis omtrent zelfevaluaties in (Vlaamse) scholen. Daarbij werd uitgegaan van de volgende vier onderzoeksvragen:

(OV1) Aan de hand van welke indicatoren kan men de kwaliteit van een zelfevaluatie operationaliseren?

(OV2) Hoe kwaliteitsvol verlopen zelfevaluaties in Vlaamse scholen en wat zijn de verschillen tussen scholen terzake?

(OV3) In welke mate wordt het verloop van een zelfevaluatie bepaald door (a) kenmerken van het functioneren van de school en (b) de ondersteuning die scholen genieten?

(OV4) In welke mate worden de resultaten van een zelfevaluatie bepaald door (a) het verloop van de zelfevaluatie en (b) de ondersteuning die scholen genieten?

Om de bovenstaande onderzoeksvragen te beantwoorden werd geopteerd voor een 'multimethode onderzoeksdesign' (Brewer & Hunter, 1989). Dit is een procedure voor het verzamelen, analyseren en combineren van kwantitatieve en kwalitatieve gegevens in één studie. Deze werd ingegeven door de meerwaarde van een gepaste sequentiële én concurrente combinatie van verschillende onderzoeksmethoden (Creswell e.a., 2003). In concreto omvatte het design een literatuuronderzoek, een Delphi-studie, surveyonderzoek en gevalstudies.

1. Het *literatuuronderzoek* bracht in kaart wat er over (het verklaren en beoordelen van) de kwaliteit van zelfevaluatie in de literatuur reeds beschreven is.
2. De *Delphi-studie* omvatte een systematische kwalitatieve bevraging van bevoorrechte getuigen die bij het uitvoeren van een zelfevaluatie betrokken (kunnen) zijn.
3. Het *surveyonderzoek* bestond in hoofdzaak uit een schriftelijke bevraging van ruim 2700 respondenten (schoolleiders, leerkrachten en ondersteunend personeel) in een 95-tal scholen over hoe zij het verloop en de kwaliteit van de zelfevaluatie die in hun school plaatsvond percipiëren.
4. De *gevalstudies* omvatten een diepgaande studie van 3 scholen doorheen het uitvoeren van een zelfevaluatie door middel van diepte-interviews met verschillende teamleden.

De belangrijkste bevindingen uit deze onderzoeksactiviteiten en de aanbevelingen die er uit voortvloeien worden in de volgende paragrafen toegelicht.

2 EMPIRISCHE EVIDENTIE: ANTWOORDEN OP DE ONDERZOEKSVRAGEN

Na elk deelonderzoek werden de waarnemingen gerapporteerd en besproken. In dit synthetiserend deel kijken we over de verschillende deelonderzoeken heen. Dat doen we aan de hand van wat wij als de belangrijkste bevindingen van deze studie naar voren schuiven. Het onderzoek brengt ons tot een aantal vaststellingen die we telkens kort onderbouwen aan de hand van evidentie uit de verschillende onderzoeksactiviteiten. De eerste vaststellingen zijn een antwoord op de onderzoeksvragen bij deze studie. Daarna wordt ook een aantal bevindingen gepresenteerd dat niet rechtstreeks aan een onderzoeksvraag gelinkt is. Omdat de volgorde van de vaststellingen los staat van hun belang werden ze doelbewust niet genummerd.

Je kan meten of een zelfevaluatie kwaliteitsvol is door na te gaan of deze aan de vooropgestelde verwachtingen voldoet (bij OV1)

De verschillende onderzoeksactiviteiten hebben alle aangetoond dat scholen zeer uiteenlopende verwachtingen kunnen koesteren bij het uitvoeren van een zelfevaluatie. Op

verschillende plaatsen werd er daarom voor gepleit om te werken met een *school- of zelfs respondenteigen indicator (Indicator 1)*. Ook binnen scholen blijken de verwachtingen immers sterk te kunnen verschillen. Een zelfevaluatie is dan kwaliteitsvol wanneer ze de verwachtingen van de betrokkenen weet te realiseren (cf. Harrington & Harrington (1994)). Wanneer men – in de context van schoolontwikkeling[35] - uitspraken wil doen over de kwaliteit van zelfevaluaties in individuele scholen neemt men dus best de individuele verwachtingen van de betrokkenen als uitgangspunt. Onder meer in de gevalstudies bleek duidelijk dat het van cruciaal belang is om daarbij een onderscheid te maken tussen de verwachtingen van teamleden en de verwachtingen van schoolleiders.

De school- en/of respondenteigen benadering mag conceptueel dan wel het meest zuiver zijn, echt werkbaar is deze niet wanneer men over de zelfevaluaties in grote groepen scholen uitspraken wenst te doen. Naar aanleiding van de bevindingen in de gevalstudies pleitten we daarom voor een focus op de grootste gemene deler van de veelheid aan verwachtingen. Deze blijkt – zo stelden we in de Delphi-studie en de gevalstudies vast - te bestaan uit de verwachting dat zelfevaluaties moeten *resulteren in 'passende actiepunten' (Indicator 2) en in 'een aantal concrete resultaten' (Indicator 3)*. Deze indicatoren noemden we de schooloverstijgende kwaliteitsindicatoren. Onder 'passende actiepunten' verstaan we actiepunten die specifiek, meetbaar, acceptabel, realistisch en tijdsgebonden zijn (Conzemius & O'Neill, 2002). Voorbeelden van 'concrete resultaten' zijn een goed beeld geven van het functioneren van onze school, de wil om kritischer tegen de zaken aan te kijken bevorderen, het vergroten van de motivatie om kwaliteitsvol onderwijs te verzorgen en het versterken van de samenhang in de school. Het survey-onderzoek toonde aan dat het mogelijk is beide kwaliteitsindicatoren aan de hand van een betrouwbare meetschaal te operationaliseren.

We willen in deze conclusies nog meegeven dat noch in de Delphi-studie, noch in de gevalstudie expliciet de link werd gelegd tussen kwaliteitsvolle zelfevaluatie en verbeterde leerlingenprestaties. Het is aannemelijk dat de respondenten van de andere verwachtingen ten aanzien van zelfevaluatie een indirecte impact verwachten op het leren van leerlingen maar het effect van zelfevaluatie op de leerresultaten van leerlingen werd in deze studie niet als een kwaliteitsindicator gehanteerd. Ander onderzoek wijst in gelijkaardige contexten op het bestaan van een klein positief (Tymms & Albone, 2002) of geen effect (Schildkamp, 2007).

---

[35] De door ons gehanteerde definitie van zelfevaluatie – en ook de invulling die de respondenten er in de Delphi-studie aan gaven – plaatst zelfevaluatie in een context van schoolontwikkeling en benadrukt het belang van de eigen verantwoordelijkheid van scholen in het zelfevaluatieproces. Door de nadruk op het ontwikkelingsaspect werd in deze studie niet sterk ingegaan op de indicatoren voor kwaliteitsvolle zelfevaluatie in een verantwoordingscontext. De situatie is daar – conceptueel tenminste - eenvoudiger. De verwachtingen ten aanzien van de zelfevaluatie worden dan immers niet door de school zelf opgelegd, maar komen van externen. (Pas) wanneer deze externe verwachtingen duidelijk geëxpliciteerd worden en transparant naar scholen toe gecommuniceerd worden, kunnen zij de basis vormen van de kwaliteitsbeoordeling.

Scholen verschillen wat de kwaliteit van hun zelfevaluaties betreft (bij OV2)

De Delphi-studie, de kwantitatieve analyses en de gevalstudies hebben duidelijk aangetoond dat scholen in sterke mate verschillen wat de kwaliteit van hun zelfevaluaties betreft. De vraag naar de kwaliteit van zelfevaluaties in Vlaamse scholen behoeft dus een genuanceerd antwoord. Vandaar ook het belang van het gebruik van verschillende invalshoeken. Bij de normgerichte invalshoek vergeleken we de kwaliteit van zelfevaluaties in scholen met deze van zelfevaluaties in andere scholen. Daaruit bleek dat er (grote) verschillen zijn tussen scholen onderling. Sommige scholen scoren statistisch significant beter dan anderen op de drie vooropgestelde indicatoren. Ook de criteriumgerichte invalshoek maakte duidelijk dat er grote verschillen zijn (Davies & Rudd, 2001). Als we de kwaliteit van een zelfevaluatie afmeten tegen een te verwachten standaard dan bevestigde dit opnieuw de grote variatie in kwaliteitsbeoordelingen. Wanneer men een weinig veeleisende standaard hanteert (zoals in de GOK-inspectie (MVG, 2006)) dan scoren de meeste scholen 'voldoende tot goed' of zelfs 'zeer goed'. Wanneer men de lat hoog legt (zoals gesimuleerd in de kwantitatieve analyses bij zelfevaluatie binnenstebuiten) dan blijkt de zelfevaluatie in de meeste scholen weinig tot niet kwaliteitsvol. Dit sluit aan bij de conclusies bij andere studies (Schildkamp, 2007). In de schoolgerichte invalshoek tenslotte beargumenteerden we dat ook rekening moet worden gehouden met de eerdere ervaringen en realisaties van scholen inzake zelfevaluatie. De gevalstudies illustreerden enerzijds dat het uitvoeren van een zelfevaluatie in sommige scholen momenteel nog veel weg heeft van een vernieuwingsproces. Een beperkte kwaliteit kan in die context aanvaardbaar zijn. De meerniveau-analyses toonden anderzijds aan dat het al dan niet hebben van ervaring met zelfevaluaties (na controle) niet statistisch significant samenhangt met de vooropgestelde drie kwaliteitsindicatoren. De verschillen inzake kwaliteitsvolle zelfevaluatie kunnen dus niet zomaar teruggebracht worden tot het al dan niet hebben van eerdere ervaring met zelfevaluatieprocessen.

Er zijn zowel verschillen tussen als binnen scholen (bij OV2)

Het beeld dat sommige scholen het beter doen dan andere, behoeft nuancering. Dit onderzoek heeft duidelijk aangetoond dat ook de verschillen binnen scholen groot zijn. Dat geldt zowel voor de kwaliteitsindicatoren als voor de zelfevaluatiegebonden kenmerken. Wat de 'houding ten aanzien van zelfevaluatie' betreft, bleek 7% van de totale variantie in de scores van respondenten toe te schrijven aan verschillen tussen scholen. De resterende 93% moet toegeschreven worden aan verschillen tussen respondenten binnen scholen. Het beeld van de ene school waar iedereen een positieve houding heeft, en de andere waar iedereen een negatieve houding heeft strookt dus niet met de realiteit. In elke school zijn er grote verschillen tussen respondenten. Ook uit de gevalstudies bleek dat dit implicaties heeft voor meta-evaluatie. Deze implicaties gaan verder dan het onderscheid tussen schoolleiders en teamleden. Ook tussen teamleden zijn er verschillen. Dat maakt zowel het opzetten als het beoordelen van een zelfevaluatie moeilijk. Wat door de één als een beargumenteerde en wenselijke aanpak van een zelfevaluatie gezien wordt, kan

door de ander als betuttelend en overbodig beschouwd worden. Daarbij denken we bijvoorbeeld aan het belang dat moet besteed worden aan sensibilisering, communicatie en gezamenlijke besluitvorming. Dit onderzoek leert dat het gevaarlijk is te denken in termen van stereotypen van scholen en van de mensen die erin actief zijn.

Het verloop van zelfevaluaties hangt samen met het ruimere functioneren van scholen (bij OV3a)

De manier waarop zelfevaluaties verlopen, staat niet los van het ruimere functioneren van de school waarin deze plaatsvinden. De meerniveau-analyses hebben aangetoond dat informatie over het functioneren van de school samenhangt met de zelfevaluatieprocesgebonden kenmerken (met name de houding ten aanzien van zelfevaluatie, 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad'). Wanneer men niet controleert voor andere kenmerken (cf. in de bruto-modellen) blijken zelfevaluaties volgens de respondenten meer aan deze drie kenmerken te voldoen naarmate de respondenten meer van mening zijn dat er in hun school sprake is van diepgaand leren, van kenniscreatie, van een hoge score op de effectiviteitsperspectieven (menselijke relaties, interne processen, open systeem en rationele doel) of van een sterke sociaalconstructivistische visie op het eigen leren. Indirect zal ook de zelfevaluatie hierdoor als kwaliteitsvoller beschouwd worden (zie volgende vaststelling). Wanneer men de verschillende schoolgebonden kenmerken gezamenlijk modelleert (én controleert voor de functie van respondenten en de impact van ondersteuning), blijken slechts drie van de bevraagde schoolgebonden kenmerken een statistisch significante impact te hebben op de zelfevaluatieprocesgebonden kenmerken. Deze zijn de score op het effectiviteitperspectief 'rationele doel', de score op 'diepgaand leren' en de score op 'kenniscreatie'. Dit wijst erop dat scholen - wanneer men garanties wil inbouwen voor een goed verloop van de zelfevaluatie - eerst op deze kenmerken voldoende hoog moeten scoren (Schildkamp, 2007). De schoolgebonden kenmerken zijn als het ware een indicatie voor de 'rijpheid', de 'startpositie' of zo men wil 'de basiscompetenties' waarover een school moet beschikken om van zelfevaluatie ten volle de vruchten te kunnen plukken. Gegeven onze onderzoeksbevindingen bieden deze kenmerken van het functioneren van de school garanties voor een goed verloop van het zelfevaluatieproces. Het belang van onder meer innovatiebereidheid (Geijsel e.a., 1998), het onderschrijven van gezamenlijke doelen (Leithwood, Aitken & Jantzi, 2001), reflectiebereidheid en de wil om samen te werken (Mortimore & MacBeath, 2001) wordt in dit onderzoek bevestigd.

De kwaliteit van zelfevaluaties hangt samen met het verloop van het zelfevaluatieproces (bij OV4a)

In de verkennende Delphi-studie bleek er een grote eensgezindheid bij de respondenten omtrent de relevantie van verschillende voorspellers van kwaliteitsvolle zelfevaluaties. Er werd benadrukt dat het uitvoeren van een zelfevaluatie *een positieve houding* vergt ten aanzien van zelfevaluatieprocessen. Verder werd aangegeven dat in het uitvoeren van een zelfevaluatie dezelfde kenmerken een rol spelen als bij het voeren van andere vor-

men van schoolbeleid. Zelfevaluaties werden in het verdere verloop van het onderzoek daarom beschouwd als *'het uitvoeren van een beleidsdaad'*. Ze vergen een voldoende groot beleidsvoerend vermogen van scholen. Ten derde werd aangegeven dat zelfevaluaties ook *'als een onderzoeksdaad'* benaderd moeten worden. Op die manier benadrukt men het belang van een systematische aanpak van het gebeuren. Scholen moeten daartoe in voldoende mate inzicht hebben in het uitvoeren van sociaalwetenschappelijk onderzoek. Ook Blok et al. kwamen in hun onderzoek tot deze conclusie (2005).

De hypothesen die we op basis van de Delphi-studie en het literatuuronderzoek formuleerden, werden bevestigd in de kwantitatieve analyses én in de gevalstudies. Dit onderzoek leidde tot sterke empirische evidentie voor de 'houding ten aanzien van zelfevaluatie', 'zelfevaluatie als beleidsdaad' en 'zelfevaluatie als onderzoeksdaad' als krachtige voorspellers van de kwaliteit van zelfevaluaties. De drie voorspellers hebben alle een statistisch significante impact. Respondenten die aangeven dat hun school hoog scoort op deze drie kenmerken van het verloop van de zelfevaluatie geven ook aan dat de zelfevaluatie in hun school kwaliteitsvol was. De kwaliteit van zelfevaluaties hangt dus duidelijk samen met het verloop van het zelfevaluatieproces. In de kwantitatieve analyses stelden we vast dat het aangewezen is om in de eerste plaats uit te gaan van het concept 'zelfevaluatie als beleidsdaad' wanneer men de kwaliteit van een zelfevaluatie op een zo spaarzaam mogelijke manier wil voorspellen.

Goede ondersteuning brengt scholen tot meer kwaliteitsvolle zelfevaluaties (bij OV3b en OV4b)

Een voorwaarde voor kwaliteitsvolle zelfevaluaties is dat de noodzakelijke expertise in de school aanwezig is en/of door een beroep te doen op externe ondersteuning in de school wordt binnengehaald. De Delphi-studie en de gevalstudies toonden aan dat de expertise momenteel nog onvoldoende in scholen zelf aanwezig is. Dat verklaart het belang van de beschikbaarheid van goede externe ondersteuning. De verschillende onderzoeksactiviteiten leverden elk op hun manier evidentie aan voor de heilzame impact van goede ondersteuning. In het volledige verklaringsmodel (in de kwantitatieve analyses) bleek dat ondersteuning zowel een direct als een indirect effect heeft op de kwaliteit van de bestudeerde zelfevaluaties. Wanneer de ondersteuning tot wenselijke resultaten leidde, spreken respondenten zich doorgaans positiever uit over het verloop van de zelfevaluatie. Precies door de impact op het verloop van de zelfevaluatie leidt ondersteuning indirect tot een betere kwaliteit. Daarnaast is er het directe effect. Goede ondersteuning leidt dus ook rechtstreeks tot een betere score op de kwaliteitsindicatoren. Dat kan verklaard worden door de invloed van ondersteuning op het interpreteren van zelfevaluatieresultaten, het komen tot actiepunten en het implementeren van actiepunten. We stelden ook vast dat respondenten die geen ondersteuning genoten doorgaans geen negatiever beeld hebben over de zelfevaluatie in hun school dan zij die verhoudingsgewijs minder kwaliteitsvolle ondersteuning genoten. Ondersteuning heeft dus enkel een impact als deze voldoende kwaliteitsvol is. Scholen die hun ondersteuning als minder goed omschrijven

komen dus doorgaans niet tot minder goede zelfevaluatie dan scholen die geen ondersteuning genoten.

Het belang van voldoende kwaliteitsvolle ondersteuning kunnen we op basis van de resultaten van de Delphi-studie en de gevalstudie inhoudelijk invullen. Het gaat om ondersteuning die op een ervaringsgerichte en interactieve manier een zo schooleigen aanpak als mogelijk weet te realiseren. Daarbij werd sterk benadrukt dat ondersteuning best het gehele proces van de zelfevaluatie in ruime zin omvat, inclusief de interpretatie, het komen tot actiepunten en het implementeren ervan. Wie de ondersteuning inricht blijkt in wezen niet van belang, zolang de externe ondersteuners door scholen maar niet gelinkt worden aan een verantwoordingsperspectief en inspelen op de eigenheid en behoeftes van de individuele school.

Schoolleiders verschillen van teamleden, maar ook weer niet

De verschillen tussen schoolleiders en teamleden lopen als een rode draad doorheen de resultaten van de kwantitatieve analyses en de gevalstudies. Schoolleiders hebben bijvoorbeeld een meer positieve houding ten aanzien van zelfevaluatie en zij spreken zich in sterke mate positiever uit over de kwaliteitsindicatoren van de zelfevaluatie dan teamleden. We stelden evenwel ook vast dat zelfevaluaties momenteel in sterke mate een zaak zijn van de schoolleiding en in mindere mate van de teamleden. Doordat ze sterker betrokken worden bij en vaak de architecten zijn van het hele gebeuren, hoeft het niet te verbazen dat zij zich positiever uitspreken over het verloop van de zelfevaluatie en de resultaten ervan hoger inschatten. Het is dan niet noodzakelijk zo dat schoolleiders de zelfevaluatie anders (lees: positief gekleurd) ervaren. Zij ervaren bij wijze van spreken gewoonweg een andere zelfevaluatie. De verschillen tussen schoolleiders en teamleden zijn in deze logica eerder een gevolg van de verschillende manieren waarop beide groepen momenteel bij zelfevaluaties betrokken worden dan van verschillen die eigen zijn aan de functie of aan de personen die deze functies vervullen. Uit het volledige verklaringsmodel in de kwantitatieve analyses leerden we immers dat het effect van functie op de kwaliteitsindicatoren verdwijnt wanneer gecontroleerd wordt voor zelfevaluatiegebonden kenmerken. Dat betekent dat schoolleiders die dezelfde houding hebben als teamleden ten aanzien van zelfevaluatie en die dezelfde mening hebben over het verloop van de zelfevaluatie (als beleidsdaad en onderzoeksdaad) niet hoger scoren op de kwaliteitsindicatoren dan deze teamleden. Maar, gezien de verschillen tussen teamleden en schoolleiders op deze punten momenteel groot zijn, is ook het verschil in de kwaliteitsbeoordeling groot.

De terughoudendheid ten aanzien van zelfevaluatie is eerder toe te schrijven aan onzekerheid over het proces dan aan een gebrek aan geloof in het product

In de verschillende gegevensverzamelingen kwam telkens naar voren dat er een terughoudendheid bestaat in Vlaamse scholen ten aanzien van het uitvoeren van zelfevalua-

ties. De Delphi-studie en de gevalstudies toonden aan dat zelfevaluaties door velen nog als een vreemd iets worden ervaren. Vormen van reflectie – zoals de hier bestudeerde zelfevaluaties - die een grote openheid naar anderen en een kwetsbare opstelling vereisen, liggen nog steeds moeilijk. Toch is er een groot geloof in de mogelijkheden van reflectie en zelfevaluatie. De bevraagde respondenten koesteren positieve verwachtingen ten aanzien van zelfevaluatie. Men erkent dat het proces tot waardevolle resultaten kan leiden (instrumenteel, conceptueel en symbolisch) maar de aard van de activiteiten die met zelfevaluatie gepaard gaan maakt dat de stap om er zich spontaan in te engageren of er geëngageerd in mee te stappen vaak nog te groot is. De oorzaak van de afwachtende houding lijkt dus eerder bij het proces dan bij het product van zelfevaluatie te liggen. Men vindt bijvoorbeeld dat zelfevaluatie veel tijd kost en dat het moeilijk uit te voeren is. Bovendien staat de handelingsgerichtheid en een gebrek aan openheid in veel scholen het vrijuit spreken over de eigen bezorgdheden en problemen in de weg. Men gelooft dus wel in de potentiële kracht en meerwaarde van zelfevaluaties, maar het bijhorende proces maakt dat velen aarzelen.

3 VOORBIJ DE EMPIRISCHE EVIDENTIE: (BELEIDS)AANBEVELINGEN NAAR AANLEIDING VAN DEZE STUDIE

We komen aan het einde van dit onderzoeksrapport ook tot een aantal aanbevelingen. Daarbij baseren we ons op de onderzoeksresultaten waarover in de voorgaande delen gerapporteerd werd. Wij hebben getracht op een wetenschappelijk verantwoorde manier gegevens te verzamelen, deze te ordenen en in hun samenhang te presenteren. Toch is de inventarisatie van mogelijke beleidsaccenten en beleidsdaden niet waardevrij. Zo is het bijvoorbeeld geen toeval dat in de eerste plaats verwezen zal worden naar de rol van scholen zelf. Het is ons inziens een illusie te denken dat een overheid op korte termijn tot fundamentele (mentaliteits)wijzigingen kan bijdragen. Een overheid kan in het komen tot kwaliteitsvolle zelfevaluatie hoogstens faciliteren en prikkelen door de juiste randvoorwaarden te scheppen. Daarbij wordt ons inziens best uitgegaan van een getuigende dan van een overtuigende aanpak. Het volgende Arabisch spreekwoord verwoordt ons standpunt zeer treffend: *'Men kan een paard wel in 't water trekken, maar niet dwingen te drinken. Of het moet dorst hebben'*. Het komt er op aan de scholen dorst te doen krijgen, eerder dan ze tot drinken te dwingen. De volgende aanbevelingen kunnen hier ons inziens toe bijdragen. Ze lichten toe wat we bedoelen en zorgen meteen ook voor nuancering.

Maak van reflectie een strategisch doel in het schoolfunctioneren

Het uitvoeren van zelfevaluaties zoals wij ze bestudeerden is als het ware een systematische vorm van reflectie op team- of schoolniveau. Wij aanzien reflectie daarom als een opstap naar en voorwaarde tot het gebruik van geïnstrumentaliseerde zelfevaluaties. Werken aan kwaliteitsvolle zelfevaluaties vergt dus in de eerste plaats een voldoende

groot reflectief vermogen. Niet toevallig gaat onze eerste aanbeveling daarom niet rechtstreeks over zelfevaluatie maar over reflectie.

Reflectie is momenteel zelden een strategisch doel in het functioneren van scholen. Individuen, groepen en organisaties kunnen echter maar leren op voorwaarde dat ze bereid zijn te reflecteren op hun eigen handelen (Mitchell & Sackney, 2000; Verbiest & Vandenberghe, 2002). De dominante cultuur in veel scholen is er echter één van doen en handelen, met weinig of geen aandacht voor reflectie. Er is in de bestudeerde scholen een sterke handelingsgerichtheid. Men geeft in de gevalstudies aan dat er weinig tijd wordt gespendeerd aan reflectie. Tijdsgebrek wordt hierbij vaak als verontschuldiging ingeroepen. We stelden echter vast dat ook de houding ten aanzien van reflectie en gewoontevorming reflectie bemoeilijken. Te veel leerkrachten aanzien het als een vorm van sociale controle. Dit bedreigend karakter van zelfevaluatie werd ook elders vastgesteld (Clift e.a., 1989). Er heerst vaak een gebrek aan openheid binnen het team en een gebrek aan een kritische ingesteldheid ten opzichte van het eigen functioneren. Alvorens scholen tenvolle de vruchten van zelfevaluatie zullen kunnen plukken, zal er dus sprake moeten zijn van een mentaliteitswijziging. We concluderen dat van een cultuur van reflectie die bijdraagt tot een positieve houding ten aanzien van zelfevaluatie bij de teamleden weinig of geen sprake is. Onze aanbeveling is daarom om eerst van reflectie een strategisch doel in het schoolfunctioneren te maken.

Het is ons inziens in de eerste plaats aan scholen zelf om de mogelijkheden te creëren om op regelmatige tijdstippen stil te staan bij de eigen verantwoordelijkheden en hoe deze ingevuld worden. Deze studie toonde aan dat men zich bewust is van het belang hiervan maar dat de retoriek om allerlei redenen niet in daden wordt omgezet. Wij zijn ervan overtuigd dat de schoolleiding in dit proces een belangrijke rol heeft te vervullen. Die rol bestaat in de eerste plaats niet uit complexe managementtechnieken maar uit eenvoudige menselijke gedragingen. Voorbeelden daarvan zijn het voeren van een opendeurpolitiek, zichtbaar aanwezig zijn, duidelijk maken en communiceren over waar je mee bezig bent en jezelf als schoolleider kwetsbaar en als lerende durven opstellen (McEwan, 2003). Meer structurele suggesties bestaan uit het betrekken van leerlingen in de evaluatie van het school- en klasgebeuren, het betrekken van kritische vrienden of het engageren in actieonderzoek.

In deze aanbeveling schuilt ook een waarschuwing. Scholen met een beperkt reflectief vermogen doen er ons inziens goed aan niet over één nacht ijs te gaan vooraleer ze zich engageren in een geïnstrumentaliseerde zelfevaluatie. Het opzetten van kwaliteitsvolle zelfevaluaties vergt de juiste kennis, vaardigheden en vooral attitudes. Scholen zonder of met weinig ervaring inzake reflectie starten dan ook best kleinschalig en met bescheiden ambities. Het werken rond een beperkt en veilig maar relevant thema kan getuigen van de zinvolheid van reflectie en zelfevaluatie duidelijk maken. Een incrementele aanpak is aangewezen. Zelfevaluatie verkoopt namelijk zichzelf indien ze kwaliteitsvol gebeurt. Helaas maakt ze zichzelf in de toekomst bij wijze van spreken onmogelijk indien dit niet het geval is. De regel ‘baat het niet dan schaadt het niet’ gaat hier niet op. Het heeft dus maar zin om te investeren in geïnstrumentaliseerde zelfevaluaties wanneer de cultuur in

de school uitnodigt tot een reflectieve benadering van het individuele handelen, van het handelen met collega's, van het handelen in team en het handelen als school.

Zelfevaluatie: van een nood naar een behoefte

Teamleden stappen momenteel nog te zeer mee in een zelfevaluatieproces omdat dat dit van hen verwacht wordt. De zelfevaluaties worden in scholen in de eerste plaats uitgevoerd omdat ze door de schoolleiding als een nood ervaren worden. Dat is een fundamenteel andere invalshoek dan het uitvoeren van een zelfevaluatie vanuit een behoefte. Noden zijn 'tekorten die door anderen aangevoeld worden'; behoeftes zijn 'tekorten die door de betrokkene(n) zelf worden aangevoeld en erkend' (Schollaert e.a., 1994, p. 13-14). Dit onderscheid is cruciaal in het implementeren van vernieuwingen zoals het uitvoeren van zelfevaluaties. Door teamleden op te leggen dat zij aan zelfevaluatie moeten doen, getuigt men van een zogenaamde adoptiebenadering. Daarmee wordt bedoeld dat teamleden in sterke mate aangezien worden als leden van een rationele organisatie die uitvoeren wat door anderen als belangrijk (en dus vernieuwingswaardig) vooropgesteld wordt. Het proces van invoeren dreigt dan weinig aandacht te krijgen, waardoor innovaties te weinig aansluiting vinden bij de praktijk (Güthe, 1997). Wanneer men zelfevaluaties ingang wil doen vinden in scholen, zal het er op aan komen aandacht te hebben voor de wisselwerking die er is tussen de inhoud van de innovatie (als teamlid mee de verantwoordelijkheid dragen voor kwaliteitszorg) en kenmerken van de individuele teamleden. De percepties van betrokkenen bijvoorbeeld, en in het bijzonder de betekenis die gegeven wordt aan zelfevaluatie, is van groot belang voor het succes van dit vernieuwingsproces (Fullan, 1991). We stelden in dit onderzoek een terughoudendheid vast wat het uitvoeren van zelfevaluaties betreft. Staessens (1993, p. 127) zegt hieromtrent: '*De reactie van een school op een verandering kan deels begrepen worden door te verwijzen naar de eigen schoolcultuur, of preciezer, naar de discrepantie tussen de waarden en normen die in de school bestaan en de waarden en normen die van de innovatie uitgaan.*' De discrepantie is momenteel nog groot. En zolang een vernieuwing zoals een zelfevaluatie door veel teamleden ervaren wordt als een opgelegd en (daardoor) bedreigend iets, dreigt deze discrepantie nog te vergroten (Van den Berg & Vandenberghe, 1999). Het dient benadrukt dat dit alles niet betekent dat het startpunt van de verandering (c.q. de zelfevaluatie) steeds bij de teamleden moet liggen (cf. de bottom-up strategie). Ook de noden die door anderen ervaren worden kunnen het uitgangspunt zijn. Door een verstandig beleid te voeren, kan men de weerstanden die men nogal eens met een top-down-aanpak associeert immers zo goed als mogelijk vermijden (Reezigt & de Jong, 2001).

Gebruik zelfevaluatie enkel in een ontwikkelingsperspectief

Zowel schooldoorlichtingen als zelfevaluaties zijn noodzakelijke onderdelen van een kwaliteitszorgsysteem. Dit gaat terug op het respectievelijke belang van een verantwoordings- en een schoolontwikkelingsperspectief in zulk een systeem. Beide kunnen ons inziens evenwel niet zomaar gecombineerd worden. Als men tot een goed functionerend samenspel wil komen, zullen beide perspectieven én instrumenten duidelijk van elkaar

moeten worden onderscheiden. Wij onderschrijven dat zelfevaluatie in het teken moet staan van schoolontwikkeling. Intrinsieke motivatie is een noodzakelijke voorwaarde om een degelijk zelfevaluatieproces op gang te brengen. Zelfevaluatie is in dit perspectief geen geschikte methodiek om gegevens te verzamelen met het oog op externe verantwoording. Een zelfevaluatie kan niet onder grote druk van buiten de school plaatsvinden. Zulk een invulling van zelfevaluatie zou immers in strijd zijn met de manier waarop we het concept hoger definieerden. Meer nog: als de functie van een externe evaluatie summatief is, dan is het de vraag of een zelfevaluatie daar überhaupt een rol in kan spelen. Het gevaar is niet denkbeeldig dat scholen zelfevaluatie zien als een verplichting en als een taak los van schoolontwikkeling (Renkema, 2002). De zelfevaluatieresultaten zullen wellicht ver af staan van het feitelijke functioneren van de school en zowel in het kader van verantwoording als schoolverbetering onbruikbaar zijn. Dat alles heeft duidelijke implicaties naar de inzetbaarheid van zelfevaluaties.

We willen hierbij twee nuances aanbrengen. Er is immers een fundamenteel verschil tussen een overheid die haar scholen verplicht om een kwaliteitszorgsysteem te hanteren waarin zelfevaluatie een rol kan spelen en het gebruiken van de eigenlijke zelfevaluatiegegevens in een verantwoordingscontext. De rol van zelfevaluatie kan dus op zeer verscheiden manieren ingevuld worden. Daarnaast mag men niet uit het oog verliezen dat verplichte zelfevaluaties - zoals deze in het kader van het GOK-beleid - in een aantal scholen een katalysatoreffect kunnen hebben gehad. In bepaalde scholen is een vorm van externe druk wellicht noodzakelijk. Scholen met een beperkt beleidsvoerend vermogen kunnen op die manier verplicht worden om het pad van de zelfevaluatie op te gaan. En vermits een pad maar ontstaat door er op te lopen, kan het aangewezen zijn voor deze groep scholen een meer dwingende strategie te hanteren. Het bestaan van een dergelijke groep scholen mag echter niet resulteren in een aanpak die alle scholen op een sturende wijze verplicht om aan zelfevaluatie te doen. De uitdaging bestaat erin een goede balans te vinden tussen ‘externe druk’ en ondersteuning (Fullan, 1991; Visscher, 2002).

Variatie tussen scholen vergt variatie in verwachtingen én variatie in autonomie

Vele onderwijsvernieuwingen hebben in scholen slechts een oppervlakkige betekenis gekregen (Van den Berg & Vandenberghe, 1999). Als het uitvoeren van zelfevaluaties gezien wordt als een waardevolle vernieuwing in het Vlaamse onderwijs, dan stelt zich de vraag hoe deze vernieuwing in bepaalde groepen scholen voor hetzelfde lot behoed kan worden. Scholen verschillen immers in sterke mate, en in het uitstippelen van strategieën worden ze best niet over één kam geschoren. We stelden in dit onderzoek bijvoorbeeld statistisch significante verschillen vast tussen scholen in de houding ten aanzien van zelfevaluatie en in het verloop en de resultaten van zelfevaluaties. Belangrijk daarbij is bovendien dat deze zaken blijken samen te hangen. Dat werpt een enigszins fatalistisch licht op zelfevaluaties in scholen. De beste resultaten worden als het ware geboekt door scholen die er het minst behoefte aan hebben. Binnen de grenzen van de pedagogische

vrijheid van scholen pleiten wij daarom voor variatie in verwachtingen en variatie in autonomie.

Met *variatie in verwachtingen* bedoelen we dat het op dit moment onrealistisch is om van alle scholen hetzelfde te verwachten wat de zelfevaluaties die ze uitvoeren betreft. De invalshoek waarin zelfevaluaties nog als een innovatie benaderd worden, noopt tot een schoolgebonden referentiekader. De uitdaging bestaat erin aan te sluiten bij de zogenaamde (door Vygotsky geïntroduceerde) 'zone van de naaste ontwikkeling' van individuele scholen (zie ook Verbiest, 1995). Dat betekent dat de verwachting door de school als zinvol en haalbaar ervaren wordt en dat deze mits ondersteuning kan gerealiseerd worden. De prestaties van scholen inzake zelfevaluatie worden in die zin langzaamaan omhoog getrokken. Scholen moeten de tijd krijgen om zich op een eigen tempo te ontwikkelen. De variatie in verwachtingen kan zowel vanuit de leerpsychologie als vanuit het innovatiedenken bepleit worden.

Toch is het principe van de 'zone van de naaste ontwikkeling' niet absoluut. Er moet immers een minimum aan verwachtingen gerealiseerd worden om onderwijskwaliteit te garanderen. Omdat de bovenstaande logica niet voor alle scholen opgaat, pleiten we ook voor *variatie in autonomie*. Het effect van een strategie is immers niet in alle scholen hetzelfde. Ook deze studie geeft bewijskracht voor de veronderstelling dat de te volgen strategie best verschilt naarmate scholen over een laag dan wel hoog beleidsvoerend vermogen beschikken (Onderwijsraad, 2001; Van Petegem e.a., 2006). In scholen met een laag beleidsvoerend vermogen en met beperkte zelfevaluatiecapaciteiten is een strikte reglementering een garantie voor onderwijskwaliteit. Het is te verantwoorden dat deze scholen bij het handje gehouden worden. Daarmee bedoelen we bijvoorbeeld dat de schoolleider verplicht wordt zich te professionaliseren, dat er sprake is van strikte overheidsreglementering, dat de school tezamen met haar begeleidingsdienst komt tot beleidsplan met formele mijlpalen of dat formatieve evaluatiemomenten ingelast worden. Zulke strikte vormen van sturing zijn in scholen met een gemiddeld of hoog beleidsvoerend vermogen niet functioneel. Een overheid kan zich daar bij wijze van spreken beperken tot het scheppen van een stimulerend kader en het vooropstellen van minimumdoelen. De verschillen tussen scholen maken het uittekenen van een overheidsbeleid omtrent de rol van zelfevaluaties tot zulk een moeilijke evenwichtsoefening. Het is dansen op een slappe koord met aan de ene zijde het verlenen van autonomie en aan de andere zijde het voorzien van sturing. De verwevenheid met de pedagogische vrijheid van scholen maakt dat een maatschappelijk debat met alle stakeholders zich hier opdringt.

Rust scholen uit met de nodige middelen

Men kan van scholen niet zomaar verwachten dat ze zonder aangepaste middelen succesvol vorm geven aan activiteiten die in wezen als een grondige vernieuwing gepercipieerd kunnen worden. Zelfevaluaties zijn bovendien tijdrovend en vergen de inzet van specifieke middelen (Davies & Rudd, 2001; Kimball, 2002). Die middelen zijn evenwel onvoldoende ter beschikking. Dat is zowel toe te schrijven aan de manier waarop de taakbelasting

van leerkrachten in het leerplichtonderwijs ingevuld wordt als aan de tijd- en dus kostintensieve activiteiten die zelfevaluaties met zich mee brengen.

Een lerarenopdracht bestaat nog steeds uit een lesopdracht waarop de leerkracht wordt afgerekend. Leerkrachten zijn in de eerste plaats dan ook (terecht) begaan met de eigen kerntaken en zijn minder geïnteresseerd in activiteiten op schoolniveau. Toch wordt er ook steeds meer een beroep gedaan op ruimere vaardigheden van de leraar als teamlid. Het probleem is dat de aandacht voor deze nieuwe uitdagingen ten koste kan gaan van de kwaliteit van de oorspronkelijke kerntaak. Bovendien maakt het huidige systeem het moeilijk om zelfevaluatie-activiteiten op een goede manier in te plannen (Davies & Rudd, 2001). Daarbij denken we bijvoorbeeld aan de nefaste impact van de timing van zelfevaluatie-activiteiten waarover we in de gevalstudies rapporteerden. Vandaar de wenselijkheid om te evolueren naar een schoolopdracht waarbinnen de mogelijkheid en de tijd geschapen wordt om zich vrij te maken voor en deskundig te worden in meer team- en schoolgebonden inhouden en competenties zoals reflectie en zelfevaluatie. Wij pleiten ervoor om van overleg, samenwerking en reflectie een structureel onderdeel van de opdracht van leerkrachten te maken. Dat creëert de ruimte om het belang van andere leerkrachtenactiviteiten te valoriseren.

Herhaaldelijk wordt het belang van coördinatie en opvolging in de verf gezet om tot kwaliteitsvolle zelfevaluaties te komen. De vraag voor veel scholen is echter hoe de tijdsbesteding door personeelsleden die een zelfevaluatie coördineren, verrekend kan worden. Omdat dergelijke taken vaak bovenop een bestaand takenpakket terecht komen, wordt één en ander op de lange baan geschoven en houden veel teamleden de boot af. Uit angst voor de planlast die met zelfevaluaties gepaard gaat, zijn potentiële kandidaten vaak zeer terughoudend om zich te engageren. Een noemenswaardig extra aan (financiële) middelen nodig is om aan de verwachtingen inzake zelfevaluatie te voldoen. Zonder deze middelen is er te weinig tijd beschikbaar en zijn er te veel administratieve of logistieke beslommeringen om de voordelen van zelfevaluaties aan den lijve te kunnen ondervinden (Schildkamp, 2007). Het is echter ook aan de school zelf om de beschikbare middelen voldoende flexibel te besteden. Dat betekent bijvoorbeeld dat zelfevaluaties op het meest rendabele moment uitgevoerd worden. Investeringen in professionele ontwikkeling en externe ondersteuning omtrent zelfevaluatie dienen in een omvattend plan te worden gesitueerd.

Scholen hebben/gebruiken weinig moderne instrumenten om een kwaliteitszorgbeleid te voeren (zoals de aanwezigheid van een kwaliteitszorgcoördinator, ontwikkelingsgerichte doorlichtingen, collegiale visitaties, mogelijkheden tot onderzoek). Het ontwikkelen van een middenkader wordt wel door veel scholen terecht als een belangrijke beleidskeuze naar voren geschoven. Doorgaans gebeurt dit door mensen deels of volledig vrij te stellen van een lesopdracht. Ongeacht de voordelen van zulk een aanpak, willen we benadrukken dat men moet bewaken dat deze keuze niet tot een nieuwe hiërarchie in scholen leidt. We stelden in deze studie immers een contrast vast tussen de beleidsmakers (schoolleiding en middenkader) die overtuigd en soms doordrongen zijn van recente benaderingen

van schoolbeleid, die het jargon spreken en die de lijnen uitzetten enerzijds en de teamleden waarvoor dit niet opgaat anderzijds.

Expliciteer de doelen van zelfevaluaties en breng ze in overeenstemming

Volgens de stakeholders in de Delphi-studie ontbreekt het in de meeste scholen nog aan visie omtrent de plaats van zelfevaluatie in het geheel van kwaliteitszorg. Scholen zien bijvoorbeeld nog vaak te weinig in dat zelfevaluatie niet los staat van de rest van de schoolwerking. De linken worden te weinig geëxpliciteerd. Dat geldt bij uitbreiding ook voor de doelen van zelfevaluatie in het algemeen (Clift e.a., 1989). De doelen werden, zoals blijkt uit de gevalstudie, bijvoorbeeld niet doelbewust 'op papier gezet'. Dat leidt tot onduidelijkheid en onzekerheid. Zelfevaluaties worden mede daardoor als eenmalige en losstaande initiatieven gepercipieerd. De kans dat scholen succesvol zijn in het beleid dat ze voeren, is echter afhankelijk van de mate waarin ze erin slagen om als gemeenschap te komen tot een gezamenlijke doelgerichtheid (Levine & Lezotte, 1990; Sammons e.a., 1995; Hooge, 1998). Daaronder verstaan we de mate waarin de school alle participanten weet te richten op de vooraf bepaalde doelen. Bij het uitvoeren van een zelfevaluatie is dat niet anders (Schildkamp, 2007).

De doelen van de verschillende stakeholders bij een zelfevaluatie zijn niet noodzakelijk in overeenstemming (Davies & Rudd, 2001). De prioriteiten van een leerkracht zijn niet altijd gelijklopend met deze van het schoolbestuur of de overheid, ook al hoeven ze elkaar niet uit te sluiten. Bij de start van een zelfevaluatie komt het er op aan om de verschillende stakeholders tot overeenstemming te brengen omtrent de functie en doelen van de zelfevaluatie (van Aanholt & Buis, 1990; Scheerens, 2004). Het is van groot belang dat de betrokkenen de doelen onderschrijven (SICI, 2003). We stelden in deze studie vast dat aan deze voorwaarde vaak niet vervuld is. De verwarring over de motieven om tot zelfevaluatie te komen, en over de plaats van die zelfevaluatie in het schooleigen kwaliteitszorgsysteem steekt in veel scholen de kop op. Men doet er goed aan bij de start van een zelfevaluatie op een gepaste (SMARTe) manier te expliciteren wat men wil bereiken. Dit bevordert de opvolging én de evaluatie van de zelfevaluaties. Wij volgen Kyriakides & Campbell (2004) in hun verwachting dat het expliciteren van doelen en in overeenstemming brengen van doelen, zal bijdragen tot het initiëren van beleidscycli waarvan zelfevaluaties inherent deel uitmaken.

Meta-evaluatie vergt breedte en diepgang én dus een kwantitatieve en kwalitatieve benadering

De onderzoeksactiviteiten waarover we in dit proefschrift rapporteerden, werpen elk op hun eigen manier een licht op het verloop en de kwaliteit van zelfevaluaties in scholen. We stelden vast dat kwantitatieve en kwalitatieve gegevens elkaar aanvullen om normatieve uitspraken over zelfevaluaties mogelijk te maken. Meta-evaluatie vergt immers breedte en diepgang. Met 'breedte' pleiten we voor het betrekken van verschillende en

voldoende betrokkenen bij de zelfevaluatie in de school. Bijvoorbeeld schriftelijke vragenlijsten kunnen uitmonden in een representatief beeld van hoe het proces en de resultaten van de zelfevaluatie in een school beleefd worden. Het maakt vergelijking met een referentiegroep mogelijk. 'Diepgang' verwijst naar de noodzaak om op zoek te gaan naar verklaringen voor de zienswijzen van respondenten en naar het specifieke karakter van zelfevaluaties in verschillende scholen. Zelfevaluaties die gelijkaardig scoren uitgaande van een schriftelijke vragenlijst kunnen door de betrokkenen zeer verschillend beleefd worden. Vandaar ook ons pleidooi om de context in een meta-evaluatie te betrekken. Om in kaart te brengen in welke mate zelfevaluaties de vooropgestelde verwachtingen realiseren, doet men er goed aan de zelfevaluatie op een holistische en gecontextualiseerde manier te benaderen. Meta-evaluatie kan in dit opzicht beschouwd worden als een vorm van multimethodeonderzoek. De ene werkwijze compenseert dan de tekorten van een andere werkwijze. De verschillende onderzoeksmethoden bestuderen dan hetzelfde fenomeen zodat hun bevindingen in de analyse geïntegreerd kunnen worden. In de praktijk betekent dit bijvoorbeeld dat de bevindingen van een vragenlijst dan kunnen getoetst worden aan de resultaten van contextinclusieve interviews, focusgroepen en documentanalyses. Meta-evaluatie vergt een multimethode én multisubject aanpak.

Het versterken van de kwaliteit van zelfevaluaties vergt interne professionalisering en externe ondersteuning

Dit onderzoek toont aan dat het uitvoeren van zelfevaluaties zowel beleidsmatige als onderzoeksmatige verwachtingen stelt aan scholen. Een zelfevaluatie vergt met andere woorden kennis, vaardigheden en attitudes van scholen die momenteel nog niet verondersteld mogen worden maar doelbewust moeten ontwikkeld worden (Visscher, 2002). Scholen, hun ondersteuners en hun overheden mogen daarom niet voorbijgaan aan het belang van interne professionalisering en externe ondersteuning. Wij zijn van mening dat beide niet alleen noodzakelijk zijn maar elkaar ook nog eens nodig hebben.

Onder *interne professionalisering* verstaan we het opzetten van initiatieven door de school zelf om de kennis, vaardigheden en attitudes van schoolleiding en teamleden te ontwikkelen. Daarbij denken we zowel aan formele en informele als aan individuele en gezamenlijke activiteiten. Met de nadruk op interne professionalisering willen we beklemtonen dat scholen in de eerste plaats zelf verantwoordelijk zijn om zich zodanig te ontwikkelen dat kwaliteitsvolle zelfevaluaties kunnen gegarandeerd worden. Daarbij denken we bijvoorbeeld aan de centrale rol van scholen in het inventariseren van de eigen bekwaamheden en tekorten wat het uitvoeren van zelfevaluaties betreft. Reflectie op (initiatieven tot) zelfevaluatie moet aan het licht brengen over welke kwaliteiten, vaardigheden en visies individuen en teams al dan niet beschikken (Schildkamp, 2007). Als scholen zichzelf willen professionaliseren in het uitvoeren van zelfevaluaties zullen zij zich dus moeten engageren in meta-evaluatie. Om werk te maken van de actiepunten die daaruit voortvloeien kunnen zij een beroep doen op talrijke strategieën en methodieken waartoe zij zelf het initiatief kunnen nemen en die zij zelf in goede banen kunnen leiden. Daarbij denken wij aan het stimuleren van het lezen van vakliteratuur en het bijwonen

van nascholingen, het expliciteren van eigen overtuigingen en het rapporteren over eigen ervaringen, het bespreken van bedenkingen en succesfactoren tijdens intervisies of supervisies, het observeren en coachen van anderen. Het komt er in grote lijnen op neer dat men de expertise die in de school aanwezig is zichtbaar maakt en ter beschikking stelt van anderen.

De verwachting dat alle scholen over het ontwikkelingsniveau beschikken om volledig autonoom hun zelfevaluaties op een kwaliteitsvolle manier vorm te geven, is evenwel niet realistisch (Davies & Rudd, 2001). Dat impliceert dat ook aanzienlijk dient geïnvesteerd te worden in de *externe ondersteuning* van scholen inzake zelfevaluatie. Ook expertise die binnen de school niet aanwezig is, moet in het ontwikkelingstraject opgenomen kunnen worden. Betrokkenheid van externen heeft bovendien als voordeel dat sterktes en zwaktes aan het licht kunnen worden gebracht waarvan de school zichzelf niet bewust is. Deze studie heeft duidelijk aangetoond dat scholen baat hebben bij goede externe ondersteuning. In het kader van zelfevaluatie pleiten wij voor intense ondersteuningsrelaties die bij voorkeur contactmomenten binnen de school omvatten. Dat geeft de kans om in te spelen op de eigenheid van en de interne professionaliseringsinitiatieven in de school. De ondersteuning waaraan scholen behoefte hebben kan variëren van het ter beschikking stellen van conceptuele raamwerken, het aanleveren van gebruiksklare instrumentaria en stappenplannen voor het komen tot gepaste actiepunten tot training en ondersteuning ter plaatse. Ondersteuning omvat best het gehele proces van de zelfevaluatie in ruime zin. Wij pleiten daarom voor een zo goed mogelijke integratie van de ondersteuning van de zelfevaluatie in de ruimere systeembenadering van de (pegagogische) begeleiding waarop scholen een beroep kunnen doen.

Uit het bovenstaande kan geconcludeerd worden dat interne professionalisering en externe ondersteuning een geïntegreerde rol behoren te vervullen in het versterken van de kwaliteit van zelfevaluaties. Beide blijken elkaar dienstig te kunnen zijn om binnen de school tot kenniscreatie te komen (cf. de vier vormen van kennisconversie (Nonaka & Takeuchi, 1995)). De meest gepaste verhouding tussen interne professionalisering en externe ondersteuning bestaat dus in ieder geval niet uit een optelsom van beide afzonderlijke aspecten. Het ziet er naar uit dat denken in termen van vermenigvuldiging aan de orde is. Dat betekent zowel dat beide aspecten elkaar kunnen versterken als dat bij afwezigheid van één van beide de ander aan waarde inboet. Externe ondersteuning die bijvoorbeeld niet aansluit bij de interne professionalisering in scholen verliest aan effect.

4 REFLECTIES BIJ HET ONDERZOEK EN SUGGESTIES VOOR TOEKOMSTIG ONDERZOEK

Dit onderzoek startte vanuit de vaststelling dat de verwachtingen omtrent de mogelijkheden van zelfevaluatie om scholen tot ontwikkeling te brengen groot zijn. Allerhande beleidsteksten – van overheden maar ook van scholen – zijn doordrongen van retoriek omtrent het waardevolle van zelfevaluaties in scholen. Er zijn veel claims waarvoor de

wetenschappelijke onderbouw tot voor kort in grote mate onbrak. De waarde van deze studie ligt in de systematische en veelzijdige benadering van zelfevaluaties in scholen. Het 'vernieuwende' karakter van deze studie bestaat in eerste instantie uit het toetsen van inzichten die uit eerder onderzoek groeiden en in het verbreden en verdiepen van de bestaande kennisbasis.

Het eerdere onderzoek in het veld van zelfevaluatie door scholen leerde bijvoorbeeld dat het realiseren van de bestaande verwachtingen ten aanzien van zelfevaluatie voor scholen geen vanzelfsprekendheid is (Visscher, 2002; Devos & Verhoeven, 2003; Blok e.a., 2005). Doorgaans is er sprake van beperkte effecten in de scholen die zelfevaluatie-instrumenten gebruikten. De gevonden effecten hebben voornamelijk betrekking op de condities om tot schoolontwikkeling te komen en niet op de prestaties van leerlingen als dusdanig (Schildkamp, 2007). Deze studie bevestigt dat zelfevaluaties in vele scholen nog niet in sterke mate tot de vooropgestelde verwachtingen bijdragen. We hebben daarbij evenwel beargumenteerd dat het gebruikte referentiekader doorslaggevend is wanneer men tot beoordelende uitspraken over de kwaliteit van zelfevaluaties in Vlaamse scholen wil komen. De verwachtingen van verschillende betrokkenen werden in kaart gebracht en we toetsten – zowel kwalitatief als kwantitatief – in welke mate aan deze verwachtingen tegemoet werd gekomen.

Ook omtrent de vraag welke kenmerken van het verloop van zelfevaluatieprocessen en van de scholen waarin de zelfevaluatie plaatsvinden van invloed zijn op de kwaliteit van zelfevaluaties, was de empirische evidentie schaars. De bestaande evidentie werd samengebracht in het theoretisch kader. Deze evidentie wordt bevestigd en uitgebreid in deze studie. Onder meer de cruciale rol van reflectie- en innovatiebereidheid in scholen, van gezamenlijke doelgerichtheid en van gedeeld leiderschap (zie onder meer (McBeath & Myers, 2002; Devos & Verhoeven, 2003; Schildkamp, 2007)) wordt door de bevindingen van deze studie ondersteund. Deze relaties waren voorheen (in Vlaanderen) nog niet op dergelijk grote schaal in kaart gebracht en bijgevolg ook nog niet aan hedendaagse statistische analysetechnieken onderworpen. De relaties die in het kwantitatieve luik van belang bleken, werden in het kwalitatieve luik verdiept. Daarbij denken we bijvoorbeeld aan de manier waarop de dragers van beleidsvoerend vermogen een rol spelen in de manier waarop zelfevaluaties in scholen vorm krijgen, beleefd worden en al dan niet tot resultaten leiden. Zonder de beperkingen van deze studie inzake representativiteit te willen verbloemen, kunnen we stellen dat de resultaten een beeld geven van de stand van zaken omtrent het uitvoeren van zelfevaluaties in Vlaamse scholen.

Deze realisaties zijn in grote mate een gevolg van onze resolute keuze om zelfevaluaties aan de hand van een multimethodebenadering te bestuderen. Toch is ook een reflectie op de methodologie van deze studie aangewezen. De keuze voor een bepaald onderzoeksdesign gaat immers telkens gepaard met specifieke voor- en nadelen. In eerste instantie willen we benadrukken dat de vooropgestelde meerwaarde van het combineren van verschillende onderzoeksactiviteiten ons inziens gerealiseerd werd. In de argumentatie van de opzet van deze studie was sprake van een verkennende en verklarende relatie enerzijds en datatriangulatie anderzijds. Deze opbouw heeft in belangrijke mate bijge-

dragen tot de validiteit en de bewijskracht van de onderzoeksbevindingen. Toch dringen zich ook een aantal methodologische bedenkingen op. Aan deze reflecties knopen we meteen suggesties voor toekomstig onderzoek vast. Ons inziens zijn de volgende methodologische elementen het meest fundamenteel.

- *Het concurrente karakter van de gegevensverzameling in 'Zelfevaluatie Binnenstebuiten'.* De gegevens in dit luik van de kwantitatieve analyses werden noodgedwongen op één moment verzameld. Het is echter aan te bevelen om informatie omtrent de houding ten aanzien van zelfevaluatie en omtrent de schoolgebonden kenmerken te verzamelen alvorens het zelfevaluatieproces in scholen opgestart wordt. Ook de vraag naar de resultaten van de zelfevaluatie zou beter in de tijd losgekoppeld worden van de vraag naar het verloop van de zelfevaluatie. Het opzetten van longitudinaal onderzoek naar het verloop en de kwaliteit van zelfevaluatie zou dan ook tot meer betrouwbare en valide gegevens kunnen bijdragen. De gefaseerde gegevensverzameling tijdens de gevalstudies is in die zin inspirerend voor toekomstig surveyonderzoek. Een longitudinale benadering maakt het naast het vaststellen van verbanden ook makkelijker om uitspraken over causaliteit te doen.
- *Het gebrek aan een 'harde' output maat*. In het kader van deze studie konden wij onmogelijk de link kunnen leggen met wat nochtans het finale doel van zelfevaluatie dient te zijn, namelijk de prestaties van leerlingen. In bijvoorbeeld een longitudinale aanpak kan ook onderzocht worden welke impact de kwaliteit van zelfevaluaties heeft op leerlingenprestaties.
- *De afhankelijkheid van percepties.* Uitgezonderd de informatie over de GOK-inspecties zijn gegevens in deze studie verkregen door zelfrapportage bij betrokkenen. Cuyvers (2002, p. 26) maakt in de context van kwaliteitszorg evenwel terecht een onderscheid tussen een verwachte kwaliteit, een beleefde kwaliteit en een feitelijke kwaliteit. Om een volledig beeld te krijgen van de kwaliteit van zelfevaluaties is er nood aan onderzoek dat de beleefde kwaliteit koppelt aan de verwachte en feitelijke kwaliteit. Het loont de moeite om te onderzoeken in welke mate een inspecterende overheid, onderzoekers, schoolleiders en leerkrachten tot gelijkaardige oordelen komen wanneer zij vanuit een gezamenlijk referentiekader uitspraken doen over de kwaliteit van eenzelfde zelfevaluatieproces. Vervolgonderzoek kan bijvoorbeeld een beroep doen op focusgroepsgesprekken waarin op zoek gegaan wordt naar verklaringen voor de verschillende percepties van schoolleiders en teamleden van de zelfevaluatie die in hun school plaats vond. Door deels af te stappen van zelfrapportage kunnen ook betere garanties voor het vermijden van sociaal wenselijk (antwoord)gedrag ingebouwd worden.
- *De beperkte generaliseerbaarheid van de bevindingen.* In wezen werd enkel in het geval van de GOK-zelfevaluatie gewerkt met een representatief staal van Vlaamse scholen. De gegevensverzameling in het kader van 'Zelfevaluatie binnenstebuiten' beperkte zich tot scholen die reeds een zelfevaluatieproces achter de rug hadden en die bereid waren aan het onderzoek deel te nemen. Representatief onderzoek dringt zich dus op. Daarbij denken we aan de nood aan een representatief beeld omtrent de

houding ten aanzien van zelfevaluatie in Vlaamse scholen, omtrent de initiatieven inzake (geïnstrumentaliseerde) zelfevaluatie die men momenteel neemt en omtrent de effecten daarvan.

- *De wederzijdse exclusiviteit van deelname aan de survey 'zelfevaluatie binnenste-buiten' en aan de gevalstudies*. De survey richtte zich op scholen die reeds een zelfevaluatieproces achter de rug hadden. De gevalstudie richtte zich op scholen tijdens het uitvoeren van een zelfevaluatie. De duurtijd van deze studie liet ons niet toe dezelfde scholen op beide manieren te bestuderen. Met het oog op het verfijnen van de multimethode en multisubject benadering van meta-evaluatie is onderzoek dat dit wel mogelijk maakt aangewezen.
- *De tijdsgebondenheid*. Tenslotte willen we nog benadrukken dat vervolgonderzoek ook kan bestaan uit een replicatie van wat in deze studie gebeurde. We stelden vast dat zelfevaluatie voor veel scholen nog kan beschouwd worden als een vernieuwing. Dat maakt het bijzonder interessant om na te gaan of er zich de komende 5 à 10 jaar een evolutie zal voordoen en of dezelfde variabelen een rol zullen blijven spelen.

Zowel vanuit theoretische als vanuit methodologische hoek kunnen verscheidene waardevolle suggesties worden gedaan om de kennisbasis omtrent zelfevaluaties in (Vlaamse) scholen verder uit te breiden. Over zelfevaluaties in scholen is het laatste woord nog niet gezegd.

5 AFRONDEND

Thales van Milete stelde ooit dat 'ken uzelf' de moeilijkste vraag is die je iemand kan voorleggen. Het advies 'Ken uzelf' stond in de Griekse oudheid gebeiteld boven de ingang van het orakel van Delphi. Velen kwamen de raad van dit orakel inroepen in moeilijke, onrustige omstandigheden en onzeker van zichzelf en hun toekomst. Voor scholen lijkt 'Ken uzelf' momenteel vervangen door het modieuzere 'zelfevaluatie'. Maar verder...

LITERATUUR

Adler, M., & Ziglio, E. (1996). *Gazing into the oracle: The Delphi method and its application to social policy and public health*. London: Jessica Kingsley Publishers.

Airasian, P. W. (1994). *Classroom assessment*. New York: McGraw-Hill.

Argyris, C., & Schön, D. A. (1978). *Organizational learning: a theory of action perspective*. Reading: Addison-Wesley.

Argyris, C., & Schön, D. A. (1996). *Organizational learning II: Theory, method and practice*. Reading, Mass: Addison Wesley.

Arter, J. A., & Bond, L. (1996). Why is assessment changing? In R. E. Blum & J. A. Arter (Eds.), *Student Performance assessment in an Era of Restructuring*. Alexandria, VA: Association for Supervision and Curriculum.

Arts, J., Kok, J., Sleegers, P., Verbiest, E., & de Wit, C. (2003). *Speelbal of spelbepaler, Over professionele schoolontwikkeling en kwaliteit*. Den Haag: Q*Primair.

Baarda, D. B., de Goede, M. P. M., & Teunissen, J. (2001). *Basisboek Kwalitatief onderzoek. Praktische handleiding voor het opzetten en uitvoeren van kwalitatief onderzoek*. Groningen: Stenfert Kroese.

Babchuck, W. (1997). *Glaser or Strauss? Grounded theory and adult education*. (Internetbron geraadpleegd op 22 november 2005 via: http://www.anrecs.msu.edu/research/gradpr96.htm).

Bakx, E., & Van der Eerden, E. (1993). *Onderzoek naar het innovatief vermogen van scholen in het voortgezet onderwijs*. Nijmegen: KUN.

Ballet, K., Kelchtermans, G., Berens, G., & Janssens, S. (2000). Professionele belangen in de beroepssocialisatie. Micropolitiek in de loopbaanverhalen van beginnende leerkrachten Basisonderwijs. *Pedagogische Studiën, 77*(2), 67-84.

Barnett, K., McCormick, J., & Conners, R. (2001). Transformational leadership: panacea, placebo or problem? *Journal of Educational Administration, 39*, 24-46.

Bergman, M. M. (2007). Multimethod Research and Mixed Methods Research: Old Wine in New Bottles? *Journal of Mixed Methods Research, 1*(1), 101-102.

Birenbaum, M. (1995). Assessment 2000: Towards a pluralistic approach to assessment. In M. Birenbaum & F. Dochy (Eds.), *Alternatives in assessments of achievements, learning processes and prior knowledge*. Dordrecht: Kluwer Academic Publishers.

Blanquaert, A., Maes, B., & Ver Eecke, E. I. (2001). De relatie tussen zelfevaluatie en externe evaluatie. Evenwichtsoefening tussen verantwoording en schoolverbetering. *Impuls voor onderwijsbegeleiding, 32*(1), 11-16.

Blok, H., Sleegers, P., & Karsten, S. (2005). *Schoolzelfevaluatie in het basisonderwijs; een terugblik op de zelfevaluatiefase van Ziezo*. Amsterdam: SCO-Kohnstamm Instituut.

Boekaerts, M., & Simons, P. R. J. (1995). *Leren en instructie; psychologie van de leerling en het leerproces*. Assen: Van Gorcum.

Bolhuis, S., & Simons, P. R. J. (2001). Naar een breder begrip van leren. In J. W. M. Kessels & R. F. Poell (Eds.), *Human resource development: organiseren van het leren* (pp. 37-51). Houten: Samsom.

Bolhuis, S. M. (2000). *Naar zelfstandig leren: wat doen en denken docenten?* Leuven - Apeldoorn: Garant.

Bolhuis, S. M., & Simons, P. R. J. (1999). *Leren en werken*. Deventer: Kluwer.

Bomers, G. B. J. (1990). De lerende organisatie: de enige zekerheid voor organisaties is permanente verandering. *Harvard Holland Review, 22*, 21-31.

Bosker, R. J., & Witziers, B. (1996). *The magnitude of school effects, or: Does it really matter which school a student attends?* Paper gepresenteerd op 'annual meeting of the American Educational Research Association', New York.

Braster, J. F. A. (2000). *De kern van Casestudy's*. Assen: Van Gorcum.

Brewer, J., & Hunter, A. (1989). *Multimethod research: A synthesis of styles*. Thousand Oaks, CA: Sage.

Bright, B. (1996). Reflecting on 'reflective practice'. *Studies in the education of adults, 28*(2), 162-184.

Bunt, A., & Van Hees, T. (1991). Zelfdiagnose door scholen. *Meso Magazine* (54), 13-17.

Byrne, B. M. (2001). *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Campbell, J. P., Bownas, E. A., Peterson, N. G., & Dunnette, M. D. (1974). *The measurement of organizational effectiveness: A review of relevant research and opinion*. San Diego: Naval Personnel Research and Development Center.

Cheng, Y. (1996). *The pursuit of school effectiveness*. Hong Kong: Hong Kong Institute of Educational Research.

Chou, S.-W., & Tsai, Y.-H. (2004). Knowledge Creation: Individual and Organizational Perspectives. *Journal of Information Science, 30*(3), 205-218.

Christie, C. A., Ross, R. M., & Klein, B. M. (2004). Moving toward collaboration by creating a participatory internal-external evaluation team: A case study. *Studies in Educational Evaluation, 30*, 125-134.

Clift, P. S., Nuttall, D. L., & McCormick, R. (Eds.). (1989). *Studies in school self evaluation*. London: The Falmer Press.

Cobb, P. (1996). Constructivism and learning. In E. De Corte & F. E. Weinert (Eds.), *International encyclopedia of developmental and instructional psychology* (pp. 338-341). Oxford: Pergamon.

Coe, R., & Visscher, A. J. (2002). Introduction. In A. J. Visscher & R. Coe (Eds.), *School Improvement through Performance Feedback* (pp. xi-xix). Lisse: Swets & Zeitlinger.

Conzemius, A., & O'Neill, J. (2002). *The handbook for SMART school teams*. Bloomington, IN: National Educational Service.

Cousins, J., & Earl, L. (1995). *Participatory evaluation in education: Studies in evaluation use and organizational learning*. London: Falmer.

Cremers-Van Wees, L. M. C. M., Rekveld, I. J., Brandsma, H. P., & Bosker, R. J. (1996). *Instrumenten voor zelfevaluatie: inventarisatie en beschrijving*. Enschede: Universiteit Twente, Onderzoek Centrum Toegepaste Onderwijskunde.

Creswell, J. W. (2005). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. New Jersey: Pearson Prentice Hall.

Creswell, J. W., Clark, V., Gutmann, M., & Hanson, W. (2003). Advanced Mixed Methods Research Designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of Mixed methods in social & behavioural research* (pp. 209-240). Thousand Oaks: Sage.

Cuyvers, G. (2002). *Kwaliteitsontwikkeling in het onderwijs*. Antwerpen - Apeldoorn: Garant.

Davies, D., & Rudd, P. (2001). *Evaluating school self-evaluation*. Berkshire: National Foundation for Educational Research.

De Klerk, L. F. W., & Simons, P. R. J. (1988). Diverse opvattingen over het leren. In *Onderwijskundig Lexicon* (pp. A 1103-1118). Alphen aan den Rijn: Samsom.

De Laat, M., Poell, R., Simons, R. J., & van der Krogt, F. (2001). Organiseren van leren op de werkplek. In A. Wald & J. van der Linden (Eds.), *Leren in perspectief* (pp. 69-84). Leuven / Apeldoorn: Garant.

De Maeyer, S., & Rymenans, R. (2004). *Onderzoek naar kenmerken van effectieve scholen*. Gent: Academia Press.

De Maeyer, S., & Vanhoof, J. (2005). De vijfpuntenschaal in onderwijsonderzoek: de regel, maar ook in regel? In M. Valcke (Ed.), *Meten en onderwijskundig onderzoek* (pp. 458-461). Gent: Academia Press.

De Wolff, A. (2003). De taal van Kwaliteitszorg. Hebben we het wel over hetzelfde? *De Q-vijver, 4*(5), 17-19.

DeFleur, M. L., Kearny, P., & Plax, T. G. (1993). *Mastering communication in contemporary America*. Mountain View, CA: Mayfield Publishing Company.

Denzin, N. K., & Lincoln, Y. S. (2000). *Handbook of qualitative research*. Thousand Oaks: Sage.

Devos, G. (2000). *Schoolmanagement. Een reflectie op de praktijk van de schoolleider*. Diegem: Kluwer.

Devos, G., & Verhoeven, J. (2000). Het proces van zelfevaluatie. Een verkennend onderzoek in secundaire scholen. *T.O.R.B., 6*, 425-436.

Devos, G., & Verhoeven, J. (2003). School self-evaluation - conditions and caveats - The case of secondary schools. *Educational Management & Administration, 31*(4), 403 -420.

Devos, G., Verhoeven, J., Opbrouck, W., & Verbeeck, B. (2000). *Handboek zelfevaluatie in het secundair onderwijs*. Diegem: Kluwer.

Devos, G., Verhoeven, J., Van den Broeck, H., Gheysen, A., Opbrouck, W., & Verbeeck, B. (1999). *Interne zelfevaluatie-instrumenten voor kwaliteitszorg in het Secundair Onderwijs*. Leuven: KUL: Departement Sociologie.

Dodgson, M. (1993). Organizational learning: A review of some literatures. *Organization Studies, 14*(3), 375-394.

Donche, V. (2005). *Leren, onderwijzen en leren onderwijzen: onderzoek naar opvattingen en handelingen van studenten en docenten*. Gent: Academia Press.

Donnenberg, O. H. J. (1999). Nieuwe werkstructuren - nieuwe leervormen. In J. M. J. Dekker (Ed.), *Opleiders in organisaties, Netwerkend leren* (pp. 116-137). Deventer: Kluwer.

Duurkoop, M., Majoor, D., & de Pinth, M. (2002). Kwaliteitszorg, inventarisatie en beschrijving van instrumenten.

Eraut, M. R. (2000). Non-formal learning and tacit knowledge in professional work. *British Journal of Educational Psychology, 70*, 113-136.

Erzberger, C., & Kelle, U. (2003). Making inferences in mixed methods: the rules of integration. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioural research* (pp. 457-488). Thousand Oaks: Sage.

Forss, K., Cracknell, B., & Samset, K. (1994). Can Evaluation Help an Organisation to Learn? *Evaluation Review, 18*(5), 574-592.

Friedman, V. J., Lipshitz, R., & Overmeer, W. (2001). Creating conditions for organizational learning. In M. Dierkes, A. B. Antal, J. Child & I. Nonaka (Eds.), *The handbook of organizational learning* (pp. 757-774): Oxford University Press.

Fullan, M. (1991). *The New Meaning of Educational Change*. London: Cassell.

Geijsel, F. (2001). *Schools and innovations. Conditions fostering the implementation of educational innovation*. Nijmegen: University Press.

Geijsel, F., & Sleegers, P. (2001). Aangrijpingspunten voor innovatie: schoolorganisatie en schoolleiding. In B. Creemers & A. Houtveen (Eds.), *Onderwijsinnovatie. Onderwijskundig lexicon, editie III* (pp. 57-74). Alphen aan de Rijn: Kluwer.

Geijsel, F., Sleegers, P., Van den Berg, R., & Kelchtermans, G. (2001). Conditions fostering the implementation of large-scale innovation programs in schools: teachers' perspectives. *Educational Administration Quarterly, 37*(1), 130-166.

Geijsel, F., van den Berg, R. M., & Sleegers, P. (1998). The innovative capacity of schools in primary education: a qualitative study. *International Journal of Qualitative Studies in Education, 12*(2), 175-192.

Gemmeke, M., & van der Vegt, A. L. (1997). *Professionalisering van de schoolorganisatie in het primair en voortgezet onderwijs*. Amsterdam: Regioplan.

Giesbers, J., & Sleegers, P. (1994). The Marx models as conceptual models in school effectiveness. *School Organisation, 14*(1), 91-102.

Glaser, B. G. (1992). *Basics of grounded theory analysis*. Mill Valley, CA: Sociology Press.

Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of Grounded Theory. Strategies for Qualitative Research*. Chicago: Aldine Publishing Company.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48*, 26-34.

Gombeir, D. (1998). Zelfevaluatie als schakel in het proces van schoolbeleidsvoering. *Schoolleiding en -begeleiding (Deel Zelfevaluatie, Algemeen)*, 22, 51-94.

Goulding, C. (1999). *Grounded theory: some reflections on paradigm, procedures and mis-conceptions*. Wolverhampton: University of Wolverhampton.

Green, S. B., & Salkind, N. J. (2004). *Using SPSS, for Windows and Macintosh, analyzing and understanding data*. New Jersey: Pearson Education.

Greenwood, J. (1998). The role of reflection in single and double loop learning. *Journal of Advanced Nursing, 27*(5), 1048-1053.

Griffith, J. (2003). Schools as Organizational Models: Implications for Examining School effectiveness. *The elementary School Journal, 104*(1), 29-47.

Güthe, K. A. K. (1997). *Indicatoren van schoolontwikkeling*. Enschede: OCTO, Universiteit Twente.

Hanson, E. M. (1979). *Educational Administration and Organizational Behaviour*. Boston: Allyn en Bacon.

Hargreaves, A. (1992). Cultures of teaching: A focus for change. In A. Hargreaves & M. Fullan (Eds.), *Understanding teacher development* (pp. 216-240). New York: Teachers College Press.

Hargreaves, A., & Fullan, M. (1992). *Understanding teacher development*. New York: Teachers College Press.

Hargreaves, A., & Fullan, M. (1998). *What's worth fighting for out there?* New York: Teachers College Press.

Harrington, H. J., & Harrington, J. S. (1994). *Total Improvement Management: The Next Generation in Performance Improvement*. New York: McGraw-Hill.

Harris, A. (2002). *School Improvement: what's in it for schools?* London: Routeledge Falmer.

Harris, A. (2003). Teacher Leadership as Distributed Leadership: heresy, fantasy or possibility? *School Leadership & Management, 23*(3), 313-325.

Heck, R. H., & Marcoulides, G. A. (1996). Organizational Culture and Performance: The Invariance of an Organizational Model. *School Effectiveness & School Improvement, 7*(1), 1-20.

Hendriks, M., & Bosker, R. (2003). *ZEBO instrument voor zelfevaluatie in het basisonderwijs. Handleiding bij een geautomatiseerd hulpmiddel voor kwaliteitszorg in basisscholen.* Enschede: Twente University Press.

Hendriks, M. A. (2000). *Kwaliteitszorg voortgezet onderwijs. Instrumenten en organisaties*. Utrecht: VVO/Q5, project kwaliteitszorg voortgezet onderwijs.

Hendriks, M. A., Doolaard, S., & Bosker, R. J. (2002). Using school effectiveness as a knowledge base for self-evaluation in Dutch schools: The ZEBO-project. In A. visscher & R. Coe (Eds.), *School improvement through performance feedback* (pp. 114-142). Lisse: Swets & Zeitlinger.

Herman, L., Aschbacher, R., & Winters, L. (1992). *A practical Guide to Alternative Assessment*. Alexandria Virginia: Association for Supervision and Curriculum Development.

Hofman, R. H., Dijkstra, N. J., de Boom, J., & Hofman, W. H. A. (2004). *Kwaliteitszorg in het primair onderwijs. Deelstudie 1: kwaliteitszorg in het primair onderwijs; peiling 2003/2004.* Groningen: GION.

Hooge, E. (1998). *Ruimte voor beleid. Autonomievergroting en beleidsuitvoering door basisscholen.* Amsterdam: SCO-Kohnstamm Instituut.

Hord, S. M. (1997). *Professional learning communities: Communities of continuous inquiry and improvement*. Austin: Southwest Educational Development Laboratory.

Hox, J. (2002). *Multilevel analysis. Techniques and applications*. Mahwah, New Jersey/London: Lawrence Erlbaum Associates Publishers.

Hoy, W., & Miskel, C. (2001). *Educational Administration: Theory, Research and Practice*. Boston: McGraw-Hill.

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Hulpia, H. (2004). The use of performance indicators in a school improvement policy. What do experts say about it? *AERA Annual conference*. San Diego.

Hutjes, J. M., & Van Buuren, J. A. (1992). *De gevalstudie. Strategie van kwalitatief onderzoek*. Meppel: Boom.

Janssens, J. M. A. M. (1991). *'Ogen' doen onderzoek. Een inleiding in de methoden van sociaal-wetenschappelijk onderzoek*. Amsterdam/Lisse: Swets & Zeitlinger.

Jeris, L. (2003). Double-loop Learning in Teams. *Lifelong Learning in Europe, 8*(1), 8-17.

Johnson, B., & Turner, L. A. (2003). Data-collection strategies in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioural research* (pp. 297-320). Thousand Oaks: Sage.

Jongmans, C. T., Beijaard, D., & Biemans, H. J. A. (1996). Betrokkenheid van leraren bij schoolbeleid en effectiviteit van scholingsbeleid. *Pedagogisch Tijdschrift, 21*(3), 211-224.

Jorna, R. J. (2004). Kennis als basis voor innovatie: management en creatie. In R. J. Jorna, J. M. L. Engelen & H. Hadders (Eds.), *Duurzame innovatie. Organisaties en de dynamiek van kenniscreatie*. Assen: Van Gorcum.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141-151.

Kelchtermans, G. (1994). *De professionele ontwikkeling van leerkrachten basisonderwijs vanuit het biografisch perspectief*. Leuven: Universitaire Pers Leuven.

Kessler, R. (1992). Shared Decision Making Works! *Educational Leadership, 50*(1), 36-38.

Kimball, S. M. (2002). Analysis of Feedback, Enabling Conditions and Fairness Perceptions of Teachers in Three School Districts with New Standards-Based Evaluation Systems. *Journal of Personnel Evaluation in Education, 16*(4), 241-268.

Kolb, D. (1984). *Experiential learning*. Englewood Cliffs: Prentice Hall.

Kolb, D. A. (1993). The process of experiential learning. In M. Thorpe, R. Edwards & A. Hanson (Eds.), *Culture and process of adult learning*. London.

Koster, E., & Bouman, W. (1997). De 'Balanced Change Card': Een raamwerk voor het vormgeven en evalueren van veranderingsorganisaties. *Mens en Onderneming, 51*(5), 7-21.

Kupek, E. (2005). Log-linear transformation of binary variables: A suitable input for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 12*, 35-47.

Kwakman, C. H. E. (1999). *Leren van docenten tijdens de beroepsloopbaan: Studies naar professionaliteit op de werkplek in het voortgezet onderwijs*. Nijmegen: Katholieke Universiteit Nijmegen.

Kyriakides, L., & Campbell, R. J. (2004). School self-evaluation and school improvement: A critique of values and procedures. *Studies In Educational Evaluation, 30*(1), 23-36.

Lagerweij, N., Lagerweij, J., & Voogt, J. (2004). *Anders kijken. De dynamiek van een eeuw onderwijsverandering*. Antwerpen: Garant.

Lagerweij, N., & Voogt, J. (1997). Schoolontwikkeling: meer dan het invoeren van vernieuwingen. In T. Bergen, A. Knoers & P. Sleegers (Eds.), *Perspectieven op de school in dynamische ontwikkeling*. Alphen aan den Rijn: Samsom HD Tjeenk Willink.

LeCompte, M. D. (1993). A Framework for Hearing Silence: What Does Telling Stories Mean When We Are Supposed to Be Doing Science? In D. McLaughlin & W. Tierney (Eds.), *Naming Silenced Lives: Personal Narratives and the Process of Educational Change* (pp. 9-27). New York, NY: Routledge.

Leenheer, P. (2002). Networks of secondary schools in the Netherlands: Learning communities in the context of innovations. In R. Schollaert (Ed.), *In search of the treasure within. Towards schools as learning organisations*. Antwerpen: Garant.

Leithwood, K., Aitken, R., & Jantzi, D. (2001). *Making schools smarter: A system for monitoring school and district progress*. Thousand Oaks, CA: Corwin Press.

Leithwood, K., & Duke, D. (1998). Mapping the conceptual terrain of leadership: A critical point of departure for cross cultural studies. *Peabody Journal of Education, 73*(2), 31-50.

Levine, D. U., & Lezotte, L. W. (1990). *Unusually Effective Schools: A Review and Analysis of Research and Practice*. Madison: National Centre for Effective Schools Research and Development.

Liket, T. M. E. (1992). *Vrijheid & rekenschap: zelfevaluatie en externe evaluatie in het voortgezet onderwijs*. Amsterdam: Meulenhoff Educatief.

Linstone, H. A., & Turoff, M. (Eds.). (1975). *The Delphi method: Techniques and applications*. Cambridge: Addison-Wasley.

Little, J. (1990a). The persistence of privacy. Autonomy and initiative in teachers' professional relations. *Teachers College Record, 91*(4), 504-536.

Little, J. W. (1990b). Teachers as colleagues. In A. Liebermann (Ed.), *Schools as collaborative cultures: creating the future now.* (pp. 165-193). New York/Philadelphia/London: The Falmer Press.

Love, A. (1991). *Internal Evaluation: Building Organizations from Within*. London: Sage.

Luke, D. A. (2004). *Multilevel Modelling*. Thousand Oaks: Sage Publications.

Mace, J. (Ed.). (2002). *Self-Assessment - Take a look at yourself*. Londen: Learning and Skills Development Agency.

Maes, B., Ver Eecke, E., & Zaman, M. (2004). Gezocht: vruchtbare relatie tussen interne en externe evaluatie. *Impuls voor onderwijsbegeleiding, 34*(4), 215-237.

Mahieu, P. (1999). De school in de wereld; verantwoordingsstrategieën voor scholen. In P. Mahieu, C. Dietvorst & P. A. Peene (Eds.), *Organisatiecultuur van een extraverte school; naar verantwoordelijk onderwijs*. Alphen a/d. Rijn: Samson.

Majoor, D. (2000). *Vooruitgang in autonomie. Een studie naar organisatorische gevolgen van financiële en personele beleidsbenutting in het basisonderwijs*. Nijmegen: Iken.

Marlow, M., Kyed, S., & Connors, S. (2005). Collegiality, collaboration and Kuleana: complexity in a professional development school. *Education, 125*(4), 557 - 589.

Maslowski, R. (2001). *School culture and school Performance*. Enschede: Twente University Press.

McBeath, J. (1999). *Schools must speak for themselves*. London: Routledge.
McBeath, J., Meuret, D., Schratz, M., & Jakobsen, L. B. (1999). *Evaluating quality in school education; final report*. (Internetbron geraadpleegd op 18 december 2005 via: europa.eu.int/comm/education/ archive/poledu/finalrep/rep.pdf).
McBeath, J., Moos, L., & Riley, K. (1998). Time for a change. In J. MacBeath (Ed.), *Effective school leadership. Responding to change* (pp. 20-31). London: Paul Chapman/Sage.
McBeath, J., & Myers, K. (2002). *Self-evaluation: what's in it for schools?* London: Routeledge Falmer.
McEwan, E. K. (2003). *7 Steps to Effective Instructional Leadership*. Thousand Oaks, California: Corwin Press.
McMillan, J. H. (2001). *Essential assessment concepts for teachers and administrators*. Thousand Oaks, CA: Corwin Publishing Company.
Meuret, D., & Morlaix, S. (2003). Conditions of Success of a School's Self-Evaluation: Some Lessons of an European Experience. *School Effectiveness and School Improvement, 14*(1), 53-71.
Michielsens, P. (1999). Bezorgd om Kwaliteit! Concepten, ideeën en visies met betrekking tot externe evaluatie. *Informatie Vernieuwing Onderwijs, 74*, 33-42.
Michielsens, P. (2002). Zelfevaluatie versus externe evaluatie: zinvolle beleidsaspiraties? *Kwaliteitszorg in het onderwijs (rubriek Beleid en visies), 1*, 35-51.
Miles, M., & Huberman, M. (1994). *Qualitative Data Analysis*. London: Sage.
Miles, M. B., & Huberman, A. M. (1984). Qualitative Data Analysis: A Sourcebook of New Methods.
Mintzberg, H. (2001). *Organisatiestructuren*. Schoonhoven: Academic Service.
Mitchell, C., & Sackney, L. (2000). *Profound improvement. Building capacity for a learning community*. Lisse: Swets & Zeitlinger Publishers.
Morse, J. M. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioural research* (pp. 189-208). Thousand Oaks: Sage.
Mortimore, P., & MacBeath, J. (2001). School effectiveness and improvement: the story so far. In P. Mortimore & J. MacBeath (Eds.), *Improving School Effectiveness* (pp. 1-21). Philadelphia: Open University Press.
Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1988). *School Matters: The Junior Years*. Somerset: Open Books.
Muthén, L. K., & Muthén, B. O. (2003). *Mplus: statistical Analysis with Latent Variables. User's Guide*. Los Angeles, CA: Muthén & Muthén.
MVG. (2006). *Onderwijsspiegel Schooljaar 2004-2005. Verslag over de toestand van het onderwijs*. Brussel: Ministerie van de Vlaamse Gemeenschap, Onderwijsinspectie.
Nevo, D. (1995). *School-Based Evaluation: A Dialogue for School Improvement*. Oxford: Pergamon.
Nevo, D. (2001). School evaluation: internal or external? *Studies in Educational Evaluation, 27*(2), 95-106.

Nevo, D. (2002). Dialogue evaluation: combining internal and external evaluation. In D. Nevo (Ed.), *School-based evaluation: An International Perspective*. Oxford: Elsevier Science Ltd.

Nisbet, J. (1988). Rapporteur's report. In Scottish Council for Research in Education (Ed.), (pp. 1-9). Amsterdam & Lisse: Swets & Zeitlinger.

Nonaka, I., & Takeuchi, H. (1995). *De kenniscreërende onderneming: hoe Japanse bedrijven innovatieprocessen in gang zetten*. Schiedam: Scriptum.

Norton Grubb, W. (1999). Improvement or Control? A U.S. View of English Inspection. In C. Cullingford (Ed.), *An Inspector Calls: OFSTED and its Effect on School Standards*. London: Kogan Page.

Onderwijsraad. (2001). *Wat scholen vermogen. Autonomie, beleidsvoerend vermogen en bestuurlijke inrichting in het primair en voortgezet onderwijs*. Den Haag: Onderwijsraad.

Oost, H., & Markenhof, A. (2002a). *Een onderzoek uitvoeren*. Baarn: HBuitgevers.

Oost, H., & Markenhof, A. (2002b). *Een onderzoek voorbereiden*. Baarn: HBuitgevers.

Paslowsky, P., Forslin, J., & Reinhardt, R. (2001). Practices and tools of organizational learning. In M. Dierkes, A. B. Antal, J. Child & I. Nonaka (Eds.), *The handbook of organizational learning* (pp. 775-793). Oxford: Oxford University Press.

Patriotta, G. (2004). *Organizational knowledge in the making: how firms create, use, and institutionalize knowledge*. Oxford: University Press.

Peeters, K., Van Cauwenberghe, A., & Schollaert, R. (1996). *Spiegeleffecten. De Vlaamse decretale navorming doorgelicht*. Leuven: Garant.

Peters, C. M. T. (1997). *De professionaliteit van schoolleiders: een taakspecifieke benadering*. Nijmegen: Katholieke Universiteit Nijmegen.

Peters, T. J., & Waterman, R. H. (1982). *In search of excellence*. New York: Harper & Row.

Poell, R. F., & van der Krogt, F. J. (1999). Leerstrategieën van werknemers in de kenniscreërende organisatie: Van kennisdragers naar actoren in netwerken. In J. M. J. Dekker, W. M. Kessels, C. A. Smit & A. N. Papas-Talen (Eds.), *Opleiders in Organisaties, Netwerkend leren* (pp. 81-98). Deventer: Kluwer.

Potter, D., Reynolds, D., & Chapman, C. (2002). School Improvement for Schools Facing Challenging Circumstances: a review of research and practice. *School Leadership & Management, 22*(3), 243-256.

Quinn, R., Faerman, S. R., Thompson, M. P., & McGrath, M. (2003). *Becomming a master manager. A competency framework*. New York: Wiley.

Quinn, R., & Rohrbaugh, J. (1983). A spatial model of effectiveness criteria: Towards a competing values approach to organizational analysis. *Management Science, 29*, 363-377.

Quinn, R. E. (1988). *Beyond rational management: mastering paradoxes and competing demands of high performance*. San Francisco, CA: Jossey-Bass.

Raykov, T., & Marcoulides, G. A. (2000). *A First Course in Structural Equation Modeling*. Mahwah NJ: Erlbaum.

Reezigt, G., & de Jong, R. (2001). Kenmerken van effectieve schoolverbeteringsprojecten. In B. Creemers & A. Houtveen (Eds.), *Onderwijsinnovatie. Onderwijskundig lexicon, editie III* (pp. 99-118). Alphen aan de Rijn: Kluwer.

Reichardt, C. S., & Cook, T. D. (1979). Beyond qualitative versus quantitative methods. In C. S. Reichardt & T. D. Cook (Eds.), *Qualitative and quantitative methods in evaluation research*. Beverly Hills, CA: Sage.

Renkema, G. (2002). Kwaliteitszorg: wat levert het op? *Basisschoolmanagement, 15*(8), 14-19.

Rice, E. M., & Schneider, G. T. (1994). A decade of teacher empowerment: an empirical analysis of teacher involvement in decision making, 1980-1991. *Journal of Educational Administration, 32*(1), 43-58.

Rogers, P. J., & Hough, G. (1995). Improving the effectiveness of evaluations: making the link to organizational theory. *Evaluation and Program Planning, 18*(4), 321-332.

Rudd, P., & Davies, D. (2000). *Evaluating school self-evaluation.* Paper gepresenteerd op 'British Educational Research Association Conference', Cardiff University.

Sammons, P., Hillman, J., & Mortimore, P. (1995). *Key characteristics of effective schools: A review of school effectiveness research.* London: Office for Standards in Education.

Scheerens, J. (1990). School effectiveness research and the development of process indicators of school functioning. *School Effectiveness and School Improvement, 1*, 61-80.

Scheerens, J. (1996). Beoordeling en evaluatie in het kader van kwaliteitszorg in het onderwijs. In J. Scheerens (Ed.), *Kwaliteitszorg in het onderwijs. Onderwijskundig lexicon.* Alphen aan den Rijn: Samsom H.D. Tjeenk Willink.

Scheerens, J. (1997). *De bevordering van schooleffectiviteit in het basisonderwijs. Mogelijkheden tot flankerend beleid bij klassenverkleining*. Enschede: OCTO - Universiteit Twente.

Scheerens, J. (2004). The evaluation culture. *Studies In Educational Evaluation, 30*(2), 105-124.

Scheerens, J., Van Amelsfoort, H. W. C. G., & Donoughue, C. (1999). Aspects of the organizational and political context of school evaluation in four European countries. *Studies in Educational Evaluation, 25*, 79-108.

Schildkamp, K. (2007). *The Utilisation of a Self-Evaluation Instrument for Primary Education*. Enschede: PrintPartners Ipskamp.

Schollaert, R. (2000). Nascholing als aanzet tot vernieuwing op school. *Tijdschrift voor Onderwijsrecht en Onderwijsbeleid, 1*, 33-42.

Schollaert, R., Peeters, K., & van Cauwenberghe, A. (1994). *De navorming voor de spiegel. Conclusies uit vier case-studies*. Brussel: Vlaamse Gemeenschap, Departement Onderwijs.

Scriven, M. (1969). An introduction to meta-evaluation. *Educational Product report, 2*, 36-38.

Scriven, M. (1991). *Evaluation thesaurus (4th ed.)*. Newbury Park, CA: Sage.

Senge, P. (1997). *The fifth discipline: the art and practice of the learning organization*. London: Century Business.

Senge, P. (2001). *Lerende scholen: een vijfde discipline-handboek voor onderwijzers, ouders en iedereen die betrokken is bij scholing*. Schoonhoven: Academic Press.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational researcher, 27*(2), 4-13.

SICI. (2003). *Effective School Self-Evaluation*. Belfast: The Standing International Conference of Central and General Inspectorates of Education.

Simons, P. R. J. (1991). Het leren van volwassenen. *Losbladig onderwijskundig Lexicon II, A 2500*, 1-20.

Simons, P. R. J. (1997). De school als lerende organisatie. In T. Bergen, A. Knoers & P. Sleegers (Eds.), *Perspectieven op de school in dynamische ontwikkeling* (pp. 37-57). Alphen a/d Rijn: Samsom Tjeenk Willink.

Simons, P. R. J. (2003). Leren van docenten: een methodiek voor professionele ontwikkeling. In B. P. M. Creemers (Ed.), *Handboek Schoolorganisatie* (pp. 1-11). Alphen a/d Rijn: Kluwer.

Simons, P. R. J., & Ruijters, M. (2001). *Learning professionals: Towards an integrated model*. Paper gepresenteerd op 'the biannual conference of the European Association for Research on Learning and Instruction', Fribourg.

Sleegers, P. (1991). *School en beleidsvoering: een onderzoek naar de relatie tussen het beleidsvoerend vermogen van scholen en het benutten van de beleidsruimte door scholen*. Nijmegen: Katholieke Universiteit Nijmegen.

Smaling, A. (1987). *Methodologische objectiviteit en kwalitatief onderzoek*. Lisse: Swets en Zeitlinger.

Smith, S. C., & Piele, P. K. (1997). *School leadership. Handbook for excellence*. Washington D.C.: ERIC.

Smylie, M. (1992). Teacher Participation in School Decision Making: Assessing Willingness to Participate. *Educational Evaluation and Policy Analysis, 14*(1), 53-67.

Snijders, T., & Bosker, R. J. (1999). *Multilevel Analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.

Snoek, M. (2004). *Van veranderd worden naar zelf veranderen. Veranderingsbekwaamheid als metacompetentie van leraren*. Amsterdam: Hogeschool van Amsterdam.

Spillane, J., Halverson, R., & Diamond, J. B. (2001). Investigating School Leadership Practice: A Distributed Perspective. *Educational Researcher, 30*(3), 23–28.

Staessens, K. (1991). *De professionele cultuur van basisscholen: elke school heeft haar eigen verhaal*. Leuven: Universitaire Pers.

Staessens, K. (1993). Identification and description of professional culture in innovative schools. *Qualitative Studies in Education, 6*, 111-128.

Standaert, R. (1999). De verleiding van het simplificeren! Een pleidooi voor integratie van interne en externe kwaliteitsbewaking. *Informatie Vernieuwing Onderwijs, 74*, 5-20.

Stijnen, P. (1993). *Transformatie van schoolorganisaties*. Alphen aan den Rijn: Samsom.

Strauss, A. L., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage.

Struebing, J. (1999). *Grounded theory*. (Internetbron geraadpleegd op 22 november 2005 via: venus.soci.niv.edu/~archives/SSSITalk/jan99/0198.html).

Swanborn, P. G. (1991). *Basisboek sociaal onderzoek*. Boom: Boompers Drukkerijen.
Swanborn, P. G. (2000). *Case-study's: Wat, wanneer en hoe?* Amsterdam: Boom.
Swieringa, J., & Wierdsma, A. F. M. (1990). *Op weg naar een lerende organisatie*. Groningen: Wolters Noordhoff Management.
Tashakkori, A., & Teddlie, C. (2002). *Handbook of Mixed Methods in Social & Behavioral Research*. Thousand Oaks, CA: Sage Publications.
Teddlie, C., Kochan, S., & Taylor, D. (2002). The ABC+ Model for School Diagnosis, Feedback, and Improvement. In A. Visscher & R. Coe (Eds.), *School improvement through performance feedback*. Lisse: Swets & Zeitlinger B.V.
Tjepkema, S. (1993). *De profiel van de lerende organisatie en haar opleidingsfunctie*. Enschede: Universiteit Twente.
Tjepkema, S. (2002). Bouwen aan een lerende organisatie via het opsporen van leerblokkades. *Handboek 'Schoolorganisaties en Management'*. Online geraadpleegd op 25 maart 2006 (http://www.kessels-smit.nl).
Tjio, I., & Linssen, L. (2003). Zelfevaluatie: theorie en praktijk. *Q-vijver, 4*(3), 18-20.
Tschannen-Moran, M. (2001). Collaboration and the need for trust. *Journal of Educational Administration, 39*(4), 308-331.
Tymms, P., & Albone, S. (2002). Performance Indicators in Primary Schools. In A. J. Visscher & R. Coe (Eds.), *School Improvement Through Performance Feedback*. (pp. 191- 218). Lisse: Swetz & Zeitlinger.
van Aanholt, T., & Buis, T. (1990). *De school onder de loep*. Culemborg: Educaboek.
Van den Berg, R., Sleegers, P., Bakx, E., & van der Eerden, E. (1994). Het innovatief vermogen van scholen in het voortgezet onderwijs. Een kwalitatief onderzoek. *Pedagogische Studiën, 71*, 402-419.
Van den Berg, R., & Vandenberghe, R. (1999). *Succesvol leiding geven aan onderwijsinnovaties*. Alphen aan de Rijn: Samson.
Van der Haar, R. (2001). Constructivisme; ofwel hoe leert een mens. *VVO Magazine, 7*, 14-16.
Van der Veen, T., & Van der Wal, J. (1989). *Van leertheorie naar onderwijspraktijk*. Leiden: Martinus Nijhoff.
Van Hoyweghen, D. (2002). Hoe complementair zijn interne en externe evaluatie? Een antwoord vanuit het Vlaams-Nederlands project Zelfevaluatie. *Kwaliteitszorg in het onderwijs (rubriek Beleid en visies), 2*, 39-70.
Van Muijen, J. J. (1994). *Organisatiecultuur en organisatieklimaat: de ontwikkeling van een meetinstrument op basis van het 'competing values' model*. Ridderkerk: Ridderprint.
Van Petegem, P. (1996). Gewikt en gewogen, en hoe bevonden? Zelfevaluatie als strategisch instrument voor schoolontwikkeling. *Impuls voor onderwijsbegeleiding, 26*(2), 49-57.
Van Petegem, P. (1998). *Vormgeven aan schoolbeleid: effectieve-scholenonderzoek als inspiratiebron voor de zelfevaluatie van scholen*. Leuven: Acco.
Van Petegem, P. (1999). Zelfevaluatie ter verbetering en/of ter verantwoording van scholen? *Informatie vernieuwing onderwijs, 74*, 21-32.

Van Petegem, P. (2005). *Vormgeven aan schoolbeleid: effectieve-scholenonderzoek als inspiratiebron voor de zelfevaluatie van scholen*. Leuven: Acco.

Van Petegem, P., & Cautreels, P. (2006). Activerende werkvormen voor zelfevaluatie. *Kwaliteitszorg in het onderwijs (rubriek Instrumenten), 14*, 79-101.

Van Petegem, P., Devos, G., Mahieu, P., Dang Kim, T., & Warmoes, V. (2006). *Hoe sterk is mijn school? Het beleidsvoerend vermogen van Vlaamse scholen*. Mechelen: Wolters Plantyn.

Van Petegem, P., Imbrecht, I., & Brandt, W. (2005a). *Leer- en onderwijspsychologische theorieën*. Gent: Academia Press.

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

Van Petegem, P., & Vanhoof, J. (2004). Een pleidooi voor een geïnformeerde schoolontwikkeling. *Impuls voor onderwijsbegeleiding, 34*(4), 192-197.

Van Petegem, P., Vanhoof, J., & De Maeyer, S. (2007). Beleidseffectiviteit als motor voor kwaliteitsvolle zelfevaluaties: empirische evidentie uit het Vlaamse gelijkeonderwijskansenbeleid. *Tijdschrift voor onderwijsrecht en onderwijsbeleid, 1*, 33-45.

Van Petegem, P., Verhoeven, J. C., Buvens, I., & Vanhoof, J. (2005b). *Zelfevaluatie en beleidseffectiviteit in Vlaamse scholen. Het gelijke onderwijskansenbeleid als casus*. Gent: Academia Press.

Van Tulder, M., Veenman, S., & Roelofs, E. (1989). *Nascholing leraren basisonderwijs: Vormgeving en resultaten*. Nijmegen: Instituut voor Onderwijskunde.

van Wessum, L. (1996). *De sectie als eenheid: samenwerking en professionaliteitsopvattingen van docenten in het voortgezet onderwijs*. Utrecht: Universiteit Utrecht.

Vandenberghe, R. (1995). Creative management of a school: A matter of vision and daily interventions. *Journal of Educational Administration, 33*(2), 31-51.

Vandenberghe, R., Van de Velde, J., & Maes, F. (2000). *Nascholing in Basis- en Secundair Onderwijs: follow-up onderzoek en ontwikkelen van Instrument Secundair Onderwijs*. Leuven: K.U.Leuven - Centrum voor Onderwijsbeleid en -vernieuwing.

Vandenberghe, R., & Van der Vegt, R. (1992). *Scholen in de vernieuwingsarena*. Leuven-Apeldoorn: Garant.

Vandenbroucke, F. (2004). *Vandaag kampioen in wiskunde, morgen ook in gelijke kansen. Beleidsnota van de Vlaamse minister van Onderwijs en Vorming, 2004-2009*. Brussel: Ministerie van de Vlaamse Gemeenschap.

Vanderpoorten, M. (2000). *Beleidsnota 2000-2004*. Brussel: Ministerie van de Vlaamse Gemeenschap.

Vanhoof, J., & Van Petegem, P. (2006a). *Kwaliteit van scholen (openbaar) vergelijken. Het zwart/wit denken voorbij*. Mechelen: Wolters Plantyn.

Vanhoof, J., & Van Petegem, P. (2006b). *Pei/ijlen naar succesvol schoolbeleid. Praktijkboek voor de beleidseffectieve school*. Mechelen: Wolters Plantyn.

Vanhoof, J., Van Petegem, P., Mahieu, P., & Daems, F. (2004). Omtrent de wenselijkheid van publiek rapporteren over scholen: een Delphi-studie naar de visie van Vlaamse stakeholders. *Tijdschrift voor onderwijsrecht en onderwijsbeleid, 1*, 63-73.

Verbiest, E. (1995). Externe Ondersteuning - invoering en uitwerking van de basisvorming als voorbeeld. In B. Creemers, J. Giesbers, M. Krüger & C. van Vilsteren (Eds.), *Handboek Schoolorganisatie en Onderwijsmanagement* (pp. D1220/1221-1210). Alphen a/d Rijn - Deurne: Samsom H.D. Tjeenk Willink.

Verbiest, E. (2002). *Collectief leren in schoolorganisaties*. Tilburg: Fontys Hogescholen.

Verbiest, E. (2003). Leren reflecteren. *School en begeleiding - Personeel en Organisatie (rubriek Human Resources Management), 2*, 85-106.

Verbiest, E. (2004). *Collectief leren in de context van professionele leergemeenschappen - een vragenlijst (CLIS)*. (Niet gepubliceerd onderzoeksrapport). Tilburg: Fontys Hogescholen - Kenniskring schoolontwikkeling en schoolmanagement.

Verbiest, E., & Vandenberghe, R. (2002). Professionele leergemeenschappen. Een nieuwe kijk op permanente onderwijsvernieuwing en ontwikkeling van leraren. *Schoolleiding en Begeleiding - Personeel en organisatie (rubriek Human Resources Management), 1*, 57- 86.

Verhoeven, J. (1986). *Sociologie van de planmatige sociale verandering*. Leuven: Acco.

Vermunt, J. (1998). Leeractiviteiten van leerlingen. In L. Verschaffel & J. Vermunt (Eds.), *Het leren van leerlingen. Onderwijskundig Lexicon* (pp. 29-45). Alphen aan den Rijn: Samsom Tjeenk Willink.

Vermunt, J. D. H. M. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs: naar procesgerichte instructie in zelfstandig denken*. Amsterdam: Swets en Zeitlinger.

Verschaffel, L., & de Corte, E. (1998). Actief en constructief leren binnen krachtige leeromgevingen. In L. Verschaffel & J. Vermunt (Eds.), *Het leren van leerlingen. Onderwijskundig Lexicon.* (pp. 15-29). Alphen aan den Rijn: Samsom.

Verstegen, R. (2005). Autonomie en (externe) sturing. In VLOR (Ed.), *Beleidsvoerend vermogen van scholen ontwikkelen, een verkenning*. Antwerpen-Apeldoorn: Garant.

Visscher, A. J. (1993). De technologie van het primaire proces in onderwijsinstellingen. In P. J. J. Stijnen, J. Scheerens, A. M. L. van Wieringen & H. G. W. Münstermann (Eds.), *Transformatie van schoolorganisaties* (pp. 53-77). Alphen aan den Rijn: Open University.

Visscher, A. J. (2002). A Framework for Studying School Performance Feedback Systems. In A. J. Visscher & R. Coe (Eds.), *School Improvement through Performance Feedback* (pp. 41-71). Lisse: Swets & Zeitlinger.

Vleminckx, G., & Van Nuland, Y. (1998). *Op weg naar een excellente onderneming. Zelfevaluatie concreet uitgewerkt*. Blanden: Comatech.

von Krogh, G., Ichijo, K., & Nonaka, I. (2000). *Enabling Knowledge Creation. How to unlock the mystery of tacit knowledge and release the power of innovation*. Oxford: University Press.

Voogt, J. C. (1989). *Scholen doorgelicht: een studie over schooldiagnose*. De Lier.

Watling, R., & Arlow, M. (2002). Wishful Thinking: Lessons from the Internal and External Evaluations of an Innovatory Education Project in Northern Ireland. *Evaluation & Research in Education, 16*(3), 166-181.

Wenger, E. (1998). *Communities of practice: learning, meaning and identity*. Cambridge: Cambridge University Press.

Wester, F. (1996). The analyses of qualitative interviews. In F. Wester (Ed.), *The deliberate dialogue: qualitative perspectives on the interview* (pp. 63-87). Brussel: VUB Press.

White, P. A. (1992). Teacher empowerment under 'ideal' school site autonomy. *Educational Evaluation and Policy Analysis, 14*(1), 69-82.

Wikeley, F., Stoll, L., & Lodge, C. (2002). Effective School Improvement: English Case Studies. *Educational research and evaluation: an international journal on theory and practice, 8*(4), 363-385.

Wilkins, R. (2002). School as organisations: some contemporary issues. *The International Journal of Educational Management, 16*(3), 120-125.

X. (2002). *Decreet betreffende gelijke onderwijskansen I. Het Belgisch Staatsblad (18.09.2002)*.

Yin, R. (1994). *Case study research: Design and methods*. Beverly Hills, CA: Sage Publishing.

York-Barr, J., Sommers, W. A., Ghere, G. S., & Montie, J. (2001). *Reflective practice: an action guide for educators*. Thousand Oaks, CA: Corwin Press.

Zuijderhoudt, R., Wobben, J. J., & Ten Have, S. (2002). De logica van chaos in veranderingsprocessen. *Belgium management review, 19* (82), 59-67.