

Gerdy ten Bruggencate

MAKEN SCHOOLLEIDERS HET VERSCHIL ?

ONDERZOEK NAAR DE INVLOED VAN
SCHOOLLEIDERS OP DE SCHOOLRESULTATEN



Maken schoolleiders het verschil?

Onderzoek naar de invloed van schoolleiders op de
schoolresultaten

Gerdy ten Bruggencate

Samenstelling promotiecommissie:

Voorzitter	Prof. dr. H.W.A.M. Coonen
Promotor	Prof. dr. J. Scheerens
Assistent promotor	Dr. J.W. Luyten
Leden	Prof. dr. J. Van Damme Prof. dr. C.A.W. Glas Dr. M.L. Krüger Prof. dr. P.J.C. Slegers Prof. dr. A.M.L. van Wieringen

Dit onderzoek is mogelijk gemaakt met een subsidie van NWO (411-21-404)

Bruggencate, Gerdy C. ten
Maken schoolleiders het verschil?
Onderzoek naar de invloed van schoolleiders op de schoolresultaten
Proefschrift Universiteit Twente, Enschede. - Met lit. opg. - Met samenvatting in het Engels.
ISBN: 978-90-365-2783-5

Druk: PrintPartners Ipskamp B.V., Enschede
Omslagontwerp & beeld De Ondernemer (brons): Jet van der Tol
Omslagfoto: Jet van der Tol & Tom Schepers (creative bastards)
Copyright © 2009, G.C. ten Bruggencate. Alle rechten voorbehouden.
e-mail: g.c.tenbruggencate@utwente.nl

Neither this book nor any part may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, microfilming, and recording, or by any information storage and retrieval system, without written permission of the author. Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvuldigd, in enige vorm of op enige wijze, zonder voorafgaande schriftelijke toestemming van de auteur.

MAKEN SCHOOLLEIDERS HET VERSCHIL?

ONDERZOEK NAAR DE INVLOED VAN SCHOOLLEIDERS OP DE SCHOOLRESULTATEN

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. H. Brinksma,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op donderdag 12 februari om 16.45 uur

door

Gerrarda Corinne ten Bruggencate
geboren op 20 januari 1958
te Alkmaar

Dit proefschrift is goedgekeurd door de promotor prof. dr. J. Scheerens en de assistent promotor dr. J.W. Luyten.

*Notre bonheur, c'est celui des alpinistes: se sentir bien en haut,
savoir qu'on l'a fait, qu'on s'est dépassé.*

Robert Aubéry¹

Voorwoord

Een van de leuke aspecten van het onderwerp voor dit proefschrift is dat het gemakkelijk aan iedereen is uit te leggen. Het onderwerp spreekt vele mensen aan en vaak hebben zij ook een mening over de uitkomsten. Sommigen gaan ervan uit dat de schoolleider wel een substantiële invloed moet hebben op de schoolprestaties, terwijl anderen juist menen dat die invloed slechts beperkt kan zijn. Vooral in de grote schoolorganisaties in het voortgezet onderwijs lijkt het niet voor de hand te liggen dat de schoolleider veel invloed kan uitoefenen op de prestaties van individuele leerlingen. Toch wordt er in de onderzoeksliteratuur vanuit gegaan dat schoolleiderschap één van de belangrijkste effectiviteitsbevorderende schoolfactoren is. De resultaten van wetenschappelijk onderzoek naar de invloed van schoolleiders op de leerprestaties blijken echter net zo uiteen te lopen als de meningen van niet-onderzoekers.

Bij de aanvang van het onderzoek ging ik er zelf vanuit dat het een hele toer zou zijn om in het voortgezet onderwijs de effecten van het gedrag van schoolleiders te traceren... en dat bleek te kloppen. Gelukkig gaf mijn promotor Jaap Scheerens mij de nodige ruimte om het onderzoek tot een goed einde te kunnen brengen. Dat waardeer ik zeer.

Het verloop van mijn onderzoek vertoonde achteraf gezien nogal wat overeenkomsten met een beklimming van de Mont Ventoux op de fiets (zie ook bovenstaand citaat). Het begin van de route is afwisselend en niet zo steil. Het is dan nog gemakkelijk om in een groepje bij elkaar te blijven fietsen. Dit deel is te vergelijken met de eerste fase van het onderzoek, vanaf de formulering van de onderzoeksvragen tot en met de dataverzameling (beschreven in de hoofdstukken 1, 2 en 3). In deze fase werd door de onderzoekers binnen het aandachtsgebied waarin mijn project was ondergebracht intensief samengewerkt. Femke Geijssel, Meta Krüger, Joachim Lyssens, Ghislaine Schmidt en Peter Slegers, hartelijk dank voor de prettige samenwerking en de inspirerende bijeenkomsten. De vergadering in de daktuin van Meta op een prachtige zomeravond was een feestje, zoals Bob al had voorspeld. Femke wil ik in het bijzonder bedanken omdat ze steeds bereid bleef om haar vakkennis te delen, ook nadat haar project binnen het aandachtsgebied was afgerond. Verder heb ik genoten van de gezellige en opbeurende e-mails. Annick Kicken dank ik voor de versnelling die zij aan het begin van het project verzorgde als studentassistent. Ook de schoolleiders en docenten die verschillende versies van de onderzoeksinstrumenten becommentarieerden wil ik hier van harte bedanken. Zij zorgden voor de reparaties die nodig waren om door te kunnen gaan.

Het vervolg van de route op de Mont Ventoux is een lang steil stuk door het bos, met weinig bochten. Hier fietst iedereen voor zichzelf. Dit deel van de route is te vergelijken met de fase van het onderzoek waarin de instrumenten werden gevalideerd (hoofdstuk 4). De ambitieuze opzet van het project leverde hier een flinke uitdaging op. Gelukkig waren er weer mensen die mij in deze fase een zetje wilden geven: Hans Vos, Cees Glas, Karin Sanders en Peter Geurts, bedankt voor jullie bijdragen.

Na het bos volgt op de Mont Ventoux het maanlandschap: stenig en kaal. Bij mooi weer is het fietsen hier niet zo zwaar, maar het kan ook heel koud zijn en zo hard waaien dat het moeilijk is om op de weg te blijven. De analyse van de kenmerken van de scholen in relatie tot de opbrengsten (hoofdstuk 5) leverde weinig problemen op, maar bij de structurele analyse

¹ Cyclist uit Mollans-sur-Ouvèze. Citaat overgenomen uit *De kale berg, op en over de Mont Ventoux*, L. Reurings en W. Janssen Steenberg, 2002

(hoofdstuk 6) werden de omstandigheden een stuk minder. In deze fase van het project overleed mijn dagelijkse begeleider Bob Witziers, veel te jong, aan de gevolgen van een ernstige ziekte. Zijn inzet, enthousiasme en betrokkenheid, werkelijk tot op het laatst, heb ik zeer gewaardeerd. Hans Luyten sprong in als begeleider. Ik wil hem bedanken voor de vele inspirerende discussies en zijn methodologische inbreng. Het was heel fijn dat ik regelmatig kon binnenvallen om snel iets te bespreken. Zowel Hans als Jaap wil ik ook nog bedanken voor hun waardevolle commentaar op de teksten voor dit proefschrift.

Het laatste stuk van de route naar de top van de Mont Ventoux is weer behoorlijk steil, maar bovenop is er de beloning van een weids uitzicht en het gevoel dat het gelukt is. Het laatste hoofdstuk van dit proefschrift (hoofdstuk 7) biedt ook een overzicht, maar dan van de uitkomsten en de conclusies van dit onderzoek en aanbevelingen voor toekomstig onderzoek.

Bij de uitvoering van een promotieonderzoek verleg je grenzen, net als bij het beklimmen van een berg. In ieder geval je eigen grenzen, maar soms ook grenzen in de wetenschap (of de sport). Het is niet voor iedereen weggelegd om een bergtop voor de eerste maal te bedwingen, maar een goede tijd neerzetten of een nieuwe route openen, is ook de moeite waard.

Tijdens het hele promotietraject waren er vele mensen mij aanmoedigden, die meehielpen bij pech of die even een duwtje gaven in de goede richting. Collega's en oud collega's van de vakgroep Onderwijsorganisatie en -Management en de groep van Wim Nijhof wil ik daarom hartelijk bedanken voor hun interesse en ondersteuning. De medewerkers van O&M hebben ook door hun werk dit onderzoeksproject mogelijk gemaakt en ik ben hen daarvoor zeer erkentelijk. Carola Groeneweg-Bouwens, Lisenka van het Reve-Alblas, Anouk van der Heij en Emmie Keijzer hebben in verschillende fasen van het project ondersteuning verleend. Dank daarvoor. Kim Schildkamp en Karin Truijten wil ik niet alleen bedanken voor de interessante inhoudelijke discussies, maar ook voor het delen van lief en leed. Verder stel ik het zeer op prijs dat zij bereid waren om mij bij de verdediging van mijn proefschrift bij te staan als paranimfen. Vele andere UT-collega's en oud-collega's toonden regelmatig hun belangstelling voor mijn promotietraject. Hiervoor hartelijk dank. Jullie interesse was voor mij ook een drijfveer om door te gaan. Mijn ex-OC-collega en -kamergenote Elly Govers wens ik veel succes met haar promotie-onderzoek in Nieuw Zeeland. Cees Terlouw en Tjeerd Plomp wil ik hier speciaal nog even noemen. Zij behoorden tot de vroegste 'supporters' in mijn promotietraject. Dank voor jullie stimulans. Dank ook aan Hubert Coonen en Jules Pieters voor hun rol in de laatste fase van het traject.

Zonder de medewerking van schoolleiders, docenten en leerlingen van meer dan honderd scholen voor voortgezet onderwijs in Nederland had dit onderzoek niet uitgevoerd kunnen worden. Hierbij wil ik alle deelnemers aan het onderzoek hartelijk bedanken voor hun bijdrage en voor de tijd die zij investeerden. Het was zeer interessant om zoveel verschillende scholen te kunnen bezoeken en inspirerend om met zoveel schoolleiders te kunnen spreken.

Dave, Albert, Trudy en de kinderen, Kees en Sylvia en andere familie en vrienden wil ik bedanken voor hun interesse in het verloop van mijn promotietraject en hun steun.

Tijdens onze huwelijksreis fietsten wij in twee uur naar de top van de Mont Ventoux. Nog steeds is met Arjan geen berg te hoog en met Bart en Sofie erbij is het echt genieten.

Mijn moeder wil ik hartelijk bedanken voor haar voortdurende belangstelling voor mijn onderzoek. Dit proefschrift draag ik op aan mijn vader. Hij overleed in de eerste week van dit promotieonderzoek. Hij zou zo trots zijn geweest.

Gerdy

Inhoud

Voorwoord

1	Introductie	1
1.1	Inleiding	1
1.2	Effectiviteit van de schoolleider	1
1.3	Verschillen in conceptuele en methodologische benadering	5
1.4	Beschrijvingsmodel van schoolleiderschap	7
1.5	Validering van het schoolleiderschapsmodel	8
1.6	Onderzoeksvragen	9
1.7	Overzicht van het proefschrift	10
2	Theoretisch en empirisch kader	13
2.1	Inleiding	13
2.2	Perspectieven op schoolleiderschap	14
2.3	Conceptuele modellen van schoolleiderschap	16
2.4	Causale modellen van schoolleiderschap	20
2.5	Het concurrerende waarden model	25
2.6	Schoolorganisatie, schoolklimaat en schoolcultuur	30
2.7	Het schooleffectiviteitsmodel	33
2.8	Onderzoeksmodel	35
3	Opzet van het onderzoek en dataverzameling	37
3.1	Inleiding	37
3.2	Constructie van onderzoeksinstrumenten	37
3.2.1	Handelingen van schoolleiders	38
3.2.2	Schoolcultuur en schoolorganisatie	40
3.2.3	Schoolbeleving van leerlingen	42
3.3	Opzet en uitvoering van de dataverzameling	43
3.3.1	De schoolleider	44
3.3.2	Werving van scholen	46
3.3.3	Gegevensverzameling en schoolrapportage	47
3.4	Representativiteit van de deelnemende scholen; achtergrondkenmerken	48
3.4.1	Achtergrondkenmerken en representativiteit van de deelnemende scholen	48
3.4.2	Achtergrondkenmerken van de schoolleiders	50
3.4.3	Achtergrondkenmerken van de docenten	50
3.5	Data-analyse	51
3.5.1	Betrouwbaarheid en validiteit van onderzoeksinstrumenten	51
3.5.2	Optimalisering van meetmodellen met behulp van factoranalyse	52
3.5.3	Meerniveau confirmatieve factoranalyse	53
3.5.4	Aanpak van de meerniveau factoranalyse; intraklasse correlatie en design effect	53
3.5.5	Berekenen van latente variabele scores; aggregeren van variabelen	55
3.5.6	Validering van de onderzoeksinstrumenten: multitrait multimethod benadering	56
3.5.7	Toetsing van het onderzoeksmodel	56
3.6	Discussie	57

4	Validering van de onderzoeksinstrumenten	59
4.1	Inleiding	59
4.2	De validering van onderzoeksinstrumenten volgens het concurrerende waarden model	62
4.3	Betrouwbaarheid en validiteit van de vragenlijsten over schoolleidersgedrag	65
4.3.1	Passing van de meetmodellen voor schoolleidersgedrag op schoolleidersdata	65
4.3.2	Passing van de meetmodellen voor schoolleidersgedrag op docentenwaarnemingen	67
4.3.3	Parameterschattingen	70
4.3.4	Verbetering van het meetmodel van schoolleidersgedrag	72
4.3.5	Betrouwbaarheid en validiteit van het nieuwe model	75
4.3.6	Evaluatie van de confirmatieve factoranalyse en de modelverbetering	78
4.4	Betrouwbaarheid en validiteit van de vragenlijsten over schoolcultuur en schoolorganisatie	82
4.4.1	Betrouwbaarheid en validiteit van de vragenlijst over schoolcultuur	82
4.4.2	Betrouwbaarheid en validiteit van de vragenlijst over de schoolorganisatiepraktijk	85
4.5	Betrouwbaarheid en validiteit van de vragenlijst over schoolbeleving	87
4.6	Analyse van de multitrait-multimethod matrix	89
4.7	Betrouwbaarheid van de geaggregeerde variabelen en verbanden tussen de variabelen	92
4.8	Discussie	95
5	Kenmerken van scholen in relatie tot de opbrengsten	99
5.1	Inleiding	99
5.2	Het concurrerende waarden model in onderwijsonderzoek	100
5.3	Kenmerken van scholen voor voortgezet onderwijs	104
5.3.1	Beschrijving van de schalen	104
5.3.2	Kenmerken van scholen	105
5.4	Prestatiegerichtheid en ontwikkelingsgerichtheid	109
5.5	Het verband tussen kenmerken van scholen en de opbrengsten	112
5.5.1	De uitkomstvariabelen	112
5.5.2	Relaties tussen schoolkenmerken en uitkomsten	113
5.6	Discussie	116
6	Een model voor de invloed van schoolleiders op de leerprestaties	119
6.1	Inleiding	119
6.2	De invloed van contextfactoren op de leerprestaties	120
6.3	Het onderzoeksmodel	122
6.4	Analyse van het onderzoeksmodel	125
6.5	De invloed van contextvariabelen nader onderzocht	132
6.6	Discussie	136
7	Samenvatting en conclusies	143
7.1	Inleiding	143
7.2	Onderzoeksmodel en variabelen	144
7.3	Dataverzameling en validering van de onderzoeksinstrumenten	147
7.4	Effecten van schoolkenmerken op de opbrengsten	150
7.5	Een causaal model voor de invloed van schoolleiders op de leerprestaties	154
7.6	Conclusies en reflectie	158
7.7	Aanbevelingen voor vervolgonderzoek	162
7.8	Maken schoolleiders het verschil?	165
	Literatuur	167
	Summary	183
	Bijlagen	187

Hoofdstuk 1

Introductie

1.1 Inleiding

Het werk van de schoolleider is in de afgelopen jaren veel complexer geworden. Door een beleid van deregulering is het aantal regels van de overheid verminderd en als gevolg van decentralisatie zijn beslissingsbevoegdheden dichterbij de scholen komen te liggen. Hierdoor is de autonomie van scholen toegenomen. Scholen kregen meer zeggenschap over financiële, organisatorische, personele en onderwijskundige zaken. Bovendien hebben fusies tussen scholen geleid tot een schaalvergroting, waardoor de managementstructuur is veranderd. Onderwijskundige processen worden nu meer dan vroeger gestuurd door verschillende personen en teams op allerlei niveaus in de organisatie. Als gevolg van deze ontwikkelingen is het functioneren van de schoolleiding een belangrijk thema geworden in het onderzoek naar schooleffectiviteit en schoolverbetering (Slegers, 1999). Dat er een relatie bestaat tussen sterk onderwijskundig leiderschap en de schoolprestaties wordt al langer aangenomen (zie Scheerens & Bosker, 1997). Hoe de onderwijskundige sturing door schoolleiders precies verloopt, is nog maar beperkt onderzocht. Die sturing zou een goed aangrijpingspunt kunnen vormen voor veranderingen, omdat managementstrategieën gemakkelijker zijn te wijzigen dan andere belangrijke schoolfactoren, zoals het curriculum of het werkklimaat voor de leerlingen (Luyten et al., 2005). Daartoe moet echter eerst worden uitgezocht wat het effect is van de verschillende strategieën die door schoolleiders kunnen worden ingezet om sturing te geven aan de onderwijskundige processen in hun school.

1.2 Effectiviteit van de schoolleider

De onderzoeksstroming die zich bezighoudt met schooleffectiviteit wordt gezien als een reactie op de uitkomsten van onderzoek in de USA in de jaren zestig en zeventig (Coleman et al., 1966; Jencks et al., 1972), waaruit werd geconcludeerd dat schoolfactoren nauwelijks invloed zouden hebben op de leerprestaties: verreweg het grootste deel van de verschillen tussen scholen moest worden toegeschreven aan de achtergrondkenmerken van de leerlingen. Op dit onderzoek werd veel kritiek geuit: een belangrijke tekortkoming was, dat er in het onderzoek van Coleman et al. voornamelijk verschillen in materiële schoolkenmerken (aantal boeken in de schoolbibliotheek, salaris van leerkrachten) werden onderzocht, terwijl andere schoolfactoren buiten beschouwing bleven.

In het schooleffectiviteitsonderzoek dat hierna op gang kwam, wordt gezocht naar (veranderbare) schoolkenmerken die de leerprestaties positief beïnvloeden. Een veel geciteerde samenvatting van de eerste resultaten van dit onderzoek is afkomstig van Edmonds (1979). Hierin worden vijf belangrijke kenmerken van een effectieve school geformuleerd (zie Luyten, 1994, p. 124):

- een sterke schoolleiding
- hoge verwachtingen van de leerlingen
- een ordelijk, maar niet onderdrukkend schoolklimaat
- een nadruk op het aanleren van basisvaardigheden
- frequente evaluatie van de schoolvorderingen.

In de jaren zeventig en tachtig was er vooral aandacht voor de schoolleider als onderwijskundig leider. De effectieve schoolleider zou veel taken uitvoeren die direct met het onderwijsleerproces in verband staan (zie bijvoorbeeld Edmonds, 1979; Sweeney, 1982; van de Grift, 1985a). De volgende leiderschapstaken zouden positief geassocieerd zijn met leerprestaties (ontleend aan Krüger, 1994, p. 32):

- het benadrukken van de leerlingprestaties
- het benadrukken van de basisvaardigheden
- bemoeienis met de onderwijsmethoden
- het coördineren van de onderwijsprogramma's
- het regelmatig evalueren van de voortgang van leerlingen
- het ondersteunen en begeleiden van het onderwijzend personeel
- het creëren van een ordelijk en op leren gericht klimaat.

Het bleek moeilijk te zijn om een direct verband aan te tonen tussen onderwijskundig leiderschap en de leerprestaties. In 1982 publiceerden Bossert, Dwyer, Rowan en Lee een overzichtsstudie van het onderzoek naar de onderwijskundige leiderschapsrol van schoolleiders. Gebaseerd op hun bevindingen, ontwikkelden zij een algemeen beschrijvingsmodel, waarin de schoolleider zowel een direct als een indirect effect uitoefent op de leerprestaties (Dwyer et al. 1985, Dwyer, Barnett & Lee, 1987). De indirecte invloed van de schoolleider verloopt volgens dit model via het onderwijskundig klimaat en de onderwijsorganisatie. Het gedrag van de schoolleider wordt daarbij bepaald door persoonlijke kenmerken, de omgeving van de school en wettelijke kaders en regelgeving. Het algemene beschrijvingsmodel wordt nader besproken in paragraaf 1.4. (zie figuur 1.1).

Begin jaren negentig werd onderwijskundig leiderschap nog steeds gezien als de belangrijkste taak van effectieve schoolleiders, naast beheersmatig management. Veel schoolleiders bleken echter niet of nauwelijks aan onderwijskundig leiderschap toe te komen. Murphy identificeerde in een overzichtsstudie (1990) een aantal belemmeringen voor onderwijskundig leiderschap binnen het onderwijs in de USA (zie Van Vilsteren, 1999, p. 19 e.v.): ten eerste zou het schoolleiders ontbreken aan voldoende onderwijskundige expertise, ten tweede werden onderwerpen op het gebied van curriculum en instructie door schoolleiders niet ervaren als kritische grootheden en ten derde zou de autonomie van docenten het onderwijskundig leiderschap belemmeren. Naar aanleiding van deze uitkomsten ontwikkelde Murphy een breder kader voor onderwijskundig leiderschap met als hoofdindeling (zie van Vilsteren, 1999, p. 57 e.v.):

- ontwikkelen van missie en doelen
- runnen van de onderwijsproductie functie (management van curriculum en instructie)
- bevorderen van een op leren gericht klimaat
- ontwikkelen van een ondersteunende werkomgeving

Met behulp van kennis die was verkregen over de activiteiten van de schoolleider, werden instrumenten ontwikkeld om grootschaliger onderzoek te doen naar de relatie tussen schoolleidersgedragingen en leerprestaties, zoals de *Principal Instructional Management*

Rating Scale (Hallinger, 1983; Hallinger & Murphy, 1985). Hallinger en Heck (1996a, 1996b, 1998) evalueerden de resultaten van dit type onderzoek tussen 1980 en 1995. Uit hun overzichtsstudie blijkt dat er in de loop van de tijd een verschuiving is opgetreden van het gebruik van ‘direct-effect-modellen’ naar complexere ‘indirect-effect-modellen’. In indirect-effect-modellen wordt ervan uitgegaan dat de invloed van de schoolleider op de leerprestaties verloopt via intermediaire (interveniërende) variabelen, zoals de schoolorganisatie, de schoolcultuur of het gedrag van docenten. In direct-effect-modellen worden intermediaire variabelen buiten beschouwing gelaten. Hallinger en Heck (1998) laten zien dat onderzoek met behulp van direct-effect-modellen weinig consistente resultaten opleverde. Een enkele keer werd een zwak effect gevonden, maar meestal bleek er geen significant verband te zijn tussen schoolleiderschap en uitkomsten. Uit de uitkomsten van onderzoek met indirect-effect-modellen kan echter worden opgemaakt dat er een klein, maar aantoonbaar indirect effect bestaat van schoolleiderschap op de leerprestaties. Hallinger en Heck merken op dat de uitkomsten voor indirect-effect-modellen consistentere werden toen onderzoekers krachtiger analysemethoden gingen toepassen. Hierbij gaat het bijvoorbeeld om causale analyse technieken, waarmee directe en indirecte effecten simultaan berekend kunnen worden. Het opnemen van contextuele variabelen in de modellen leidde daarnaast ook tot betere resultaten.

Naast direct en indirect-effect-modellen onderscheiden Hallinger en Heck (1998) nog een derde categorie van modellen voor schoolleiderschap, namelijk ‘reciproque-effect-modellen’. In deze modellen worden reciproque (wederkerige) relaties verondersteld tussen schoolleiderschaps-variabelen, intermediaire variabelen en uitkomsten. Er wordt bijvoorbeeld vanuit gegaan dat er niet alleen effecten zijn van de handelingen van schoolleiders op de schoolorganisatie en op de leerprestaties, maar ook andersom, van de leerprestaties en de schoolorganisatie op het gedrag van de schoolleider. Reciproque-effect-modellen zijn nog maar weinig onderzocht, omdat hiervoor longitudinale data nodig zijn en die zijn schaars (Hallinger & Heck, 1998). Volgens van Vilsteren (1999, p. 36) sluit de benaderingswijze van onderwijskundig leiderschap die is ontwikkeld door Leithwood en zijn collega’s, in samenwerking met de groep rond Murphy, goed aan bij het reciproque model. In de benadering die ‘cognitive perspective on educational leadership’ wordt genoemd, zijn de denkprocessen van de schoolleider opgenomen in de keten van variabelen die beschrijft op welke wijze de invloed van de schoolleider op de leerprestaties tot stand komt (zie bijv. Leithwood, 1995). Deze denkprocessen worden niet alleen gestuurd door factoren van buiten de school, maar ook door omstandigheden binnen de schoolorganisatie. Het cognitieve perspectief veronderstelt dus een reciproque relatie tussen schoolleidersgedrag en de schoolorganisatie: de schoolleider oefent een effect uit op de schoolorganisatie, terwijl het gedrag van de schoolleider ook weer wordt beïnvloed door de schoolorganisatie.

In hun beschrijving van de bijdrage van de schoolleider aan de effectiviteit van de school onderscheiden Hallinger & Heck (1998) vier domeinen, waardoor schoolleiders invloed zouden kunnen uitoefenen (zie ook De Maeyer & Rymenans, 2004, p. 277):

- visie en doelstellingen
- organisatiestructuur en sociale netwerken
- menselijk kapitaal
- organisatiecultuur

Deze indeling is gebaseerd op het werk van Leithwood (1994) en Ogawa en Bossert (1995). Hallinger en Heck geven aan dat het verband tussen de rol van de schoolleider, de organisatiecultuur en de leerprestaties nog het meest onzeker is.

De hierboven genoemde vier domeinen geven aan dat de opvatting van onderwijskundig leiderschap in de jaren negentig is verbreed. Door de schaalvergroting en de toenemende autonomie van scholen en door het dynamischer en kritischer worden van de omgeving, is het voor schoolleiders noodzakelijk om als onderwijskundig leider een meer strategische oriëntatie aan te nemen (Murphy, 1990; Marsh, 1997; zie van Vilsteren, 1999, p. 40). Terwijl eerder in het werk van de schoolleider wel een onderscheid werd gemaakt tussen onderwijskundig leiderschap ('instructional leadership') en algemene of beheersmatige leiderschapstaken (zie o.a. Scheerens & Bosker, 1997), valt die scheiding nu weg (Imants, 1996). Van Vilsteren (1999) omschrijft deze nieuwe invalshoek als volgt: "Onderwijskundig leiderschap onderscheidt zich van administratief beheer niet meer primair door de onderwijskundige aard van de objecten van leidinggeven, maar door de primair onderwijskundige intentie en deskundigheid waarmee een leidinggevende een zaak of thema in zijn school tracht te beïnvloeden." (p. 42).

Om te komen tot een overzicht van de componenten van schooleffectiviteitsbevorderende factoren, analyseerden Scheerens en Bosker (1997) operationele definities en instrumenten uit een groot aantal schooleffectiviteitsstudies en zelfevaluatiesystemen. De uitgebreide lijst van componenten van schoolleiderschap (zie Scheerens & Bosker, 1997, p. 101 e.v.) is onderverdeeld in algemene leiderschapsvaardigheden en onderwijskundig leiderschap, waarbij onderwijskundig leiderschap wordt opgevat als het uitoefenen van een indirecte controle en beïnvloeding van het primaire proces. De Maeyer en Rymenans (2004, p. 26) merken op dat de lijst een grote diversiteit aan componenten bevat, waaruit blijkt dat het bij schoolleiderschap om een zeer breed concept gaat. Scheerens (1992, p. 89) wijst erop dat het omvangrijke takenpakket van de schoolleider niet per se door één persoon hoeft te worden uitgevoerd. Taken kunnen zijn gedelegeerd aan andere functionarissen, maar kunnen bijvoorbeeld ook worden vervuld via participatieve besluitvorming, als er consensus bestaat over de schooldoelen (Kerr & Jermier, 1978).

Vanaf het begin van de jaren negentig speelt naast onderwijskundig leiderschap ook het concept 'transformationeel leiderschap' een belangrijke rol in het onderzoek naar effectief schoolleiderschap. De Maeyer en Rymenans (2004) beschrijven transformationeel leiderschap als volgt: "Een transformationeel leider is er vooral op gericht om zijn medewerkers te motiveren en te ondersteunen om hun hoogste potentieel te bereiken." (p. 246). Het begrip is ontleend aan de algemene leiderschapsliteratuur van de jaren zeventig en tachtig en het is door Leithwood en zijn collega's aangepast aan de onderwijscontext (zie bijv. Leithwood, Tomlinson & Genge, 1996; Leithwood & Jantzi, 2000, 2006). Recente publicaties wijzen erop dat het van belang is dat onderwijskundig leiderschap en transformationeel leiderschap geïntegreerd worden, om te komen tot hoge leerprestaties (Hallinger, 2003; Marks & Printy, 2003; zie ook De Maeyer & Rymenans, 2004, p. 250).

Het bovenstaande overzicht laat zien dat er in de afgelopen decennia de nodige ontwikkelingen hebben plaatsgevonden in het onderzoek naar effectief schoolleiderschap, maar hoe groot is nu precies de impact van schoolleidersactiviteiten op de leerprestaties? Inmiddels is duidelijk dat de voornaamste conclusies van het onderzoek van Coleman et al. (1966) eigenlijk niet kunnen worden weerlegd. Slechts een klein percentage van de totale spreiding in leerprestaties kan worden toegeschreven aan verschillen tussen scholen en materiële schoolkenmerken hebben nauwelijks invloed op de leerprestaties (Luyten, 1994 p. 123). Uit een diepgravende analyse door Scheerens en Bosker (1997) van de uitkomsten van het schooleffectiviteitsonderzoek, kan worden geconcludeerd dat kenmerken van de leeromgeving meer invloed hebben op de leerprestaties dan schoolorganisatiekenmerken. Het

effect van schoolleiderschap wordt geschat op 0.05, dit is een heel klein effect, ook ten opzichte van het totale schooleffect. Gebaseerd op een meta-analyse van 168 onderzoeken uit de hele wereld (Bosker & Witziers, 1996) rapporteren Scheerens & Bosker (1997, p. 77) een bruto schooleffect van 0.48 en een netto schooleffect van 0.30. Dat wil zeggen dat 19% van de variantie in het leerresultaat zou kunnen worden toegeschreven aan de school wanneer geen rekening wordt gehouden met verschillen in aanvangsniveau tussen leerlingen (bruto) en 8% van de variantie wanneer wel wordt gecorrigeerd voor instroomverschillen (netto).

Hoewel het effect van schoolleiderschap klein is, kan het toch grote gevolgen hebben. Het gaat namelijk om een effect van een of enkele personen op vele anderen (Witziers, Bosker & Krüger, 2003), hierdoor wordt het effect als het ware vermenigvuldigd (zie Scheerens & Bosker, 1997, p. 79). Scheerens en Bosker (1997) geven aan dat het totale schooleffect kan leiden tot een zodanig verschil in resultaten op de eindtoets van de basisschool in Nederland, dat leerlingen met dezelfde mogelijkheden op verschillende niveaus instromen in het voortgezet onderwijs (op de ene school haalt de leerling bijvoorbeeld vmbo-niveau en op de andere school havo). De grootte van het effect zegt dus niet zonder meer iets over de impact. Verder kan het ook moeilijk zijn om een relatie aan te tonen tussen schoolleiderschap en (netto) leerprestaties, als het schoolleiderschap op alle scholen aan bepaalde minimum-eisen voldoet (zie Scheerens & Bosker, 1997, p. 306). Het verschil in impact is dan maar heel klein.

1.3 Verschillen in conceptuele en methodologische benadering

Er is nog relatief weinig bekend over de wegen waarlangs de bijdrage van de schoolleider aan de effectiviteit van de school verloopt. Hallinger en Heck concluderen op basis van hun overzichtsstudie van onderzoek naar schoolleiderschap (1998) dat kan worden aangenomen dat de invloed van de schoolleider indirect verloopt via de schooldoelen, de schoolorganisatie, de schoolcultuur en de mensen die betrokken zijn bij de school. Met name het uitzetten van een richting voor de school, door middel van visie, missie en doelen, wordt gezien als een belangrijke manier voor schoolleiders om invloed uit te oefenen. Hallinger en Heck (1998) wijzen erop dat er inconsistenties zijn in de uitkomsten van schoolleiderschapsonderzoek, die voor een deel zijn toe te schrijven aan ontoereikende modellen en analysetechnieken (zie paragraaf 1.2). Zij geven aan dat er meer rekening gehouden zou moeten worden met het feit dat onderwijskundige databestanden meestal een geneste of hiërarchische structuur hebben: leerlingen zijn bijvoorbeeld geclusterd (genest) in klassen en klassen in scholen. Op elk niveau in deze structuur kunnen andere effecten optreden. Leerlingen brengen ieder hun eigen vaardigheden met zich mee, docenten geven vorm aan de leeromgeving binnen een klas en schoolleiders beïnvloeden de hele schoolorganisatie. Met behulp van meerniveau analysetechnieken kunnen effecten op verschillende niveaus in de structuur onderscheiden worden, zodat betrouwbaarder resultaten verkregen worden (zie ook Scheerens & Bosker, 1997). Deze technieken zijn echter pas de laatste jaren gangbaar geworden.

Conceptuele problemen spelen ook nog steeds een rol bij het vergelijken van de uitkomsten van onderzoek naar schoolleiderschap. Van Vilsteren (1999) concludeert uit zijn analyse van onderwijskundig leiderschap in het voortgezet onderwijs in Nederland, dat schoolleiderschap een ‘lastig in kaart te brengen’ onderwerp is. Hij merkt daarbij op dat “het aantal operationele categorieën waarin onderwijskundig leiderschap wordt uitgedrukt in de loop van de tijd kleiner is geworden, maar ook algemener en abstracter” (Van Vilsteren, 1999, p. 248). Een bezwaar van deze ontwikkeling is dat de abstracte categorieën voor onderzoek weer geoperationaliseerd moeten worden op handelingsniveau.

Volgens Van Vilsteren (1999) komt het ook voor dat onderzoekers “puttend uit eerder onderzoek, onderwijskundig leiderschap operationaliseren met behulp van enkele items, waaraan nauwelijks een kader ten grondslag ligt” (p. 64).

Het trekken van algemene conclusies uit de uitkomsten van het onderzoek naar effectief schoolleiderschap wordt gehinderd door de verschillende conceptuele en methodologische benaderingen (zie ook Bryk & Raudenbush, 1992, p. 155; Raudenbush, 1994). Om deze verschillen in benadering te ondervangen, voerden Witziers, Bosker en Krüger (2003) een kwantitatieve meta-analyse uit van de uitkomsten van onderzoek met direct-effect-modellen, waarbij zij het PIMRS instrument (*Principal Instructional Management Rating Scale*, Hallinger, 1989, 1994) gebruikten als raamwerk om leiderschapshandelingen te categoriseren. Bovendien hielden zij rekening met verschillende factoren (moderatoren) die van invloed zouden kunnen zijn op de effectgrootte, zoals het land waarin de studie werd uitgevoerd, het type onderwijs waarin het onderzoek werd uitgevoerd (primair of secundair) en karakteristieken van de onderzoeksopzet, zoals het soort test dat gebruikt werd om de leerprestaties vast te stellen (taal, wiskunde, of samengesteld) en de statistische techniek die werd gebruikt (wel of niet meerniveau).

Witziers et al. analyseerden eerst de resultaten van alle onderzoeken gezamenlijk aan de hand van de gemiddelde effectgrootte per onderzoek. Zij vonden dat de correlaties tussen schoolleiderschap en leerprestaties in het algemeen lager zijn dan 0.1. Dat betekent dat niet meer dan 1% van de variatie in leerprestaties is gekoppeld aan verschillen in schoolleiderschap. Uit meer gedetailleerde analyses van de invloed van moderatoren bleek dat er geen bewijs is voor een (direct) effect van schoolleiderschap in het voortgezet onderwijs, terwijl er in het basisonderwijs (buiten Nederland) wel relaties tussen schoolleiderschap en uitkomsten aantoonbaar zijn. Voor onderzoek dat is uitgevoerd in Nederland (zie bijvoorbeeld Brandsma, 1993; Houtveen, Vermeulen & Van de Grift, 1993; Van de Grift, 1990) geldt dat de gemiddelde effectgrootte vrijwel gelijk is aan nul. Door Witziers et al. werden geen verbanden gevonden tussen karakteristieken van de onderzoeksopzet en de uitkomsten van de studies. In een tweede meta-analyse werden de effecten van specifieke leiderschapsgedragingen onderzocht. Voor enkele van deze gedragingen werden kleine effecten gevonden. Het formuleren en uitdragen van een missie bleek, zoals ook door Hallinger en Heck (1998) werd geconstateerd, het meest relevante aspect van schoolleiderschap te zijn.

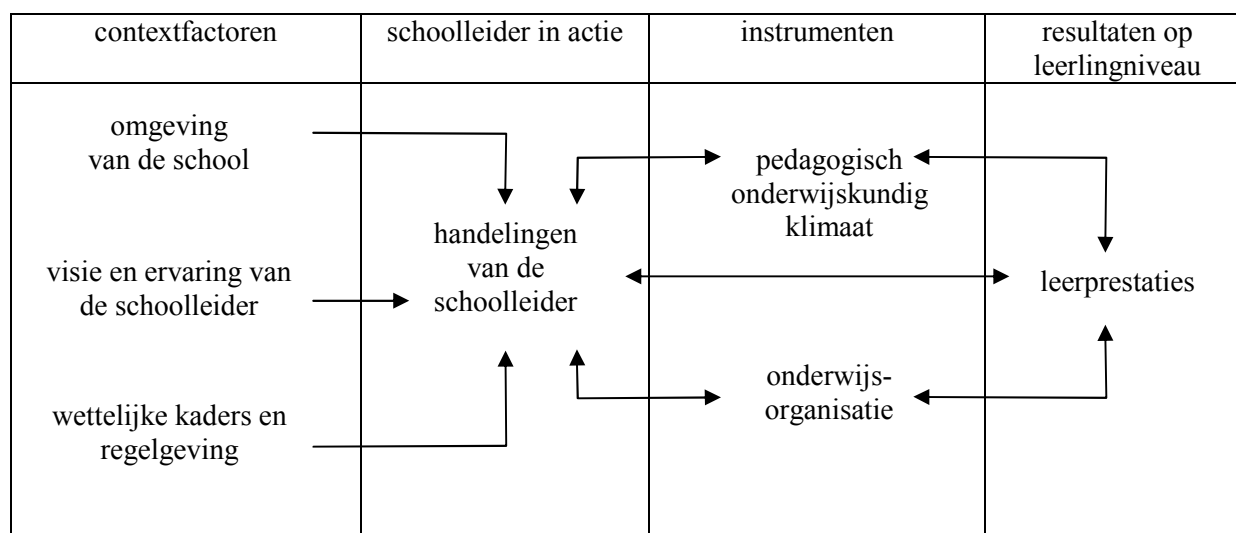
De meta-analyse van Witziers et al. (2003) laat zien dat direct-effect-modellen geen overtuigende resultaten opleverden, onafhankelijk van de onderzoeksopzet. De vraag is nu of indirect-effect-modellen meer inzicht kunnen bieden in de relatie tussen schoolleiderschap en leerprestaties. Hallinger en Heck (1998) stelden vast dat deze modellen betere resultaten opleveren wanneer geavanceerde onderzoeksmethoden worden gebruikt, zoals meerniveau analyse of *structural equation modeling* (SEM). Witziers en zijn collega's zochten daarom naar recente publicaties (vanaf 1995) van onderzoek waarin causale analyse technieken werden toegepast om indirecte effecten van schoolleiderschap op de leerprestaties te testen. Zij vonden slechts vijf onderzoeken die aan de criteria voldeden. Deze onderzoeken tonen aan dat er een verband bestaat tussen schoolleiderschap en de schoolcultuur, de schoolorganisatie, het gedrag van docenten en de praktijk in de klas, terwijl deze factoren op hun beurt weer gerelateerd zijn aan leerprestaties. Witziers et al. (2003) concluderen dat het indirect-effect-model van schoolleiderschap aanknopingspunten biedt voor toekomstig onderzoek, maar zij veronderstellen dat ook hierbij slechts een zwak verband tussen schoolleiderschap en leerprestaties zal worden gevonden.

1.4 Beschrijvingsmodel van schoolleiderschap

In hun artikel over Onderwijskundig leiderschap in moderne onderwijsinstellingen beschrijven Krüger, Witziers, Slegers en Imants (1999) de huidige kijk op onderwijskundig leiderschap als volgt:

Waar onderwijskundig leiderschap in eerste instantie nog gezien werd als het uitvoeren van relatief eenvoudige aan het onderwijs gerelateerde taken, houdt onderwijskundig leiderschap volgens de meest recente opvatting in dat de schoolleider alle aspecten van de organisatie beïnvloedt en vormgeeft op een zodanige manier dat het primaire proces wordt geoptimaliseerd en verbeterd. Het gaat dus niet om het uitvoeren van taken op het primaire proces in tegenstelling tot de uitvoering van beheersmatige taken met betrekking tot het secundaire proces. Veel belangrijker blijkt te zijn de voortdurende gerichtheid van de schoolleider op het primaire proces bij alle taken die worden vervuld, waarbij de handelingen van de schoolleider worden gerelateerd aan de context en de missie van de school. (p. C4120-21)

Krüger et al. wijzen erop dat het beschrijvingsmodel voor onderwijskundig leiderschap, dat is ontwikkeld in de jaren tachtig door onderzoekers van het Far West Laboratory (nu WestEd) in San Francisco, het verband tussen de handelingen van de schoolleider en de context goed weergeeft (Bossert, Dwyer, Rowan & Lee, 1982; Dwyer et al, 1985; Dwyer, Barnett & Lee, 1987). In het model van Bossert et al. (zie figuur 1.1) zijn de intermediaire variabelen voor de invloed van de schoolleiders samengevat in twee domeinen: pedagogisch didactisch klimaat en onderwijsorganisatie. Het model veronderstelt wederkerige relaties tussen de handelingen van de schoolleider, de intermediaire variabelen en de uitkomstvariabelen.



Figuur 1.1 Algemeen beschrijvingsmodel van onderwijskundig leiderschap (Dwyer et al., 1985; Dwyer, Barnett & Lee, 1987), overgenomen uit Krüger et al. (1999)

Reciproque effect modellen van schoolleiderschap zijn nog nauwelijks onderzocht, vanwege de longitudinale gegevens die nodig zijn en de complexe analysetechnieken (zie Hallinger & Heck, 1998). Wanneer de wederkerigheid van effecten buiten beschouwing wordt gelaten, kan het bovenstaande model worden opgevat als een indirect-effect-model met contextfactoren (zie ook Leithwood, Begley & Cousins, 1990). Voor het indirect-effect-model van schoolleiderschap bestaat enige empirische ondersteuning (voor een overzicht zie:

Hallinger & Heck, 1998; Witziers, Bosker en Krüger, 2003; Krüger, Witziers & Slegers, 2007). Er is echter weinig consistentie in de intermediaire variabelen die worden geïdentificeerd. Voorbeelden van intermediaire variabelen uit verschillende studies zijn: de missie van de school, betrokkenheid van docenten bij de school, interacties tussen docenten en leerlingen en het leerklimaat in de klas (Scheerens & Witzers, 2005).

Om een theoretische basis te verkrijgen voor de relaties tussen schoolleiderschapsvariabelen, intermediaire variabelen en leerprestaties, stellen Witziers, Bosker en Krüger (2003) voor om gebruik te maken van het *competing values framework* (concurrerende waarden model) van Quinn & Rohrbaugh (1983). Dit raamwerk is ontwikkeld aan de hand van de effectiviteitscriteria van verschillende (concurrerende of elkaar aanvullende) organisatiemodellen. Eerder onderzoek heeft aangetoond dat het model geschikt is om de schoolcultuur in kaart te brengen (Maslowski, 2001; Houtveen, 1995). Daarbij is vastgesteld dat er een verband bestaat tussen de schoolcultuur en leerlingresultaten (zie Maslowski, 2001). Het concurrerende waarden model biedt echter ook de mogelijkheid om een (reciproque) relatie te leggen tussen leiderschapsgedrag en de organisatiecultuur: leiders kunnen bijdragen aan de ontwikkeling van een bepaalde cultuur, maar hun gedrag wordt ook beïnvloed door de heersende organisatiecultuur (Quinn, 1998). De toepassing van het concurrerende waarden model op schoolleidersgedrag is nog maar beperkt onderzocht (zie bijv. Thompson, 2000).

1.5 Validering van het schoolleiderschapsmodel

Uit het voorgaande overzicht van de ontwikkelingen in het onderzoek naar schoolleiderschap kan worden opgemaakt dat er grote behoefte bestaat aan de validering van een causaal model van de invloed van schoolleiders op de leerprestaties, waarin zowel intermediaire variabelen als contextfactoren zijn opgenomen. Om een dergelijk complex model te kunnen toetsen, terwijl ook rekening wordt gehouden met de hiërarchische structuur van de data, zijn gegevens nodig van een groot aantal scholen. Daarom werd in de afgelopen jaren door de Universiteit van Amsterdam (UvA) en de Universiteit Twente (UT) samengewerkt aan een grootschalig onderzoek, waarbij de volgende vraagstelling centraal stond:

Op welke wijze geven schoolleiders in het voortgezet onderwijs sturing aan onderwijsprocessen binnen de school en wat is de invloed van deze sturing op de effectiviteit van de school?

In het samenwerkingsproject is het model van schoolleiderschap van Bossert et al. (zie figuur 1) als uitgangsmiddel gebruikt. Het onderzoeksteam van de Universiteit van Amsterdam richtte zich op het eerste deel van de keten van variabelen in het model. In een promotie-onderzoek werden de relaties onderzocht tussen contextvariabelen, de interne processen (visie) van de schoolleider en leiderschapshandelingen (Schmidt, 2009). De visie van de schoolleider werd hierbij opgevat als een intermediaire variabele tussen context en schoolleidersgedrag (zie Leithwood, Begley & Cousins, 1990). Aan de Universiteit Twente stond het tweede deel van de keten centraal. Het onderzoek naar de relaties tussen leiderschapsgedrag, intermediaire variabelen en effecten op leerlingenniveau vormt het onderwerp van deze dissertatie. De onderzoekers van de Universiteit van Amsterdam en de Universiteit Twente hebben vooral samengewerkt bij de opzet van het onderzoek, de constructie van de onderzoeksinstrumenten en de dataverzameling.

Naast de twee deelprojecten was een integrerend project gepland, om de hele causale keten te toetsen. Het integrerend project liep enigszins vooruit op de afronding van de twee promotie-onderzoeken en had daarom vooral een oriënterend karakter (Slegers & van Schooten, 2006).

1.6 Onderzoeksvragen

Het doel van dit onderzoek is om inzicht te geven in de wijze waarop schoolleiders invloed kunnen uitoefenen op de effectiviteit van de school. Er wordt verondersteld dat schoolleiders niet zozeer een directe invloed hebben op de leerprestaties, maar vooral een indirecte, via intermediaire factoren zoals de schoolorganisatie en de schoolcultuur. Veel van het onderzoek naar de relaties tussen schoolleiderschap en uitkomsten op leerlingenniveau heeft plaatsgevonden in het basisonderwijs. De vraag is, of daar gevonden resultaten ook geldig zijn in het voortgezet onderwijs (Slegers & van Schooten, 2006, p. 5). Gebaseerd op deze uitgangspunten is de volgende hoofdvraag geformuleerd:

Wat is de directe en indirecte invloed van schoolleiders op de leerprestaties, in het voortgezet onderwijs in Nederland?

In dit onderzoek zijn de schoolleidersvariabelen opgevat als onafhankelijke variabelen. De predictoren van schoolleidersgedrag vormen het onderwerp van een complementair onderzoek aan de Universiteit van Amsterdam. Contextfactoren die van invloed zouden kunnen zijn op zowel de leerprestaties, als de handelingen van schoolleiders, of de intermediaire variabelen, zijn echter wel meegenomen. Deze contextfactoren zijn van belang, omdat het grootste deel van de verschillen tussen scholen verklaard kan worden door verschillen in kenmerken van de leerlingenpopulatie. Wanneer de contextvariabelen genegeerd zouden worden, zou een te groot deel van de variantie kunnen worden toegeschreven aan verschillen in schoolleidersgedrag. Verder wordt in dit onderzoek, zoals gebruikelijk in schooleffectiviteitsonderzoek, uitgegaan van leerprestaties als uitkomstvariabele. Hierbij kan bijvoorbeeld worden gerefereerd aan Bosker, Béguin en Rekers-Mombarg (2001), die stellen dat: “de invloed van scholen op het functioneren van leerlingen afneemt naarmate we ons verder van het bedoelde curriculum - zoals vastgelegd in kerndoelen, eindtermen en exameneisen - af bewegen” (p. 122).

Om de indirecte invloed van schoolleiders in kaart te kunnen brengen moet een relatie worden gelegd tussen contextvariabelen, schoolleiderschapsvariabelen, intermediaire variabelen en leerprestaties. Voor het karakteriseren van zowel schoolleidersvariabelen als van de centrale intermediaire variabelen (schoolorganisatie en schoolcultuur) wordt gebruik gemaakt van het concurrerende waarden model (Quinn, & Rohrbaugh, 1983). Dit is een generiek model waarmee zowel de organisatie-effectiviteit, als de organisatiecultuur en leiderschapsgedrag in kaart kan worden gebracht. Met de toepassing van het concurrerende waarden model wordt voortgebouwd op het werk van Maslowski (2001) en Houtveen, Voogt, van der Vegt en van de Grift (1995). Deze onderzoekers gebruikten het model om instrumenten te ontwikkelen waarmee de schoolcultuur gemeten kan worden. Met name het instrument van Houtveen et al. is hierbij interessant omdat dit instrument zowel aspecten van de onderwijsorganisatie als de –cultuur omvat. Met behulp van het concurrerende waarden model kan een invulling worden gegeven aan de intermediaire variabelen die in het model van Bossert et al. (zie figuur 1) zijn samengevat in de categorieën: pedagogisch onderwijskundig klimaat en onderwijsorganisatie. Bovendien biedt het concurrerende waarden model de mogelijkheid om relaties te leggen tussen leiderschapsgedrag en organisatiekenmerken (Quinn, 1998).

Behalve schoolorganisatie, schoolcultuur en leerprestaties zijn er nog andere intermediaire en uitkomstvariabelen denkbaar waarin de invloed van de schoolleider doorwerkt, zoals kenmerken van het onderwijsproces in de klassen en de schoolbeleving van leerlingen (zie bijv. Heck & Hallinger, 2000). Om de wegen waarlangs de schoolleider invloed uitoefent zichtbaar te kunnen maken is het van belang om de belangrijkste variabelen te identificeren.

Naar aanleiding van de bovenstaande overwegingen zijn de volgende onderzoeksvragen afgeleid uit de hoofdvraag:

1. Welke contextvariabelen, schoolleidersvariabelen, intermediaire variabelen en uitkomstvariabelen zijn relevant voor het in kaart brengen van de effectiviteit van de schoolleider en welke relaties zijn er tussen deze variabelen te onderscheiden?
2. Is het concurrerende waarden model een geschikt kader om zowel het schoolleidersgedrag als de schoolcultuur en de schoolorganisatie te karakteriseren en relaties tussen deze variabelen bloot te leggen?
3. Zijn er (directe) relaties tussen de kenmerken van scholen voor voortgezet onderwijs in Nederland en de opbrengsten?
4. Wat zijn de directe en indirecte effecten van schoolleiderschap op de leerprestaties en hoe verlopen deze effecten langs intermediaire variabelen, gegeven de contextfactoren?

1.7 Overzicht van dit proefschrift

Door veranderingen in het onderwijs is het werk van schoolleiders complexer geworden, terwijl zij daarnaast een grotere verantwoordelijkheid hebben gekregen voor de onderwijskwaliteit. Schoolleiders moeten omgaan met een toegenomen *accountability* (verantwoordingsplicht), die in Nederland bijvoorbeeld tot uitdrukking komt in de 'kwaliteitskaarten' (vanaf najaar 2008 'toezichtkaarten'), waarmee de resultaten van scholen openbaar gemaakt worden door de Inspectie van het Onderwijs. Uit de voorgaande paragrafen blijkt echter dat het nog allerminst duidelijk is in hoeverre schoolleiders daadwerkelijk invloed kunnen uitoefenen op de leerresultaten. Dit proefschrift draait om de vraag of schoolleiders 'het verschil kunnen maken': wat kunnen schoolleiders doen om de prestaties van leerlingen te verbeteren? In de volgende hoofdstukken worden de vier onderzoeksvragen (zie paragraaf 1.6) achtereenvolgens beantwoord.

In hoofdstuk 2 staat de eerste onderzoeksvraag centraal. Het gaat hier om het identificeren van relevante context-, leiderschaps-, intermediaire en opbrengstvariabelen voor het modelleren van schoolleiderschap. Aan de hand van de resultaten van een literatuuronderzoek wordt het theoretisch en empirisch kader beschreven. Eerst wordt ingezoomd op (causale) modellen voor schoolleiderschap die eerder zijn onderzocht en intermediaire variabelen die daarbij zijn gevonden. Daarna volgt een beschrijving van het concurrerende waarden model met het oog op het karakteriseren van schoolleidersgedrag, de schoolcultuur en de schoolorganisatie. Vervolgens wordt het algemene model voor schooleffectiviteit besproken en de belangrijkste effectiviteitsbevorderende factoren. Aan het eind van hoofdstuk 2 wordt het onderzoeksmodel gepresenteerd. Hoofdstuk 3 gaat over de opzet van het onderzoek. In dit hoofdstuk worden de steekproef, de dataverzameling, de constructie van de onderzoeksinstrumenten en de analysemethoden beschreven. De betrouwbaarheid en de validiteit van de instrumenten komen aan de orde in hoofdstuk 4. Daarbij staat de tweede onderzoeksvraag centraal: is het concurrerende waarden model een geschikt model voor het karakteriseren van schoolleidersgedrag, schoolcultuur en schoolorganisatie en van de relaties hiertussen? Hoofdstuk 5 begint met een overzicht van de toepassing van het concurrerende waarden model in onderwijsonderzoek. Daarna worden scholen voor voortgezet onderwijs in

Nederland getypeerd aan de hand van de kenmerken die in dit onderzoek zijn gemeten. Vervolgens wordt de derde onderzoeksvraag behandeld. Hierbij gaat het om ‘directe’ effecten van de schoolleidersvariabelen en andere variabelen in het onderzoeksmodel op de uitkomstvariabelen. In hoofdstuk 6 komt de vierde onderzoeksvraag aan de orde. Door middel van een (direct en) indirect-effect-model met contextvariabelen worden directe en indirecte effecten van schoolleidersgedrag in kaart gebracht. Hoofdstuk 7 bevat een samenvatting van de resultaten met de conclusies die hieruit getrokken kunnen worden.

Hoofdstuk 2

Theoretisch en empirisch kader

2.1 Inleiding

Verschillende perspectieven op leiderschap, gecombineerd met wijzigingen in de maatschappelijke en beleidsmatige context waarin de schoolleider functioneert, hebben geleid tot uiteenlopende conceptuele modellen van schoolleiderschap (zie Krüger & Witziers, 2003; Heck & Hallinger, 2000; Leithwood, Jantzi & Steinbach, 1999; De Maeyer & Rymenans, 2004; Scheerens & Witziers, 2005; Van Vilsteren, 1999). In een recente beschouwing van de stand van zaken in het onderzoeksveld, maken Heck en Hallinger (2005) zich zorgen over de diversiteit aan invalshoeken en het gebrek aan grondig empirisch onderzoek bij sommige benaderingen. Zij menen dat er nog te weinig gebruik wordt gemaakt van de krachtige statistische gereedschappen die beschikbaar zijn (zie ook Hallinger & Heck, 1998; Heck & Hallinger, 2000).

Door de verschillen in conceptuele en methodologische benadering is een theoretische en empirische inkadering van het onderzoek naar de invloed van schoolleiders op de leerprestaties niet eenvoudig. In hoofdstuk 1 is het onderzoek naar effectief schoolleiderschap in vogelvlucht beschreven. Dit hoofdstuk start met een overzicht van de verschillende perspectieven op schoolleiderschap en zoomt daarna verder in op schoolleiderschapsmodellen en de variabelen die daarbij van belang zijn. Hierbij staat de volgende onderzoeksvraag centraal:

Welke contextvariabelen, schoolleidersvariabelen, intermediaire variabelen en uitkomstvariabelen zijn relevant voor het in kaart brengen van de effectiviteit van de schoolleider en welke relaties zijn er tussen deze variabelen te onderscheiden?

In paragraaf 2 wordt ingegaan op pogingen die zijn ondernomen om een ordening aan te brengen in de verschillende benaderingen van (school-)leiderschap. In paragraaf 3 komen enkele belangrijke conceptuele modellen van schoolleiderschap aan de orde. Paragraaf 4 gaat over causale modellen van schoolleiderschap en de intermediaire variabelen die daarbij zijn gevonden. Het concurrerende waarden model (Quinn & Rohrbaugh, 1983) biedt de mogelijkheid om diverse conceptuele benaderingen van leiderschap te integreren. Dit model legt bovendien een verband tussen leiderschapsgedrag, de organisatiecultuur en de effectiviteit van de organisatie. Het concurrerende waarden model wordt besproken in paragraaf 5. Volgens het algemeen beschrijvingsmodel van onderwijskundige leiderschap van Bossert et al. (1982) zou de schoolleider invloed kunnen uitoefenen op de leerprestaties via het ‘pedagogisch onderwijskundig klimaat’ en de ‘onderwijsorganisatie’. In paragraaf 6 worden de begrippen schoolcultuur, schoolklimaat en schoolorganisatie toegelicht.

In paragraaf 7 komt het schooleffectiviteitsmodel aan de orde. Uitgaande van de in dit hoofdstuk beschreven theoretische en empirische achtergrond wordt in paragraaf 8 het onderzoeksmodel gepresenteerd.

2.2 Perspectieven op schoolleiderschap

Verschillende auteurs hebben geprobeerd om een overzicht te geven van de benaderingen die worden toegepast in het onderzoek naar schoolleiderschap (bijv. Heck & Hallinger, 2000, 2005; Leithwood & Duke, 1999; Ribbins & Gunter, 2002; Richmon & Allison, 2003; Van Vilsteren, 1999). Richmon en Allison (2003) concluderen dat de ontwikkeling van leiderschapstheorie niet zozeer wordt gehandicapt door een gebrek aan kwalitatief goed onderzoek, als wel door het ontbreken van overeenstemming over wat leiderschap inhoudt (zie ook Leithwood & Duke, 1999). Zij pleiten ervoor om aan de hand van een typologie van leiderschapstheorieën, verschillen en overeenkomsten in leiderschapsbenaderingen te identificeren, om te komen tot een beter begrip van leiderschap.

Een bekend overzicht van onderzoek naar (algemeen) leiderschap is samengesteld door Yukl (1998). Een van de typologieën die hij voorstelt (zoals beschreven door Richmon & Allison, 2003), is een indeling in persoonlijkheids-, gedrags-, situationele, macht-invloed en integrerende theorieën. In het onderzoek naar persoonlijkheidskenmerken van leiders wordt geprobeerd om specifieke kenmerken die leiders zouden bezitten te identificeren. Het gedragsonderzoek is gericht op specifieke gedragingen van leiders en het verband met variabelen zoals succes, motivatie en arbeidssatisfactie. Situationele of contingentie theorieën (Fiedler, 1967) gaan ervan uit dat er een relatie is tussen het gedrag van leiders en de situatie waarin de leider optreedt. Macht-invloed theorieën beschrijven het proces van beïnvloeding tussen leiders en volgers. In integrerende theorieën worden zowel de persoonlijkheidskenmerken en het gedrag van leiders beschouwd, als de percepties en attributies van volgers, in relatie tot de situatie. Yukl (1998) geeft aan dat het heel moeilijk is om leiderschapstheorieën te ordenen. Er is geen voor de hand liggende ‘correcte’ klassificatie, vanwege de diversiteit en complexiteit van de theorieën.

Krüger en Witziers (2003) hebben een poging ondernomen om leiderschapsbenaderingen te ordenen aan de hand van de ontwikkelingen in het denken over schoolleiderschap. De indeling die zij aanbrengen overlapt voor een deel de hierboven beschreven typologie van Yukl (1998). Hieronder volgt een samenvatting van het overzicht van Krüger en Witziers (2003, zie ook Scheerens & Witziers, 2005).

Onderzoek naar leiderschap was in het begin vooral gericht op persoonlijkheidskenmerken van leiders. In de eerste helft van de vorige eeuw begonnen onderzoekers zich af te vragen waarom sommigen beter dan anderen in staat zijn om leiderschap uit te oefenen. Toen deze benadering weinig licht bleek te werpen op de effectiviteit van leiderschap, verschoof de aandacht naar het gedrag van leiders. Ook schooleffectiviteitsonderzoek richtte zich in eerste instantie vooral op de taken en activiteiten van schoolleiders, aan de hand van gedragsobservaties. Verder was er vanaf het begin van de jaren zestig veel aandacht voor leiderschapsstijlen (zie ook Immegart, 1988). Leiderschapsstijl wordt wel omschreven als de consistente lijn die in de werkzaamheden van de leider te herkennen is (De Jong & Van Doorne-Huiskes, 1992). Om leiderschapsstijlen te kunnen onderscheiden, is in het schoolleidersonderzoek veel gebruik gemaakt van twee dimensies: ‘taakgerichtheid’ en ‘relatiegerichtheid’ (Blake & Mouton, 1964; Blake, Mouton, & Williams, 1981). Bij taakgerichtheid gaat het erom de organisatie zo doelmatig en effectief mogelijk te laten

functioneren. Relatiegerichtheid heeft als doel het welbevinden en het moreel van medewerkers te vergroten. Op het onderzoek naar persoonskenmerken en leiderschapsstijlen kwam veel kritiek vanwege de relatieve contextloosheid van de theorieën. Als reactie hierop ontstond een nieuwe benadering, waarbij meer rekening wordt gehouden met de context waarin het leiderschap wordt uitgeoefend. Deze benadering wordt situationeel leiderschap genoemd. De opvatting dat de schoolleider zijn/haar gedrag moet aanpassen aan de situatie waarin de school zich bevindt, is afkomstig uit de contingentietheorie (Fiedler, 1967; Mintzberg, 1979). Vervangend leiderschap is een benadering die in het verlengde ligt van situationeel leiderschap. Krüger en Witziers halen hierbij het werk aan van Kerr en Jermier (1978). Deze auteurs stellen dat er in een organisatie allerlei condities zijn die leiderschap vervangen of neutraliseren, zoals een hoge mate van competentie van medewerkers, of cohesieve taakgroepen. Sterke vaksecties zouden bijvoorbeeld kunnen fungeren als vervangers van schoolleiderschap (Pitner, 1981).

Terwijl er in het schooleffectiviteitsonderzoek aanvankelijk vooral veel aandacht was voor wat schoolleiders doen, kwam er in de jaren negentig steeds meer belangstelling voor wat effectieve schoolleiders denken. In het zogenaamde cognitieve perspectief op leiderschap gaat het niet meer alleen om de handelingen, maar vooral om de intenties die de schoolleiders hebben met deze handelingen. Effectief gedrag van een schoolleider is volgens dit cognitieve perspectief niet zozeer gebaseerd op een beheersing van deelvaardigheden, maar op een hogere orde integratie van vaardigheden. Een sterke visie van de schoolleider blijkt belangrijk te zijn voor effectief schoolleiderschap (zie bijv. Leithwood, 1995). Hierbij wordt er vanuit gegaan dat schoolleiders zich aanpassen aan de organisatie, waardoor hun denken en doen in de loop van de tijd veranderen. Krüger en Witziers verwijzen naar het overzichtsstudie van Hallinger en Heck (1998), om aan te geven dat modellen van schoolleiderschap waarin contextfactoren en schoolkenmerken zijn opgenomen, beter inzicht geven in de invloed van de schoolleider op de leerprestaties.

Uit de bovenstaande samenvatting kan worden opgemaakt dat in het leiderschapsonderzoek achtereenvolgens de nadruk lag op wie leiders zijn, wat ze doen en hoe ze denken. Krüger en Witziers (2003) menen dat een integratie van deze benaderingen is te vinden in een relatief nieuw perspectief op leiderschap, waarin wordt uitgegaan van rollen en competenties van leiders:

Dit wordt duidelijk uit de definiëring van het begrip competentie als de integratie van kennis, vaardigheden, attitudes, overtuigingen en persoonlijkheid. Het gaat dus niet langer óf om de persoonskenmerken, óf om de visie en attitudes, óf om de taken en handelingen, óf om de stijlen van schoolleiders. Om zich competenties eigen te maken dienen schoolleiders hun visies, hun persoonlijkheid, hun taken en activiteiten en hun stijl van leiderschap op elkaar af te stemmen, te integreren en te baseren op de kennis die zij verzamelen niet alleen uit de literatuur, maar ook uit de omgeving van de school en uit de school zelf. Dat is het uitgangspunt van het competentiedenken. (p. E3000-23)

De benadering van leiderschap in de vorm van rollen en competenties is te vinden in het werk van Robert Quinn c.s. (zie bijv. Quinn, 1998; Quinn & Cameron, 1983; Quinn et al., 2003). In het competing values framework (Quinn & Rohrbaugh, 1983) worden vier verschillende cultuurtypen onderscheiden, gebaseerd op verschillende waardenoriëntaties in organisaties. Quinn en zijn collega's gaan ervan uit dat de rollen die een leider moet vervullen, afhankelijk zijn van de cultuur van de organisatie, terwijl de cultuur weer gerelateerd is aan de levensfase van de organisatie. Quinn c.s. koppelen aan elke managementrol een aantal competenties

(Quin et al., 2003). Dit is volgens Krüger en Witziers (2003): “bijzonder relevant niet alleen omdat deze een geïntegreerde benadering van schoolleiderschap betreft, maar ook omdat de competentiebenadering oog heeft voor de professionele ontwikkeling van schoolleiders” (p. E3000-24). Het concurrerende waarden model van Quinn c.s. komt verder aan de orde in paragraaf 5. Eerst worden enkele belangrijke conceptuele modellen van schoolleiderschap nader beschreven (paragraaf 3) en wordt een overzicht gegeven van causale modellen voor de invloed van de schoolleider op de leerprestaties (paragraaf 4).

2.3 Conceptuele modellen van schoolleiderschap

In de discussie over schoolleiderschap spelen twee concepten een dominante rol: onderwijskundig leiderschap en transformationeel leiderschap (zie bijv. De Maeyer & Rymenans, 2004; Hallinger, 2003; Heck & Hallinger, 2000; Krüger, Witziers, Slegers & Imants, 1999; Leithwood, Jantzi & Steinbach, 1999). Beide concepten zijn in de loop van de tijd geëvolueerd. Volgens nieuwere opvattingen zouden de twee modellen als overlappend of complementair beschouwd kunnen worden (De Maeyer & Rymenans, 2004; Hallinger, 2003; Marks & Printy (2003). Een vorm van leiderschap waarin de twee modellen zijn geïntegreerd, zou een substantieel effect op de leerprestaties hebben (Marks & Printy, 2003). De nieuwste inzichten wijzen erop dat het relevanter is om uit te gaan van leiderschapscompetenties (zie bijv. Krüger & Witziers, 2003) of van een ‘repertoire of basic leadership practices’ (Leithwood, Day, Sammons, Harris & Hopkins, 2006b) in plaats van specifieke conceptuele modellen.

In hoofdstuk 1 is de ontwikkeling van de visie op onderwijskundig leiderschap al globaal geschetst. In de jaren zeventig en tachtig werd onderwijskundig leiderschap vooral opgevat als het superviseren van docenten. Beheersmatig management werd beschouwd als een aparte taak voor de schoolleider. Later werd onderwijskundig leiderschap breder geïnterpreteerd, als de gerichtheid van de schoolleider op de primaire processen in de school. De scheiding tussen onderwijskundige en beheersmatige taken viel weg (Imants, 1996). Hierbij kwam er steeds meer aandacht voor de invloed van de context waarbinnen de schoolleider functioneert. Het algemene beschrijvingsmodel van Bossert et al. (zie figuur 1.1.) geeft vorm aan deze visie op schoolleiderschap. Krüger et al. (1999) beschrijven dit model als volgt:

De schoolleider wordt geleid door een aantal achtergrondfactoren die per school en per schoolleider verschillen. De schoolleider is cultuurmanager via de strategieën die worden gevormd door handelingen gericht op het pedagogisch klimaat in de school. De schoolleiders is beheersmatig manager en organisatie-ontwikkelaar door handelingen die gericht zijn op de onderwijsorganisatie. Via het managen van de organisatie en de cultuur is de schoolleider een onderwijskundig leider door bij alles wat hij doet de primaire processen in de gaten te houden. (p. C4120-10/C4120-11)

Verschillende elementen uit het beschrijvingsmodel van Bossert et al. (Dwyer et al., 1985; Dwyer, Barnett & Lee, 1987) zijn verder uitgewerkt in speciale benaderingen van leiderschap (zie Krüger et al., 1999). Het belang van de visie van de schoolleider komt bijvoorbeeld terug in het constructivistische of cognitieve perspectief op schoolleiderschap. Binnen dit cognitieve perspectief staan meta-cognitieve denk- en afwegingsprocessen van schoolleiders centraal (Hallinger, Leithwood & Murphy, 1993; Leithwood, 1995).

In een wijdere opvatting van onderwijskundig leiderschap die integraal leiderschap wordt genoemd, wordt ervan uitgegaan dat de schoolleider beheersmatige en onderwijskundige taken combineert vanuit een overkoepelende visie (Krüger et al., 1999). Hierbij gaat het niet alleen om een sturing van gedrag, maar juist ook om een sturing van mentale processen in de school.

Het idee dat een schoolleider indirect de leerresultaten kan beïnvloeden via het ontwikkelen van de organisatie en de cultuur is verder uitgewerkt in het concept van transformationeel leiderschap. Volgens deze benadering moet de schoolleider zorgen voor een cultuuromslag zodat een werkomgeving ontstaat waarin docenten optimaal samenwerken en zich sterk identificeren met de gezamenlijk ontwikkelde doelstellingen van de school (Krüger et al., 1999). Het ontwikkelen van een (gezamenlijke) visie speelt dus zowel een rol bij (integraal) onderwijskundig leiderschap, als bij transformationeel leiderschap (zie ook Hallinger, 2003). De transformationele leider is vooral gericht op de betrokkenheid, motivatie en capaciteiten van docenten en andere medewerkers in de schoolorganisatie, met als doel het vergroten van het vermogen van de schoolorganisatie om te veranderen en te innoveren (Geijsel, Slegers & van den Berg, 1999; Leithwood & Jantzi, 1999b, 2000a; Slegers, 1999).

Zowel de brede opvatting van onderwijskundig leiderschap, als de ideeën over transformationeel leiderschap, sluiten aan bij het relatief nieuwe concept van de school als 'lerende organisatie' (zie bijv. Krüger et al., 1999; Leithwood, Jantzi & Steinbach, 1999; Marks, Louis & Printy, 2000; Mulford & Silins, 2003). De onderliggende gedachte bij de lerende organisatie is dat een succesvolle organisatie continu leert en zich aanpast aan veranderingen in de omgeving. Het gaat hierbij niet alleen om aanpassingen die nodig zijn om de bestaande doelen te halen, maar ook om vernieuwing van de doelen (Krüger et al., 1999). Silins, Zarins en Mulford (2002) onderscheiden op basis van een literatuuronderzoek zeven dimensies waarmee scholen gekarakteriseerd kunnen worden als lerende organisaties:

Learning organisations were defined as schools that: employ processes of environmental scanning; develop shared goals; establish collaborative teaching and learning environments; encourage initiatives and risk taking; regularly review all aspects related to and influencing the work of the school; recognise and reinforce good work; and, provide opportunities for continuing professional development. (p. 24)

In de lerende organisatie is er ruimte voor 'gespreid leiderschap' of 'leiderschap als organisatorische kwaliteit' (Ogawa & Bossert, 1995; Pounder, Ogawa & Adams, 1995). Het idee is dat leiderschap kan worden uitgeoefend door allerlei personen op verschillende niveaus in de organisatie. Hier spelen aspecten van transformationeel leiderschap een rol. Door het delegeren van leidinggevende taken en bevoegdheden kan de betrokkenheid van docenten vergroot worden. De empirische onderbouwing van de effectiviteit van gespreid leiderschap is nog beperkt. Verondersteld wordt dat meer gecoördineerde vormen van gespreid leiderschap betere resultaten opleveren (zie Leithwood, Day, Sammons, Harris & Hopkins, 2006). Ook wanneer onderwijskundige taken zijn gedelegeerd, blijft de schoolleider integraal onderwijskundig leider, maar dan meer in de rol van coach. De visie of de missie van de school vormt het toetsingskader voor de onderwijskundige taken die door verschillende mensen worden uitgeoefend (Krüger, Witziers, Slegers & Imants, 1999).

Op basis van een analyse van de conceptuele en empirische ontwikkeling van de modellen van onderwijskundig leiderschap en transformationeel leiderschap concludeert Hallinger (2003) dat (inmiddels) de overeenkomsten tussen de twee modellen groter zijn dan de

verschillen, in ieder geval met betrekking tot de belangrijkste schoolverbeteringsactiviteiten. Volgens Hallinger ligt in beide modellen de nadruk op:

- een gedeelde visie creëren binnen de school
- een klimaat van hoge verwachtingen ontwikkelen en een schoolcultuur die is gericht op de verbetering van onderwijs en leren
- een beloningsstructuur creëren waarmee de doelen die zijn vastgesteld voor personeel en leerlingen worden weerspiegeld
- voorzien in een breed scala aan activiteiten voor de intellectuele stimulering en de ontwikkeling van het personeel
- zichtbaar zijn in de school en een modelfunctie vervullen met betrekking tot de waarden van de organisatie.

Drie belangrijke verschillen tussen onderwijskundig leiderschap en transformationeel leiderschap die door Hallinger (2003) werden geïdentificeerd zijn:

1. *Een gerichtheid op eerste orde vs. tweede orde veranderingen*
Onderwijskundig leiderschap zou met name gericht zijn op het beïnvloeden van omstandigheden die een directe impact hebben op de kwaliteit van het curriculum en de instructie. Dit zijn eerste orde veranderingen. Transformationeel leiderschap zou de capaciteit van anderen moeten vergroten om eerste orde effecten te bewerkstelligen. Hierbij gaat het om tweede orde veranderingen.
2. *Een coördinatie en controle strategie vs. een 'empowerment' strategie*
Bij onderwijskundig leiderschap ligt de nadruk op de handhaving van de status quo in de organisatie door coördinatie en controle, terwijl bij transformationeel leiderschap wordt voorgebouwd op ambities en ideeën van het personeel, dit wordt 'empowerment' genoemd.
3. *Individueel vs. gespreid leiderschap*
Onderwijskundig leiderschap meestal gelokaliseerd bij de schoolleider, terwijl bij transformationeel leiderschap wordt uitgegaan van een spreiding van leiderschap over meer personen.

Hallinger (2003) wijst op twee mogelijke benaderingen om het onderwijskundige en transformationele leiderschapsmodel met elkaar te verenigen. De eerste benadering is gebaseerd op de uitkomsten van een onderzoek van Marks en Printy (2003). Zij tonen aan dat een geïntegreerde vorm van transformationeel en onderwijskundig leiderschap een 'substantiële' invloed heeft op de leerresultaten. De tweede benadering komt voort uit de contingentie theorie (Fiedler, 1967). Hierbij verwijst Hallinger naar het werk van Bossert, Dwyer, Rowan en Lee (1982), die op basis van een analyse van kwantitatieve onderzoeksresultaten tot de conclusie komen dat bepaalde leiderschapsactiviteiten verschillende effecten hebben in verschillende organisatorische omgevingen. Hallinger veronderstelt dat het type leiderschap dat nodig is afhangt van de situatie waarin de school zich bevindt (zie ook Hallinger & Murphy, 1985). Het idee is dat leiderschap kan worden opgevat als een ontwikkelingsproces. Wanneer een school zich in moeilijke omstandigheden bevindt en een verbeteringsproces ingaat, zou in eerste instantie een meer directieve vorm van onderwijskundig leiderschap nodig kunnen zijn, om de organisatie in de gewenste richting te kunnen sturen. Op termijn is echter een grotere participatie van het personeel wenselijk bij het leiden van de veranderingsprocessen, om langdurige effecten te kunnen bereiken (Hallinger, 2003). De opvatting dat de geschiktheid van bepaalde leiderschapsactiviteiten is gekoppeld aan de ontwikkelingsfase waarin een organisatie zich bevindt, is ook terug te vinden in de ideeën van Quinn (1998) over het concurrerende waarden model en effectief management (zie paragraaf 5).

In de afgelopen jaren publiceerde Kenneth Leithwood, in samenwerking met verschillende andere onderzoekers, diverse documenten over de uitkomsten van een omvangrijk literatuuronderzoek, dat werd uitgevoerd om als uitgangspunt te dienen voor een (toekomstig) grootschalig empirisch onderzoek naar een aantal sterke ‘claims’ over succesvol schoolleiderschap (Leithwood, Day, Sammons, Harris & Hopkins, 2006a, 2006b, Leithwood & Riehl, 2003a, 2003b, 2005; Leithwood, Seashore Louis, Anderson & Wahlstrom, 2004). Leithwood en zijn mede-auteurs erkennen dat er de nodige kritiek is op het onderzoek naar schoolleiderschap, waarbij onder meer wordt gewezen op een gebrek aan programmatisch onderzoek, de grote variëteit aan onderzoeksmethoden en op het ontbreken van voldoende bewijsmateriaal van een zodanige kwaliteit dat het kan dienen als leidraad voor beleid en praktijk (zie bijv. Firestone & Riehl, 2005; Heck & Hallinger, 2005). Zij benadrukken echter dat er, ondanks deze duidelijke tekortkomingen, toch voldoende betrouwbare informatie beschikbaar is om een aantal conclusies over schoolleiderschap te kunnen onderbouwen.

Een van de claims waarvoor de meeste empirische ondersteuning bestaat is: “Almost all successful leaders draw on the same repertoire of basic leadership practices.” (Leithwood et al. 2006b, p. 6). Dit repertoire van leiderschapsactiviteiten is in te delen in vier categorieën (zie ook Leithwood et al. 2006a, NCSL, 2007):

1. *Visie ontwikkelen en richting geven*: identificeren en verwoorden van een visie, creëren van een gedeeld belang, hoge verwachtingen scheppen van prestaties, bevorderen van de acceptatie van groepsdoelen, monitoren van organisatieprestaties en communiceren.
2. *Mensen begrijpen en ontwikkelen*: intellectuele stimulans aanbieden, individuele begeleiding verschaffen en modelgedrag vertonen. Voortbouwen op de kennis en vaardigheden van docenten en ander personeel om schooldoelen te bereiken. De integratie van persoonlijke kwaliteiten en functioneel gedrag ontwikkelen, coachen, erkennen en belonen.
3. *Herontwerpen van de organisatie*: bouwen aan culturen en processen waarin wordt samengewerkt, managen van de omgeving en van de werkomstandigheden, creëren en onderhouden van productieve relaties met ouders en de gemeenschap, de school verbinden met de wijdere omgeving.
4. *Managen van het onderwijs- en leerprogramma*: creëren van een productieve werkomgeving voor zowel docenten als leerlingen, organisatorische stabiliteit bevorderen, waarborgen van effectief leiderschap waarbij het leren centraal staat, toewijzen van docenten en ondersteunend personeel om het curriculum uit te voeren, monitoren van schoolactiviteiten en –prestaties.

Over de contextgevoeligheid van leiderschapsactiviteiten wordt door Leithwood en collega’s het volgende gesteld (zie Leithwood et al. 2006b): “Without doubt, successful leaders are sensitive to context, but this does not mean they use qualitatively different practices in every different context. It means, rather, that they apply contextually sensitive combinations of the basic leadership practices described above.” (p. 8). Ter illustratie van deze opvatting van contextgevoeligheid, schetsen Leithwood et al. verschillen in de wijze waarop de vier categorieën van leiderschapsactiviteiten worden toegepast in verschillende stadia van een schoolverbeteringsproces. Net als Hallinger (2003) wijzen Leithwood et al. er bijvoorbeeld op dat de schoolleider in een vroeg stadium van een schoolverbeteringsproces het voortouw moet nemen bij de ontwikkeling van een visie en het uitzetten van de richting, terwijl in een latere fase meer personeel betrokken moet zijn bij het bepalen van de koers. Het begrijpen en ontwikkelen van mensen is in alle fasen van het schoolverbeteringsproces van belang. Herstructurering van de organisatie zal in eerste instantie gericht moeten zijn op een betere

communicatie, om later een cultuurverandering te kunnen bewerkstelligen. Verder zal de schoolleider er in een vroeg stadium van het schoolverbeteringsproces bijvoorbeeld voor moeten zorgen dat afgeweken kan worden van vastgesteld personeelsbeleid, om personeel aan te kunnen stellen met capaciteiten om dit proces op te starten, terwijl later weer moet worden uitgegaan van duidelijke regels bij het aanstellen van personeel.

Leithwood et al. (2004) merken op dat de generieke functies van leiderschap vaak worden gemaskeerd door de vele etiketten die worden gebruikt voor verschillende vormen en stijlen van leiderschap. De toepassing van een generiek model van leiderschap zou dus wellicht meer inzicht kunnen bieden in de effectiviteit van schoolleiders. Het concurrerende waarden model (zie bijv. Quinn et al., 2003) biedt een dergelijk generiek model, terwijl het tevens een zekere mate van contextafhankelijkheid van leiderschapsgedrag impliceert (zie paragraaf 5). In de volgende paragraaf komen causale modellen van het effect van schoolleidersactiviteiten op de leerprestaties aan de orde.

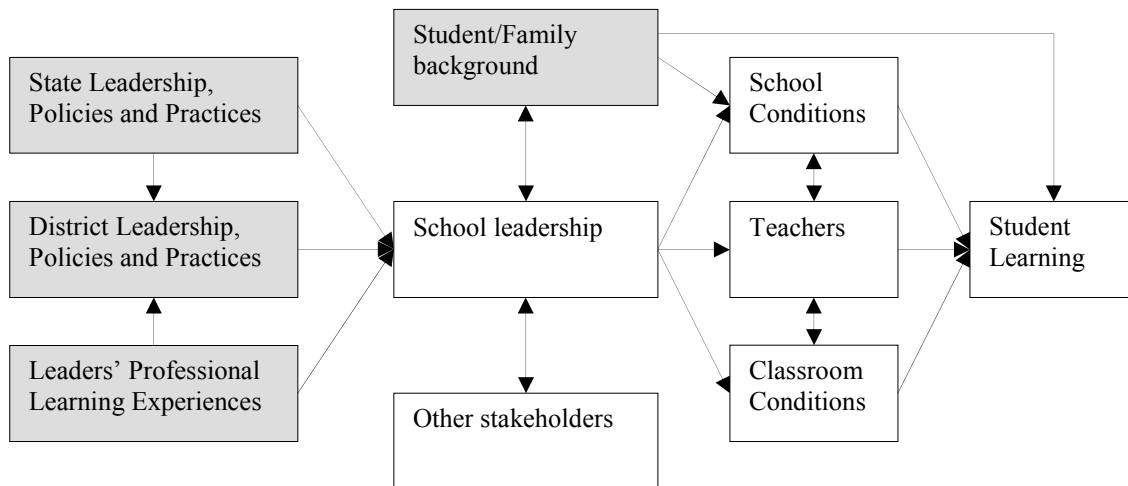
2.4 Causale modellen van schoolleiderschap

Volgens de uitkomsten van de overzichtsstudie van Hallinger en Heck (1998) zou de invloed van de schoolleider op de leerprestaties vooral indirect moeten verlopen, via intermediaire factoren. Witziers, Bosker en Krüger (2003) komen op basis van een kwantitatieve meta-analyse tot een overeenkomstige conclusie (zie hoofdstuk 1). De vraag is nu welke intermediaire variabelen een rol spelen bij het uitoefenen van schoolleiderschap.

Het algemene beschrijvingsmodel van onderwijskundig leiderschap van Bossert et al. (1982, zie figuur 1.1) biedt aanknopingspunten om schoolleidersgedrag te verklaren en de effecten hiervan te analyseren. In het model van Bossert et al. zijn het pedagogisch klimaat en de onderwijsorganisatie opgenomen als intermediaire factoren voor de invloed van de schoolleider op de leerprestaties (Bossert et al., 1982; Dwyer et al., 1985; Dwyer, Barnett & Lee, 1987). De visie en de ervaring van de schoolleider vormen in dit model verklarende factoren voor de handelingen van de schoolleider, gecombineerd met factoren uit de omgeving van de school en wettelijke kaders en regelgeving.

Leithwood, Seashore Louis, Anderson en Wahlstrom (2004) gebruikten voor een overzicht van de stand van zaken op het gebied van het onderzoek naar schoolleiderschap een raamwerk met tien variabelen. Dit model, dat is afgebeeld in figuur 2.1, kan worden gezien als een uitbreiding van het algemene beschrijvingsmodel van Bossert et al.

Volgens dit raamwerk is er een interactie tussen leiderschap, beleid en activiteiten op het niveau van de staat en op het niveau van het district, terwijl deze factoren een directe invloed uitoefenen op het gedrag van schoolleiders. In het model van Bossert et al. zijn soortgelijke contextfactoren terug te vinden in 'wettelijke kaders en regelgeving' en 'omgeving van de school'. Deze contextfactoren hebben volgens Leithwood et al. (2004) eveneens invloed op de omstandigheden in de school en in de klas en op de professionele gemeenschap van docenten. De activiteiten van de schoolleiders worden ook beïnvloed door diverse andere partijen, zoals media, vakbonden, beroepsverenigingen, maatschappelijke organisaties en bedrijven en daarnaast door de eigen professionele ervaring. De andere partijen zijn in het model van Bossert et al. niet apart opgenomen.



Figuur 2.1 Onderzoeksraamwerk van Leithwood et al. (2004) voor het verband tussen leiderschap en leren. De relaties die zijn aangegeven tussen de variabelen zijn slechts illustratief. Het schema geeft geen volledig overzicht van de complexe relaties die er werkelijk bestaan tussen de verschillende variabelen.

Achtergrondfactoren van de leerlingen en hun families hebben een belangrijke invloed op de meeste variabelen in het raamwerk van Leithwood et al. Deze factoren hebben bijvoorbeeld zowel een effect op de middelen die beschikbaar zijn voor de school, als op de activiteiten van de schoolleider, het onderwijsproces in de klas en het sociale kapitaal (voordelen door relaties met anderen) dat de leerling ter beschikking heeft. Leithwood et al. (2004) wijzen erop dat niet alleen de sociaal economische status van de leerling, maar ook de ‘family educational culture’ een belangrijke rol speelt bij het behalen van leerprestaties. In het model van Bossert et al. zijn achtergrondfactoren van de leerlingen buiten beschouwing gelaten.

In het raamwerk van Leithwood et al. omvat schoolleiderschap zowel formeel als informeel leiderschap. Dit gedistribueerde leiderschap (zie bijv. Leithwood et al., 2006a) geeft zowel vorm aan de schoolomstandigheden zoals doelen, cultuur en structuur, als aan de omstandigheden in de klas, zoals de inhoud van het onderwijs, de onderwijsvormen die worden ingezet en de grootte van de klas. De schoolleider heeft ook invloed op de mate waarin docenten een professionele gemeenschap ervaren, maar hierbij spelen nog meer factoren een rol. De factoren op schoolniveau en de omstandigheden in de klas, de professionele gemeenschap van docenten en de achtergrond van studenten zijn direct van invloed op de leerprestaties.

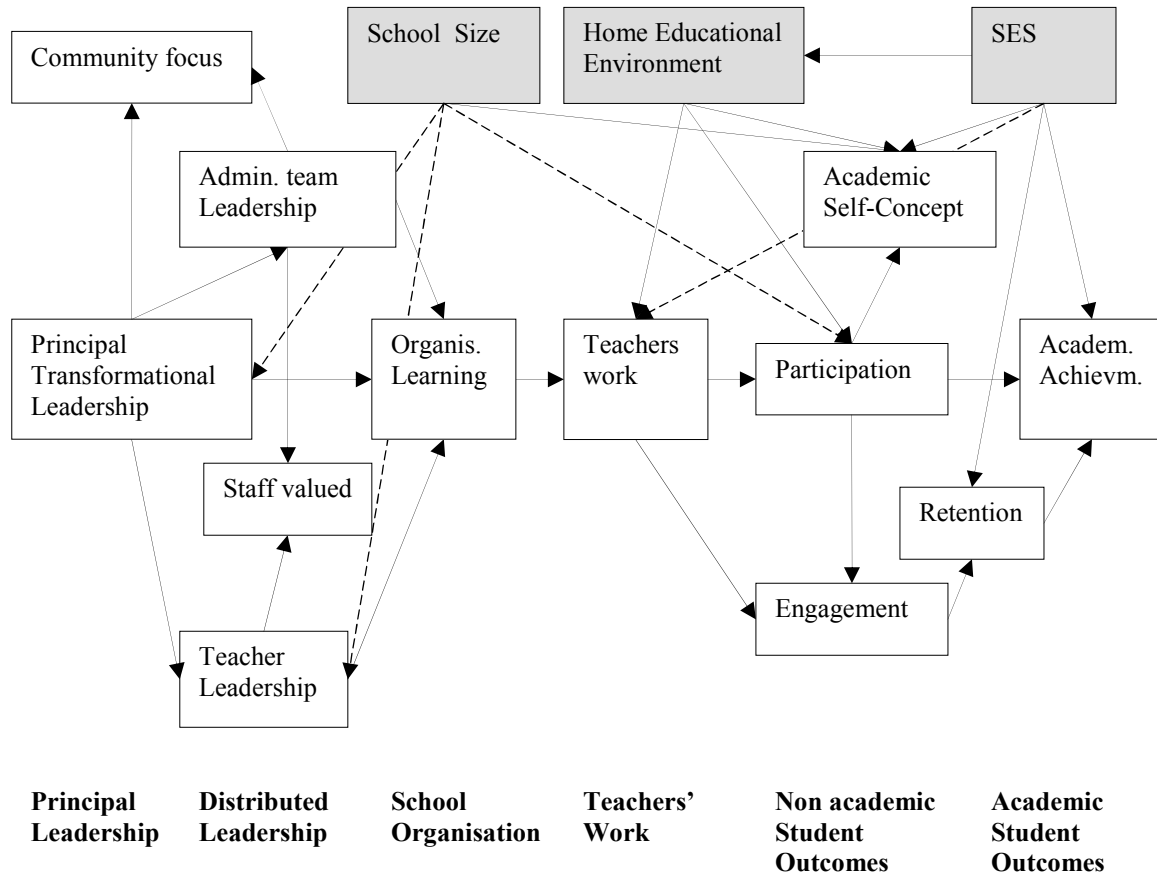
Hoewel het raamwerk van Leithwood et al. is gebaseerd op empirische gegevens, is het niet in zijn geheel als model getoetst. Om indirecte effectmodellen van de invloed van de schoolleider op de leerprestaties te kunnen toetsen, zijn causale analysetechnieken nodig. Zowel Hallinger en Heck (1998), als Witziers, Bosker en Krüger (2003), merken op dat er nog relatief weinig onderzoek is gedaan naar indirecte effecten van schoolleidersactiviteiten op de leerprestaties. Daarbij is er weinig consistentie in de intermediaire variabelen die zijn geïdentificeerd. In tabel 2.1 staat een overzicht van significante intermediaire variabelen die zijn gevonden in dit soort onderzoek (gegevens gedeeltelijk ontleend aan Hallinger & Heck, 1998; Scheerens & Witziers, 2005, p. 72; Witziers et al., 2003).

Tabel 2.1 Overzicht van onderzoeken waarin significante indirecte effecten van schoolleiderschap op leerprestaties werden gevonden, met de toegepaste leiderschapsmodellen en de significante intermediaire variabelen die werden geïdentificeerd

Referentie	Leiderschapsmodel	Significante intermediaire variabelen
Heck, Larsen en Marcoulides, 1990	Governance ¹	School climate School Instructional organization
Hill, Rowe and Holmes-Smith, 1995	Leadership	Teacher student interactions Professional climate
Hallinger, Bickman en Davis, 1996	Principal Instructional Leadership	A clear school mission Students' opportunity to learn Teachers' expectations
Hallinger en Heck, 1998	Instructional leadership	Principal supervisory activities
Bosker, De Vos en Witziers, 2000	Educational leadership	Teachers' job satisfaction Teachers' achievement orientation Evaluation and feedback practices
Mulford en Silins, 2003	Transformational and distributive leadership	Organisational learning Teachers' work Students' participation Students' engagement
De Maeyer en Rymenans, 2004	Geïntegreerd leiderschap	Prestatiegericht klimaat
Leithwood et al., 2006a	Total Leadership (distributed)	Teachers' capacity Teachers' motivation and commitment Teachers' perceived working conditions

¹ Hier gaat het eigenlijk om effecten van algemene leiderschapsactiviteiten (governance) op specifieke leiderschapsactiviteiten (intermediaire variabelen) en van specifieke activiteiten op leerprestaties.

De onderzoeken in tabel 2.1 zijn chronologisch gerangschikt. Hierdoor is de ontwikkeling in schoolleiderschapsmodellen van onderwijskundig leiderschap in nauwe zin (instructional leadership) naar geïntegreerd of gedistribueerd leiderschap zichtbaar (zie ook paragraaf 3). Verder is uit tabel 2.1 op te maken dat in de indirecte effectmodellen die werden onderzocht, intermediaire variabelen op verschillende niveaus zijn opgenomen: schoolniveau, klassen- of docentenniveau en leerlingenniveau. De meeste modellen omvatten slechts een of twee niveaus van intermediaire variabelen, alleen het model van Mulford en Silins (2003) omvat zowel intermediaire factoren op schoolniveau, als op klassenniveau en op leerlingenniveau. Dit model is afgebeeld in figuur 2.2. Het model van Mulford en Silins is ontwikkeld in het *Leadership for Organisational Learning and Student Outcomes (LOLSO) Research Project*, dat is uitgevoerd in het voortgezet onderwijs in Australië (zie o.a. Mulford, 2003, 2006; Mulford & Silins, 2003; Silins & Mulford, 2002, 2004; Silins, Mulford & Zarins, 2002).



Figuur 2.2 Compleet LOLSO model van Mulford en Silins. → = positieve relatie, ---> = negatieve relatie (Mulford & Silins, 2003)

In het LOLSO-model wordt uitgegaan van transformationeel en gedistribueerd leiderschap. Transformationeel leiderschap wordt bepaald aan de hand van de mate waarin de activiteiten van de schoolleider zijn gericht op (Mulford & Silins, 2003):

1. Gedeelde visie en doelen
2. Cultuur van onderling vertrouwen en flexibiliteit
3. Structuren voor participatieve besluitvorming
4. Intellectuele stimulering
5. Individuele ondersteuning
6. Hoge verwachtingen van prestaties

Transformationeel en gedistribueerd leiderschap blijken organisatieleren ('organisational learning') te bevorderen. De variabele 'organisational learning' wordt gemeten door middel van scores op vier factoren (Mulford & Silins, 2003):

1. Een klimaat van onderling vertrouwen en samenwerking
2. Het nemen van initiatieven en risico's
3. Een gedeelde missie die steeds in het oog wordt gehouden
4. Relevante en voortdurende professionele ontwikkeling

Uit het LOLSO-model is op te maken dat er positieve relaties bestaan tussen zowel transformationeel als gedistribueerd leiderschap en de omgeving van de school ('community focus'), maar dit heeft geen effect op de schoolorganisatie of de leerprestaties.

Het feit dat het personeel zich gewaardeerd voelt heeft wel effect op organisatieleren. Organisatieleren vormt in het model een belangrijke intermediaire variabele tussen leiderschap en het werk van de docenten.

De variabele ‘teachers’ work’ wordt in het LOLSO-model bepaald aan de hand van de percepties van leerlingen. Hierbij gaat het om de waardering van de lessen, de variëteit aan activiteiten, de organisatie van de lessen, het bespreken van werk met de leerlingen, hoge verwachtingen van de prestaties en de mate van uitdaging. Positieve percepties van het werk van docenten bevorderen het academische zelfbeeld (‘academic self-concept’) van leerlingen, de participatie (‘participation’) van leerlingen in de klas en bij andere schoolactiviteiten en hun identificatie met de school (‘engagement’). Het academische zelfbeeld van de leerling geeft onder meer aan hoe tevreden de leerling is over zijn/haar eigen mogelijkheden en prestaties. Deze factor blijkt niet van invloed te zijn op de leerprestaties, maar is waarschijnlijk wel gerelateerd aan succes in de latere loopbaan (Feinstein, 2000; Mulford, 2003). De participatie van de leerling is direct gerelateerd aan de leerprestaties en de identificatie met school indirect, via een beperktere schooluitval (‘retention’).

Mulford en Silins vonden drie belangrijke contextfactoren voor de invloed van schoolleiderschap op de leerprestaties: schoolgrootte, ‘home educational environment’ (onderwijsondersteuning van thuis) en sociaal economische status (zie Mulford, 2003, 2006; Mulford & Silins, 2003). De sociaal economische status (SES) van leerlingen is positief gerelateerd aan de ‘home educational environment’ (mate waarin het schoolwerk wordt ondersteund, bespreken van belangrijke gebeurtenissen in de wereld) en aan het academische zelfbeeld, ‘retention’ (beperking van schooluitval) en de leerprestaties. Opmerkelijk is dat er een negatief verband is tussen de sociaal economische status van de leerlingen en de perceptie van het werk van docenten. Aan de andere kant blijkt de ‘home educational environment’ een sterkere relatie te hebben met het academische zelfbeeld dan de sociaal economische status en daarnaast ook nog een sterke positieve relatie met participatie en de perceptie van het werk van docenten. Tenslotte blijkt een grotere school een beter academisch zelf-concept op te leveren, maar verder is deze omgeving niet gunstig voor transformationeel en gedistribueerd leiderschap en ook niet voor de participatie van leerlingen. In het Australische onderzoek zijn grote scholen overigens tevens grootstedelijke scholen (Mulford, 2003).

De variabelen ‘participation’ en ‘engagement’ in het LOLSO-model zijn geconceptualiseerd door Leithwood et al. (Leithwood, Dart, Jantzi & Steinbach, 1993) en gebaseerd op het werk van Finn (1989). Leithwood paste het ‘Student Participation and Identification’ (SPI) model van Finn aan en breidde het uit naar het concept ‘engagement’ (verbondenheid met de school), met zowel een gedragscomponent als een affectieve component. De gedragscomponent wordt gevormd door de participatie van leerlingen in schoolactiviteiten binnen en buiten de klas (‘participation’). De affectieve component komt overeen met de mate waarin leerlingen zich identificeren met hun school en zich er thuisvoelen (‘identification’). Enigszins verwarrend is dat de variabele ‘identification’ van Leithwood in het LOLSO-model ‘engagement’ wordt genoemd (zie figuur 2.2).

Leithwood en Jantzi (1999a, 1999b, 2000a, 2000b) kozen student engagement als afhankelijke variabele in hun onderzoek naar transformationeel en gedistribueerd leiderschap, om inzicht te krijgen in de effecten van leiderschap op andere uitkomstvariabelen dan leerprestaties. Uit eerder onderzoek bleek dat uitval van leerlingen de laatste stap is in een proces waarbij de leerlingen steeds minder participeren in de school en zich steeds minder identificeren met de school (Finn, 1989). Aan de patronen van participatie en identificatie van

leerlingen is al heel vroeg te herkennen of de leerling tot een risico-groep behoort (Lloyd, 1978). Verder is het volgens Leithwood en Jantzi (2000b) aannemelijk dat student engagement een predictor is van leerprestaties (Finn & Cox, 1992; Bredschneider, 1993; Dukelow, 1993). Mulford (2003) geeft aan dat bijvoorbeeld ook PISA onderzoek (OECD, 2001) aanwijzingen opleverde dat leerlingen die het leuk vinden op school beter presteren dan leerlingen die dat niet doen.

Leithwood en Jantzi (2000b) vonden in hun onderzoek naar de invloed van schoolleiderschap op student engagement alleen een significant indirect effect van 'principal leadership' op de identificatie van leerlingen met de school. Het effect van leiderschap op participatie was kleiner dan het effect op identificatie en niet significant. Er werden geen significante effecten van 'teacher leadership' op participatie en identificatie gevonden. In een recent onderzoek toonden Leithwood en zijn collega's overigens aan dat 'total leadership', ofwel gecombineerd leiderschap vanuit vele mogelijke bronnen, een significant indirect effect heeft op leerprestaties (Leithwood & Mascall, 2008; zie ook Leithwood et al., 2006a) in tegenstelling tot transformationeel leiderschap van de schoolleider alleen (Leithwood & Jantzi, 2006). Het indirecte effect van 'total leadership' op de leerprestaties verloopt via de capaciteit van docenten, de perceptie van de werkomstandigheden door docenten en de motivatie en het 'commitment' van docenten (zie ook tabel 2.1).

In het LOLSO-project werd eveneens geconstateerd dat schoolleiderschap een groter effect heeft op de identificatie van leerlingen dan op de participatie (Silins & Mulford, 2003). Verder vonden Silins en Mulford dat leiderschap, zowel van de formele schoolleider (principal leadership), als van andere functionarissen in de schoolleiding en de docenten (administrative team en teacher leadership), alleen een indirect effect heeft op de uitkomstvariabelen, door middel van organisatieleren en het werk van de docenten (Silins & Mulford, 2004). Hoe hoger de docenten scoorden op de vier dimensies van organisatieleren (zie hierboven), hoe positiever het werk van de docenten werd gewaardeerd door de leerlingen.

Uit het bovenstaande wordt duidelijk dat er uiteenlopende causale modellen van schoolleiderschap zijn getoetst, met verschillende intermediaire variabelen op het niveau van de school, klas en leerling. Het LOLSO model van Mulford en Silins (2003), dat gedeeltelijk is gebaseerd op het werk van Leithwood c.s., is een ver uitgewerkt model dat drie niveaus van intermediaire variabelen omvat. Dit model gaat uit van transformationeel en gedistribueerd leiderschap, waarbij organisatieleren een belangrijke intermediaire variabele is voor de invloed van de schoolleider op de leerprestaties. Het is nu de vraag of de specifieke vormen van schoolleiderschap en schoolorganisatie in dit causale schema ook vervangen kunnen worden door een generiek model van leiderschap en organisatie-effectiviteit, zoals het concurrerende waarden model.

2.5 Het concurrerende waarden model

In 1983 werd door Quinn en Rohrbaugh het *competing values framework* (concurrerende waarden model) gepresenteerd als een raamwerk van effectiviteitscriteria voor organisatie-analyse. Het model is gebaseerd op de uitkomsten van een onderzoek naar de opvattingen van theoretici en onderzoekers over organisatie-effectiviteit (Quinn & Rohrbaugh, 1983). Er werd dus niet uitgegaan van de operationele structuur van organisaties, maar van de cognitieve structuur van de organisatiedeskundigen. Organisatie-effectiviteit is een abstract begrip, waarvoor verschillende modellen gehanteerd worden. Omdat het moeilijk is om eenduidig

vast te stellen wat organisatie-effectiviteit precies is, is het ook moeilijk te meten. Quinn en Rohrbaugh (1983) zochten hiervoor een oplossing door de opvattingen van experts te vergelijken en te ordenen met behulp van multidimensionele schaling. Uit hun onderzoek bleek dat organisatiedeskundigen een impliciet theoretisch raamwerk delen. Nadat het concurrerende waarden model was gedefinieerd (Quinn & Rohrbaugh, 1983), werd het in de loop der jaren verder ontwikkeld en uitgewerkt. Het framework wordt niet alleen geschikt geacht om de organisatie-effectiviteit vast te stellen, maar ook om de cultuur van een organisatie te analyseren en om leiderschapsgedrag te beschrijven (zie bijv. Cameron & Quinn, 1999; Denison, Hooijberg & Quinn, 1995; Quinn, 1998; Quinn & Cameron, 1983; Quinn et al. 2003 ; Quinn & Spreitzer, 1991).

In het theoretisch kader van Quinn en Rohrbaugh (1983) zijn in feite vier modellen van organisatie-effectiviteit geïntegreerd die in de loop van de vorige eeuw min of meer achtereenvolgens zijn opgekomen (zie Quinn et al., 2003):

- Het *rationele doel* model, gericht op maximale output: het idee hierbij is dat een duidelijke richting leidt tot productieve resultaten. De nadruk ligt op verduidelijking van doelen, rationele analyse en handelend optreden. Bij dit model is het organisatieklimaat economisch rationeel: de eindresultaten zijn het belangrijkste.
- Het *interne proces model*, gericht op consolidatie en continuïteit: hierbij geldt de overtuiging dat routines tot stabiliteit leiden. De nadruk ligt op het vastleggen van verantwoordelijkheden, metingen en documentatie. Bij dit model is er een hiërarchisch klimaat in de organisatie.
- Het *'human relations'* model, gericht op ontwikkeling van human resources: er wordt vanuit gegaan dat betrokkenheid leidt tot inzet. De nadruk ligt op participatie, oplossen van conflicten en consensus bereiken. Het organisatieklimaat is teamgericht.
- Het *open systeem* model, gericht op expansie en verandering: continue aanpassing en innovatie leiden tot het verwerven en onderhouden van externe middelen. Hierbij gaat het om politieke aanpassing, een creatieve probleemaanpak, innovatie en management van verandering. Het organisatieklimaat is innovatief en flexibel.

De vier modellen lijken tegengesteld aan elkaar, maar ze vullen elkaar ook aan. De modellen kunnen daarom worden beschouwd als elementen binnen een breder kader: ze vormen vier deelgebieden van het begrip organisatie-effectiviteit (zie Quinn et al. 2003). Omdat de onderliggende waarden van de modellen elkaar in eerste instantie lijken tegen te spreken, is het theoretisch kader 'competing values framework' genoemd (Quinn & Rohrbaugh, 1983).

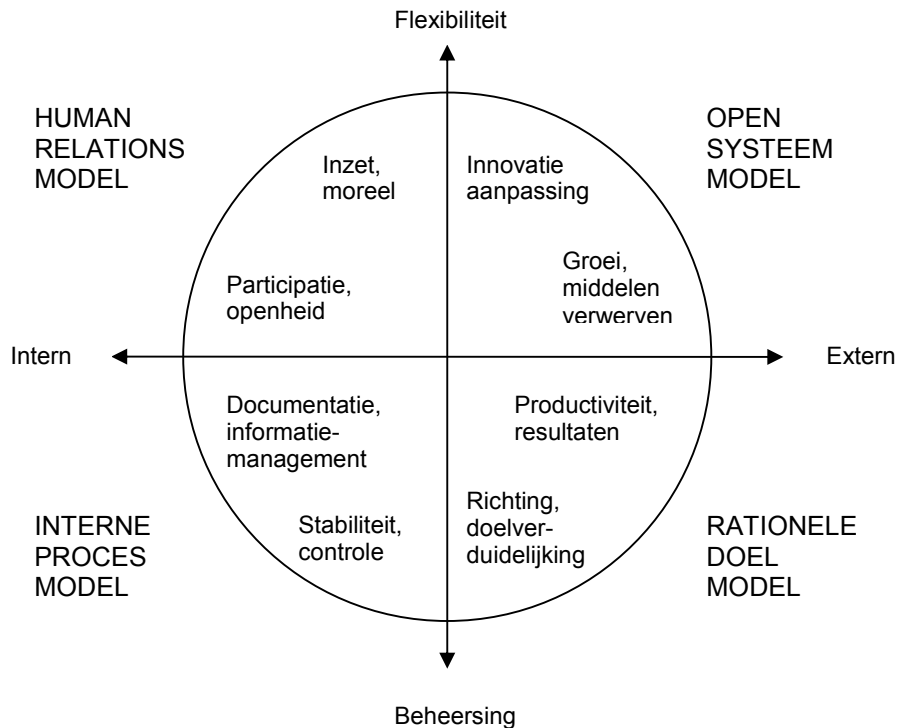
In het concurrerende waarden model worden de relaties tussen de vier organisatiemodellen weergegeven in twee dimensies, waardoor vier kwadranten ontstaan (zie figuur 2.3). Elk model past in een van de kwadranten. In figuur 2.3. loopt de verticale as van flexibiliteit bovenaan naar beheersing onderaan. De horizontale as loopt van interne gerichtheid aan de linkerkant naar externe gerichtheid rechts.

Ieder model kan als het tegenovergestelde worden gezien van een ander model, zoals beschreven door Quinn et al. (2003):

Het 'human relations' model – gedefinieerd door flexibiliteit en interne gerichtheid- staat tegenover het rationele doel model dat gedefinieerd is door beheersing en externe gerichtheid. In het eerste model worden mensen gewaardeerd omdat ze mensen zijn. In het tweede worden mensen alleen gewaardeerd als zij in hoge mate bijdragen aan het doel. Het open systeem model – gedefinieerd door flexibiliteit en externe gerichtheid – is

tegengesteld aan het interne proces model dat is gedefinieerd door beheersing en interne gerichtheid. Terwijl het open systeem model te maken heeft met aanpassing aan de voortdurende verandering van de omgeving, betreft het interne proces model het handhaven van stabiliteit en continuïteit binnen het systeem. (p. 21)

Er zijn ook overeenkomsten tussen de modellen. Het rationale doel model en het open systeem model zijn beide extern (op de buitenwereld) gericht. Bij het interne proces model en het ‘human relations’ model ligt de nadruk op de interne organisatie.



Figuur 2.3 Het concurrerende waarden model: vier organisatiemodellen met effectiviteitscriteria. Op de assen staan de kenmerken die de modellen van elkaar onderscheiden. Overgenomen uit: Quinn et al., 2003, p. 19.

Er zijn ook overeenkomsten tussen de modellen. Het rationale doel model en het open systeem model zijn beide extern (op de buitenwereld) gericht. Bij het interne proces model en het ‘human relations’ model ligt de nadruk op de interne organisatie. Het ‘human relations’ en het open systeem model benadrukken flexibiliteit en ontwikkeling, terwijl de andere twee modellen meer gericht zijn op beheersing en efficiëntie. Quinn en Rohrbaugh merken op dat het rationale doel model en het interne proces model conceptueel het dichtst bij elkaar liggen. Dit blijkt uit de uitkomsten van de multidimensionale schaalanalyse, waarbij de effectiviteitscriteria werden afgebeeld in een ruimte die is opgespannen door de gevonden dimensies (Quinn & Rohrbaugh, 1983). De effectiviteitscriteria voor het interne proces model en het rationeel model model zijn minder gedifferentieerd op de intern-extern dimensie (zie figuur 2.3) dan die voor het ‘human relations’ en open systeem model. Quinn en Rohrbaugh wijzen erop dat het interne proces model in de literatuur niet altijd wordt onderscheiden van het rationale doel model (voor referenties, zie Quinn & Rohrbaugh, 1983).

In de complexe en snel veranderende wereld van nu is elk van de vier organisatiemodellen afzonderlijk te beperkt om organisatie-effectiviteit te bereiken. Soms is stabiliteit nodig in een organisatie, soms verandering. Afhankelijk van de situatie zullen bepaalde benaderingen de nadruk krijgen, terwijl hierbij ‘tegengestelde’ of concurrerende waarden tegelijkertijd van belang kunnen zijn. Toen Quinn en zijn collega Cameron de ontwikkeling van jonge organisaties in kaart brachten met behulp van het concurrerende waarden model, bleek dat er vier overgangsfasen te onderscheiden zijn, waarin verschillende waarden en effectiviteitscriteria de nadruk krijgen (Quinn, 1998; Quinn & Cameron, 1983). De ondernemersfase wordt gekenmerkt door innovatie, creativiteit en het verwerven van productiemiddelen. Hier ligt het accent op open systeem criteria voor effectiviteit. Daarna komt de collectiviteitsfase, die zich onderscheidt door informele communicatie en structuur, een familiegevoel, samenwerking, toewijding en persoonlijk leiderschap. In deze fase ligt er veel meer nadruk op de criteria van het ‘human relations’ model dan in de andere fasen, maar open systeem criteria zijn ook nog steeds belangrijk. De derde overgangsfase is de formalisatiefase. In deze fase kenmerken organisaties zich door stabiliteit wat betreft de organisatie, regels en procedures en conservatieve trends. De effectiviteit lijkt vooral te worden afgemeten aan criteria van het interne proces model en het rationele doel model. Tenslotte is er de uitwerkingsfase, waarin de structuur wordt uitgewerkt. In deze fase houdt de organisatie vooral de externe omgeving in de gaten, om zich te vernieuwen en te groeien. In deze fase ligt de meeste nadruk op het open systeem model, terwijl er een bescheiden accent ligt op de criteria uit de andere modellen.

Voor het analyseren van de organisatiecultuur wordt door Quinn (1998) uitgegaan van dezelfde dimensies en vrijwel dezelfde criteria als voor het vaststellen van de organisatie-effectiviteit (zie figuur 2.3). Maslowski (2001) merkt op dat het onderscheid tussen beide concepten hierdoor onduidelijk wordt. Een ander punt van kritiek op het concurrerende waarden model is naar voren gebracht door Van Muijen (1994). Hij wijst erop dat bij de operationalisering van het framework door Quinn c.s. de vier waardenoriëntaties: rationele doel, interne proces, ‘human relations’ en open systeem, meestal worden geïnterpreteerd als vier onafhankelijke factoren, die verschillende aspecten van het functioneren van een organisatie vertegenwoordigen (zie Maslowski, 2001, p. 29). De dimensies intern-extern en flexibiliteit-controle (zie figuur 2.3) worden hierbij vooral beschouwd als referentiekader en niet zozeer als coördinatensysteem (Van Muijen, 1994; Boerman, 1998).

Het concurrerende waarden model is door Quinn c.s. verder uitgewerkt als managementmodel om effectief leiderschap te beschrijven (Quinn, 1998; Quinn et al., 2003). Het raamwerk specificeert de verschillende rollen die van een manager verwacht worden bij de verschillende organisatiemodellen. In tabel 2.2 staat een overzicht van de vier managementmodellen in het concurrerende waarden model, met de bijbehorende effectiviteitscriteria, leiderschapsrollen en vaardigheden (zie Quinn et al., 2003). Bij het rationele doel model horen de rollen van bestuurder en producent, bij het interne procesmodel de rollen van controleur en coördinator, bij het ‘human relations’ model de rollen van stimulator en mentor en bij het open systeem model de rollen van innovator en bemiddelaar. Bij elke rol is een specifieke set vaardigheden geformuleerd. Deze vaardigheden geven inzicht in de aard van de verschillende rollen.

De vier managementmodellen in het concurrerende waarden model (zie tabel 2.2) vertonen grote overeenkomsten met de vier categorieën van schoolleidersactiviteiten die door Leithwood worden beschreven (zie paragraaf 3): het rationele doel model overlapt voor een belangrijk deel met de categorie ‘visie ontwikkelen en richting geven’; het interne proces model vertoont overeenkomsten met de categorie ‘managen van het onderwijs- en

leerprogramma'; het 'human relations' model lijkt erg op de categorie 'mensen begrijpen en ontwikkelen' en het open systeem model vertoont dezelfde externe gerichtheid als de categorie 'herontwerpen van de organisatie'. Het ligt natuurlijk voor de hand dat deze twee generieke leiderschapsmodellen een grote gelijkenis vertonen, maar hiermee wordt ook de toepasbaarheid van het concurrerende waarden model in het schoolleidersonderzoek ondersteund.

Tabel 2.2 *Het concurrerende waarden model met managementmodellen, effectiviteitscriteria, leiderschapsrollen en leiderschapsvaardigheden (Quinn et al., 2003)*

Managementmodellen	Effectiviteitscriteria	Leiderschapsrollen	Vaardigheden
Rationele doel model	Productiviteit, resultaten	Producent	<ol style="list-style-type: none"> 1. Productief werken 2. Een productieve werkomgeving bevorderen 3. Tijd- en stressmanagement
	Richting, doelverduidelijking	Bestuurder	<ol style="list-style-type: none"> 1. Een visie ontwikkelen en communiceren 2. Doelen en doelstellingen communiceren 3. Ontwerpen en organiseren
Interne proces model	Stabiliteit, controle	Coördinator	<ol style="list-style-type: none"> 1. Projectmanagement 2. Taken ontwerpen 3. Crossfunctioneel management
	Documentatie, informatiemanagement	Controleur	<ol style="list-style-type: none"> 1. Informatie beheren door kritisch denken 2. Omgaan met een overvloed aan informatie 3. Kernprocessen beheren
Human relations model	Participatie, openheid	Stimulator	<ol style="list-style-type: none"> 1. Teambuilding 2. Participerende besluitvorming 3. Conflict managen
	Inzet, moreel	Mentor	<ol style="list-style-type: none"> 1. Inzicht in uzelf en anderen 2. Effectief communiceren 3. Ontwikkeling van werknemers
Open systeem model	Innovatie, aanpassing	Innovator	<ol style="list-style-type: none"> 1. Leven met verandering 2. Creatief denken 3. Management van veranderingen
	Groei, middelen verwerven	Bemiddelaar	<ol style="list-style-type: none"> 1. Een machtsbasis opbouwen en handhaven 2. Onderhandelen over inzet en overeenstemming 3. Ideeën presenteren

Uit een onderzoek van Quinn, Faermann en Dixit (1987, zoals beschreven in Quinn, 1998) is gebleken dat de typologie van leiderschapsrollen in het concurrerende waarden model bruikbaar is voor karakteriseren van managers op verschillende niveaus in een organisatie: zowel voor topmanagers, als voor het middenkader en eerste-lijnsmanagers. Wel zijn er verschillen in accent op een aantal rollen. De rollen van innovator en bemiddelaar nemen bijvoorbeeld in belang toe naarmate de positie van de manager op de hiërarchische ladder hoger wordt. De rollen van mentor en stimulator zijn volgens dit onderzoek wat minder belangrijk voor het middenkader. Het gewicht van de andere rollen blijft min of meer stabiel. Verder werden er door Quinn c.s. ook effectieve en ineffectieve ‘profielen’ van managers gevonden (Quinn, Faermann en Dixit, 1987). Effectieve managers hebben in het algemeen een evenwichtiger profiel dan ineffectieve managers. Effectieve managers besteden aan alle leiderschapsrollen de nodige aandacht, waarbij er wel een lichte nadruk kan zijn op een of meer rollen. Ineffectieve managers hebben vaak een profiel waarbij er een relatief veel nadruk is voor een of meer rollen, terwijl andere rollen onvoldoende worden vervuld.

Een voordeel van het concurrerende waarden model ten opzichte van andere leiderschapsmodellen is, dat het een verband legt tussen leiderschapsrollen en de organisatiecultuur. Quinn c.s. wijzen er echter ook op dat een effectieve manager alle rollen in voldoende mate moet vervullen. De mate waarin een manager bepaalde leiderschapsrollen vervult, kan afhangen van de organisatiecultuur, maar ook van de positie in de organisatie, of van de persoonlijke stijl of de competentie van de manager. Krüger en Witziers (2003) merken op dat de leiderschapsrollen niet per se allemaal door één persoon hoeven te worden vervuld: de rollen kunnen bijvoorbeeld ook verdeeld worden in een managementteam. Het concurrerende waarden model is nog maar weinig toegepast in schooleffectiviteits-onderzoek. Er zijn enkele onderzoeken naar schoolleiderschap bekend met dit model (zie bijv. Smart, 2003; Thompson, 2000). Hier wordt in hoofdstuk 5 nader op ingegaan. Verder is het model met name in Nederland gebruikt om de organisatiecultuur van scholen in kaart te brengen (Houtveen, Voogt, van der Vegt & van de Grift, 1995; Maslowski, 2001). Dit onderzoek komt in de volgende paragraaf aan de orde.

2.6 Schoolorganisatie, schoolklimaat en schoolcultuur

In het algemene beschrijvingsmodel van Bossert et al. (1982), dat als uitgangspunt wordt gebruikt voor dit onderzoek, wordt verondersteld dat de invloed van de schoolleider op de leerresultaten verloopt via het ‘pedagogisch onderwijskundig klimaat’ en de ‘onderwijsorganisatie’ (zie figuur 1.1). Het pedagogisch onderwijskundig klimaat omvat in dit model de volgende aspecten (overgenomen uit Krüger, 1994, p. 37): gebouw, grond, inrichting en uitrusting; het sociale leerplan; regels en discipline; interactiepatronen tussen leerlingen, personeel en lokale omgeving. De onderwijsorganisatie is opgebouwd uit: vakkenaanbod; differentiatiestructuur; didactiek; professionalisering van docenten. Beide concepten omvatten dus zeer uiteenlopende aspecten, op schoolniveau en op klassenniveau.

Onderzoekers blijken verschillende operationalisaties te hanteren van (school- of onderwijs-) organisatie, klimaat en cultuur. Dit kan worden geïllustreerd aan de hand van een aantal voorbeelden van onderzoek naar indirecte effectmodellen van schoolleiderschap (zie ook tabel 2.1). Heck, Larsen en Marcoulides (1990) onderzochten een causaal model waarin ‘schoolklimaat’ en ‘onderwijsorganisatie’ zijn opgenomen als intermediaire variabelen voor de invloed van schoolleiders op de leerprestaties. Hun uitkomsten bevestigen de relaties die worden verondersteld in het model van Bossert et al. (1982). Het schoolklimaat zoals geoperationaliseerd door Heck et al. wordt getypeerd door de mate waarin de schoolleider

hoge verwachtingen van de prestaties van leerlingen bevordert en de onderwijsorganisatie door superviserende activiteiten van de schoolleider. Heck et al. vonden een groot effect van onderwijsorganisatie op schoolklimaat. De Maeyer en Rymenans (2004) toonden aan dat er een verband bestaat tussen ‘geïntegreerd leiderschap’, een ‘prestatiegericht klimaat’ en de leesprestaties van leerlingen. In dit onderzoek wordt een prestatiegericht klimaat gekenmerkt door hoge verwachtingen, prestatiegerichte doelstellingen, monitoring en regelgeving. Krüger, Witziers en Slegers (2007) vonden een effect van ‘onderwijskundig leiderschap’ op ‘schoolorganisatie’ en van schoolorganisatie op ‘schoolcultuur’, maar geen verband tussen deze variabelen en de schoolbeleving van leerlingen. In dit laatste onderzoek wordt de schoolcultuur gekarakteriseerd aan de hand van de mening van docenten over de algemene sfeer en de loyaliteit tussen collega’s en de schoolorganisatie als de algemene gang van zaken met betrekking tot de informatievoorziening, gezamenlijke besluitvorming en verantwoordelijkheid.

Wat wordt nu precies verstaan onder ‘schoolcultuur’, ‘schoolklimaat’ en ‘schoolorganisatie’? Houtveen, Voogt, van der Vegt en van de Grift (1995, p. 34) constateren dat operationalisering van organisatiecultuur en -klimaat in de loop van de tijd zijn veranderd en elkaar soms overlappen. Een gangbare opvatting is dat schoolcultuur betrekking heeft op de gemeenschappelijke vooronderstellingen, waarden en normen op school en schoolklimaat op de (gemeenschappelijke) perceptie van de werkomstandigheden (zie bijv. Hallinger & Heck, 1998; Hoy et al. 1991; De Maeyer & Rymenans, 2004; Maslowski, 1997). Visscher (1999, p. 19) laat zien dat de perceptie van de werkomstandigheden door docenten (het schoolklimaat) wordt bepaald door zowel de feitelijke werkomstandigheden als de waarden en normen van docenten. Daarbij bestaat er een wisselwerking tussen de waarden en normen van docenten en de schoolcultuur (gebaseerd op Maslowski, 1997).

Houtveen et al. (1995) hanteren een enigszins afwijkende opvatting van organisatieklimaat. Zij zien organisatieklimaat niet als een ‘gemeenschappelijke perceptie’ van een situatie, maar als “een serie randvoorwaarden die van invloed zijn op het gedrag van individuen” (p. 33). Hierbij geven zij de volgende nadere omschrijving:

Deze randvoorwaarden kunnen zowel fysiek zijn als perceptueel (bijvoorbeeld: coördinatie, sociale afstand, status verschillen, participatie in besluitvorming, mandatering, etc.). Het begrip klimaat wordt in dit geval ingevuld met de reeks potentiële processen, praktijken en (organisatorische) voorwaarden die van invloed zijn op de manier waarop de organisatie functioneert.” (Houtveen et al., 1995, p. 33).

Houtveen et al. omschrijven organisatieklimaat als een “manifestatie van de cultuur van de organisatie” (p. 33) en als complementair aan cultuur. Zij geven daarbij aan dat voor deze invulling van het begrip klimaat ook wel de term ‘organisatiestructuur’ wordt gebruikt (zie Payne & Pugh, 1976).

Een bekende definitie van organisatiestructuur is afkomstig van Mintzberg (1979): “The structure of an organization can be defined simply as the sum total of the ways in which it divides its labor into distinct tasks and then achieves coordination among them.” (p. 2). Witziers (1992, 1999) geeft aan dat er verschillende perspectieven zijn op de organisatiestructuur van scholen. Aan de ene kant worden scholen gezien als bureaucratische organisaties (Mintzberg, 1979) met een grote mate van coördinatie en controle, vooral door regels en procedures vanuit hogere bestuurlijke niveaus. Een ander perspectief geeft aan dat scholen ‘losjes gekoppelde’ organisaties zijn (Weick, 1976, 1982, 1984). Hierbij wordt ervan

uitgegaan dat docenten autonoom functioneren en dat zij weinig verantwoording hoeven af te leggen, omdat formele regels en evaluatieprocedures ontbreken. Volgens een derde perspectief zijn scholen op te vatten als professionele bureaucratieën (Mintzberg, 1979), waarbij coördinatie vooral wordt gerealiseerd door de opleiding en socialisatie van docenten. Witziers (1999, p. 133) concludeert dat niet zonder meer is vast te stellen welk perspectief juist is. Er is empirische onderbouwing voor elk van deze perspectieven. Verder kunnen er binnen één school zowel strakke als losse coördinatiemechanismen gehanteerd worden en tenslotte zijn er ook nog grote verschillen tussen schooltypes en landen.

De definiëring van het begrip organisatieklimaat als “structurele randvoorwaarden waaronder de organisatie functioneert” (Houtveen et al., 1995, p. 35) overlapt gedeeltelijk de definitie van organisatiestructuur van Mintzberg (zie hierboven). Organisatiestructuur en organisatiecultuur kunnen beide worden gezien als componenten van een organisatie, naast ‘strategie’, ‘organisatieleden’, ‘kwalificaties’ en ‘werktechnologie’ (zie Scheerens, 1993, p. 23; Watson, 1986). Volgens deze invalshoek wordt de organisatiecultuur beïnvloed door zowel de strategie, als de structuur en de organisatieleden (zie Houtveen et al., 1995, p. 12).

In de optiek van Schein (1985) is de ‘organisatiecultuur’ het belangrijkste domein voor de invloed van schoolleiders. Schein meent dat het gedrag van docenten minder wordt beïnvloed door regels, procedures, structuren of training, dan door normen en waarden. Hij stelt dat: “There is a real possibility that the only thing of real importance that leaders do is create and manage culture” (in Visscher, 1999, p. 19). Voor het verband tussen schoolleiderschap, schoolcultuur en leerprestaties bestaat echter nog maar weinig empirische onderbouwing (Hallinger & Heck, 1998). Dit zal niet alleen het gevolg zijn van verschillen in de definitie van het begrip cultuur, maar ook van het feit dat de organisatiecultuur kan veranderen, afhankelijk van de ontwikkelingsfase waarin de organisatie verkeert (zie bijv. Marx, 1986; Schein, 1985, 1992; Quinn, 1998; Quinn & Cameron, 1983).

Houtveen et al. maakten gebruik van concurrerende waarden model (zie paragraaf 5) voor de ontwikkeling van een instrument om de organisatiecultuur in scholen te meten. Dit instrument is gericht op “gedragingen en belevingen van praktijken” en op “de waarden en normen die men hanteert ten opzichte van het onderwijs” (Houtveen et al., 1995, p. 36). Het instrument van Houtveen et al. werd ontwikkeld voor het basisonderwijs. Maslowski (2001) construeerde op basis van het concurrerende waarden model een instrument voor het meten van de schoolcultuur in het voortgezet onderwijs. De vragenlijst van Maslowski is erop gericht om meer direct de onderliggende waarden op organisatieniveau te bevragen.

Uitgaande van uitkomsten van eerder onderzoek naar de relatie tussen schoolcultuur en effectiviteit (bijv. Heck & Marcoulides, 1996; Gaziel, 1997; Pang, 1998) en ander schooleffectiviteits- en schoolverbeteringsonderzoek veronderstelde Maslowski (2001) dat elk van de vier cultuuriëntaties uit het concurrerende waarden model (rationele doel, interne proces, ‘human relations’ en open systeem) positief zou samenhangen met de leerprestaties. Uit zijn onderzoek bleek echter dat er geen relatie aantoonbaar was tussen de afzonderlijke cultuuriëntaties van scholen en de leerprestaties. Er werd wel een negatief effect op de leerprestaties gevonden van een overwegend flexibele cultuuriëntatie (‘human relations’ en open systeem), vergeleken met een interne oriëntatie (interne proces en ‘human relations’), maar dit gold alleen voor vwo-scholen. Verder bleek dat zowel de interne proces oriëntatie als de rationele doel oriëntatie een sterk effect hadden op het gemiddelde doorstroompercentage in de bovenbouw.

2.7 Het schooleffectiviteitsmodel

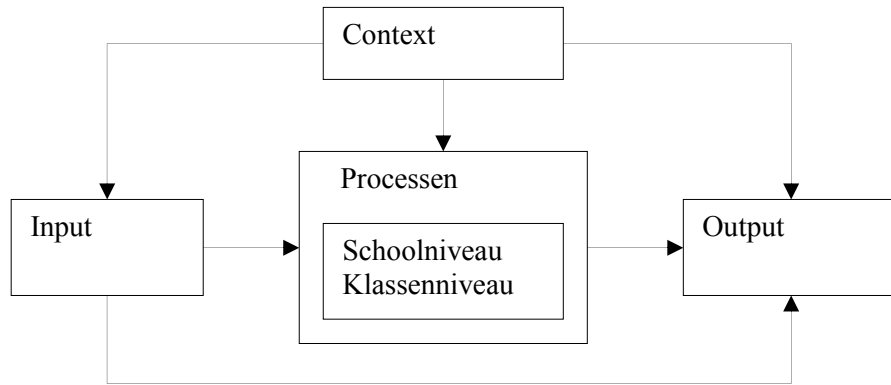
Welke aspecten van de schoolorganisatie(-cultuur) dragen nu het meest bij aan de effectiviteit van de school? Een veelgeciteerd overzicht van effectiviteitsbevorderende kenmerken is opgesteld door Scheerens en Bosker (1997). Op basis van een overzichtsstudie van schooleffectiviteitsstudies en zelfevaluatie-instrumenten komen zij tot de volgende factoren:

1. Prestatiegerichtheid / hoge verwachtingen
2. Onderwijskundig leiderschap
3. Consensus en cohesie onder de teamleden
4. Curriculumkwaliteit / gelegenheid om te leren
5. Schoolklimaat: ordelijke atmosfeer; oriëntatie op effectiviteit en goede interne relaties
6. Evaluatief vermogen / monitoring
7. Betrokkenheid van ouders
8. Klimaat in de klas
9. Effectieve leertijd (klassenmanagement)
10. Gestructureerde instructie
11. Onafhankelijk leren²
12. Differentiatie, adaptief onderwijs
13. Feedback en bekrachtiging

De eerste zeven factoren zijn op schoolniveau, de rest op klassenniveau. In het algemeen geldt dat factoren die nauw gekoppeld zijn aan het instructieproces een grotere invloed hebben op de leerprestaties dan factoren op schoolniveau (Scheerens & Bosker, 1997, p. 304). In de schooleffectiviteitsbevorderende factoren zijn aspecten van het concurrerende waarde model te herkennen, zo past 'prestatiegerichtheid' bijvoorbeeld in het rationele doel model, een 'ordelijke atmosfeer' in het interne proces model, 'consensus en cohesie onder teamleden' in het 'human relations' model en 'betrokkenheid van ouder' in het open systeem model.

In het schooleffectiviteitsmodel (Scheerens, 1990; zie ook Scheerens & Bosker, 1997) worden vier soorten variabelen onderscheiden: Context, Input, Process en Outputvariabelen. Figuur 2.4 geeft een schematische weergave van het CIP0-model. De dertien effectiviteitsbevorderende factoren uit de bovenstaande lijst zijn procesvariabelen die een aantoonbaar effect hebben op de uitkomstvariabelen. De uitkomsten worden echter ook beïnvloed door inputvariabelen en contextvariabelen. Inputvariabelen kunnen worden onderscheiden in variabelen op schoolniveau, klassenniveau en leerlingniveau (Bosker, de Vos & Witziers, 2000). Op schoolniveau gaat het bijvoorbeeld om de materiële en financiële middelen van de school en om de leerling/docent ratio. Op het niveau van de klas telt de ervaring en de betrokkenheid van de docenten mee. Inputvariabelen op leerlingniveau hebben echter de verreweg de grootste invloed op de uitkomsten, hiermee kan mogelijk tot 85% van de variantie in de leerprestaties verklaard worden (Luyten, 1994). Belangrijke achtergrondvariabelen op leerlingniveau zijn: intelligentie (IQ), eerdere prestaties en de sociaaleconomische status (SES) van het thuismilieu (zie bijv. De Maeyer & Rymenans, 2004).

² Gebaseerd op het constructivisme (Scheerens & Bosker, 1997, p.131)



Figuur 2.4 Schema van CIPPO model voor schooleffectiviteit (Bronnen: Bosker, de Vos & Witziers, 2000, p. 4; Luyten et al., 2005; Scheerens & Witziers, 2005, p. 9)

In het CIPPO-model worden door Scheerens (1990) de volgende contextvariabelen onderscheiden:

- Stimulans tot presteren vanuit hogere bestuurlijke organen
- Ontwikkeling van consumentisme in het onderwijs
- Covariabelen, zoals samenstelling van de leerlingenpopulatie, schoolgrootte, schoolsoort, stedelijkheid van de omgeving.

In een overzicht van het onderzoek naar contexteffecten binnen het schooleffectiviteitsonderzoek noemen Teddlie, Stringfield en Reynolds (2000) bijvoorbeeld ook nog de fase van de opleiding en de bestuursvorm van de school (gekoppeld aan denominatie) als belangrijke categorieën van contextvariabelen.

De effectiviteit van scholen wordt in het algemeen gemeten aan de hand van het 'productiviteitscriterium' (Scheerens & Bosker, 1997, p. 12). In schooleffectiviteitsonderzoek wordt dus meestal uitgegaan van leerprestaties als uitkomstvariabele. Het belang van additionele uitkomstvariabelen wordt echter onderkend (Reynolds & Teddlie, 2000). Aan de ene kant zouden verschillende cognitieve outputmaten gebruikt moeten worden, omdat schooleffectiviteit niet altijd consistent is over verschillende vakgebieden (Teddlie, Reynolds & Sammons, 2000). Aan de andere kant zouden ook niet-cognitieve uitkomstmaten in beschouwing moeten worden genomen. Niet-cognitieve uitkomsten, zoals de attitudes van leerlingen, worden ook wel gezien als intermediaire variabelen, maar het verband met leerprestaties blijkt meestal zwak (Fitz-Gibon & Kochan, 2000, p. 264). Alternatieve uitkomstmaten kunnen echter ook op zichzelf waardevol zijn. Het belang dat gehecht wordt aan verschillende uitkomstmaten hangt af van de effectiviteitscriteria die worden gehanteerd. De doelen van onderwijs gaan verder dan alleen cognitieve resultaten. Door verschillende cognitieve en niet-cognitieve uitkomstvariabelen te gebruiken kan meer inzicht worden verkregen in de effectiviteit van scholen (zie ook Reynolds & Teddlie, 2000).

Er is slechts een beperkt aantal schooleffectiviteitsonderzoeken te vinden, waarin zowel cognitieve als niet-cognitieve uitkomstmaten zijn meegenomen. Opdenakker en Van Damme (2000) geven hiervan een overzicht (Bosker, 1990; Brookover, Beady, Flood, Schweitzer & Wisenbaker, 1979; Grisay, 1996, Knuver & Brandsma, 1993; Mortimore, Sammons, Stoll, Lewis & Ecob, 1988; Rutter, Maughan, Morimore & Ouston, 1979). De resultaten met betrekking tot de relaties tussen cognitieve en niet-cognitieve uitkomsten op schoolniveau van deze onderzoeken lopen uiteen. Er zijn aanwijzingen dat scholen die de leerprestaties

verbeteren niet noodzakelijk even effectief zijn in het bevorderen van het welbevinden van leerlingen (Opdenakker & Van Damme, 2000). De school- en klasseneffecten voor niet-cognitieve uitkomsten zijn veel kleiner dan die voor cognitieve uitkomsten (Van Landeghem, Van Damme, Opdenakker, De Fraine & Onghena, 2002). Niet-cognitieve uitkomsten zijn moeilijker te meten en de resultaten van onderzoek zijn vaak moeilijk te vergelijken, omdat er gebruik wordt gemaakt van verschillende achtergrondvariabelen bij het bepalen van netto schooleffecten. Knuver en Brandsma (1993) onderzochten de samenhang tussen affectief en cognitief functioneren van leerlingen en de rol van klas- en schoolkenmerken hierbij, in het Nederlandse basisonderwijs. Hieruit bleek dat motivatie en attitudes geen noodzakelijke voorwaarden zijn voor het leveren van prestaties (zie Knuver, 1993, p. 144). Er lijkt een wisselwerking te zijn tussen affectief en cognitief functioneren van leerlingen, waarbij goede leerprestaties bevorderend zijn voor het ontwikkelen van positieve attitudes, prestatiemotivatie, zelfbeeld en schoolbeleving. Er is geen duidelijke set van klas- en schoolkenmerken geïdentificeerd waarmee de verschillen in het affectief functioneren van leerlingen verklaard kan worden (Knuver, 1993, p. 145). Knuver geeft aan dat verbanden tussen affectief en cognitief functioneren sterker en stabiel lijken te zijn naarmate de leerling meer tijd heeft doorgebracht in het onderwijs. In het voortgezet onderwijs zou het affectief functioneren van leerlingen daarom wel een rol kunnen spelen bij de leerprestaties (Knuver, 1993, p. 146). Opdenakker en van Damme (2000) onderzochten het effect van scholen en klassen op de leerprestaties en het welbevinden van leerlingen in het Vlaamse voortgezet onderwijs aan de hand van data die werden verzameld in het LOSO-project (Longitudinaal Onderzoek Secundair Onderwijs, zie ook Van Damme, De Fraine, Van Landeghem, Opdenakker & Onghena, 2002). Zij concluderen uit hun onderzoek dat leerprestaties en welbevinden als twee verschillende uitkomstvariabelen beschouwd moeten worden en dat schoolkarakteristieken uiteenlopende effecten hebben op deze uitkomsten.

2.8 Onderzoeksmodel

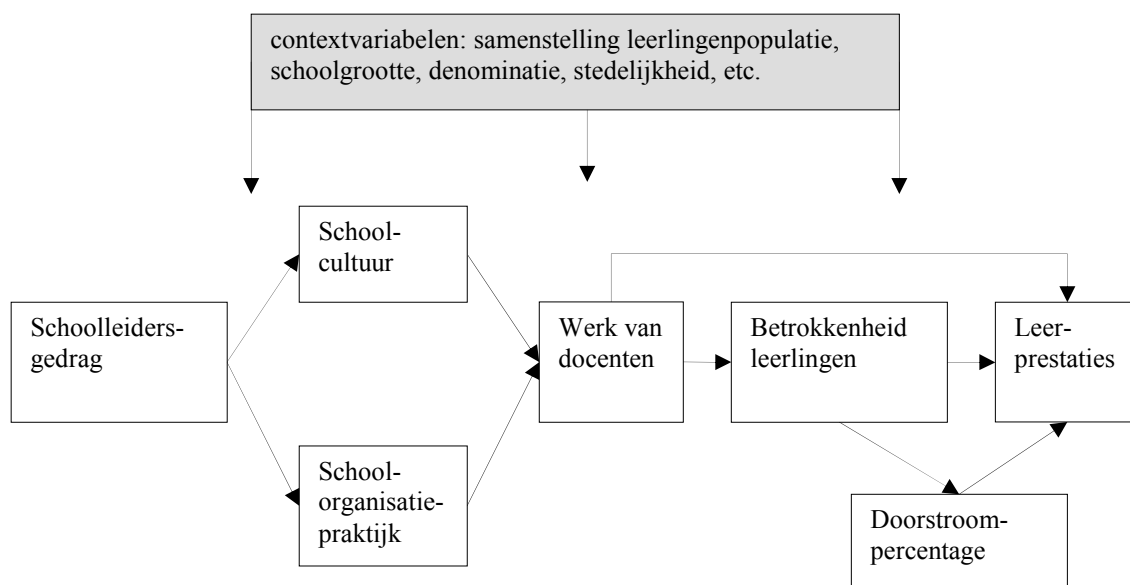
In dit onderzoek naar de effecten van schoolleiders op de leerprestaties is het algemene beschrijvingsmodel van onderwijskundig leiderschap van Bossert et al. (1982) als uitgangspunt genomen (zie figuur 1.1). Om het model van Bossert et al. nader te kunnen specificeren, werd een literatuuronderzoek uitgevoerd aan de hand van de volgende onderzoeksvraag:

Welke contextvariabelen, schoolleidersvariabelen, intermediaire variabelen en uitkomstvariabelen zijn relevant voor het in kaart brengen van de effectiviteit van de schoolleider en welke relaties zijn er tussen deze variabelen te onderscheiden?

De uitkomsten van het literatuuronderzoek zijn beschreven in dit hoofdstuk. Uit de resultaten blijkt dat er verschillende opvattingen zijn over schoolleiderschap. Verder zijn uiteenlopende intermediaire factoren geïdentificeerd voor de invloed van de schoolleider op de leerprestaties. Om de beperkingen van een specifiek leiderschapsmodel te omzeilen, is er in dit onderzoek voor gekozen om een generiek model toe te passen, met een breed spectrum aan leiderschapsactiviteiten. Het concurrerende waarden model biedt een dergelijk integraal leiderschapsmodel, dat bovendien goed past bij nieuwere opvattingen van schoolleiderschap (zie paragrafen 2 en 3). Intermediaire factoren die zijn gevonden voor het werk van de schoolleider, omvatten op schoolniveau onder meer aspecten van de schoolcultuur en de schoolorganisatie (zie paragrafen 4 en 6). Het concurrerende waarden model biedt de mogelijkheid om verbanden te leggen tussen leiderschapsgedrag, organisatiecultuur en organisatie-effectiviteit (zie paragraaf 5). Voor de operationalisatie van intermediaire factoren

is het van belang om rekening te houden met schooleffectiviteitsbevorderende factoren zoals die zijn onderscheiden door Scheerens en Bosker (1997, zie paragraaf 7).

Het algemene model van onderwijskundig leiderschap (figuur 1.1) werd aan de hand van de bevindingen in dit hoofdstuk en mede geïnspireerd door het LOLSO-model van Mulford en Silins (figuur 2.2), nader uitgewerkt in een onderzoeksmodel dat is afgebeeld in figuur 2.5. In het model zijn behalve de schoolcultuur en de schoolorganisatiepraktijk ook het werk van docenten en de betrokkenheid van leerlingen opgenomen als intermediaire variabelen. Verondersteld wordt dat de schoolleider via de schoolcultuur en de schoolorganisatie invloed uitoefent op het werk van docenten. Het werk van docenten kan vervolgens weer een effect hebben op de leerprestaties, direct, of indirect via een grotere betrokkenheid van leerlingen en een hoger gemiddeld doorstroompercentage. De variabele ‘doorstroompercentage’ is opgenomen naar aanleiding van de resultaten van Maslowski (2003, zie paragraaf 6). Hij vond geen effect van schoolcultuur op de leerprestaties, maar wel op het gemiddelde doorstroompercentage van de derde klas naar het diploma. Het doorstroompercentage is te vergelijken met de variabele ‘retention’ in het LOLSO-model (zie paragraaf 4). Voor het onderzoeksmodel dat is afgebeeld in figuur 2.5 zijn alleen die contextvariabelen van belang, die zowel invloed uitoefenen op de uitkomsten, als op andere variabelen in het model (zie paragraaf 7). De mate waarin contextvariabelen het gedrag van de schoolleider verklaren, is het onderwerp van een complementair promotie-onderzoek (Schmidt, 2009).



Figuur 2.5 Onderzoeksmodel voor het analyseren van de invloed van schoolleidersgedrag op uitkomstvariabelen

In het volgende hoofdstuk, over de opzet van het onderzoek en de dataverzameling, wordt uiteengezet hoe de variabelen in het onderzoeksmodel zijn geoperationaliseerd. De validering van de onderzoeksinstrumenten wordt besproken in hoofdstuk 4. In hoofdstuk 5 worden de kenmerken van scholen beschreven aan de hand van de variabelen in het onderzoeksmodel en de samenhang tussen deze kenmerken en de uitkomstvariabelen. De contextvariabelen komen nader aan de orde in hoofdstuk 6, waarin de toetsing van het gehele onderzoeksmodel wordt besproken. Hoofdstuk 7 bevat een samenvatting van de uitkomsten en de conclusies.

Hoofdstuk 3

Opzet van het onderzoek en dataverzameling

3.1 Inleiding

In hoofdstuk 1 is beschreven wat de aanleiding is voor het onderzoek naar de effecten van sturing door schoolleiders en welke conceptuele en methodologische problemen hierbij een rol spelen. In hoofdstuk 2 zijn uiteenlopende benaderingen van schoolleiderschap aan de orde gekomen en een aantal indirecte effectmodellen die zijn onderzocht. Uitgaande van het algemene beschrijvingsmodel van onderwijskundig leiderschap van Bossert et al. (1982) en modellen van Leithwood en Jantzi (Leithwood & Jantzi, 2000) en van Mulford en Silins (Mulford & Silins, 2003), is een onderzoeksmodel opgesteld (zie figuur 2.5). Om met behulp van het onderzoeksmodel de invloed van schoolleiderschap op de leerresultaten te kunnen onderzoeken zijn de volgende stappen uitgevoerd:

1. Constructie van onderzoeksinstrumenten
2. Dataverzameling in het voortgezet onderwijs
3. Validering van de instrumenten
4. Toetsing van het onderzoeksmodel aan de data

Dit hoofdstuk gaat over de eerste twee stappen. Deze stappen zijn door de onderzoeksgroepen van de Universiteit van Amsterdam en de Universiteit Twente in samenwerking uitgevoerd. In paragraaf 3.2 wordt beschreven hoe de onderzoeksinstrumenten zijn samengesteld (stap 1) en in paragraaf 3.3 hoe de dataverzameling is opgezet en uitgevoerd (stap 2). Daarna worden in paragraaf 3.4 beschrijvende gegevens gepresenteerd van de deelnemende scholen en de respondentengroepen. In paragraaf 3.5 wordt een overzicht gegeven van de analysemethoden voor de volgende onderzoeksstappen. Het hoofdstuk wordt afgesloten met een discussie in paragraaf 3.6.

3.2 Constructie van onderzoeksinstrumenten

Bij het operationaliseren van de variabelen in het onderzoeksmodel (zie figuur 2.5) is zoveel mogelijk uitgegaan van bestaande instrumenten. Het concurrerende waarden model (Quinn & Rohrbaugh, 1983) diende als uitgangspunt voor het meten van zowel de handelingen van schoolleiders, als de schoolcultuur en de schoolorganisatie. Hiertoe werd telkens voor elk kwadrant in het model een schaal samengesteld. Items voor het bepalen van de handelingen van de schoolleider zijn vooral ontleend aan een vragenlijst voor het meten van de Onderwijskundige Gerichtheid van Schoolleiders van Krüger (1994). Voor het in kaart brengen van de schoolorganisatie en de schoolcultuur werd het werk van Houtveen, Voogt, van der Vegt en van de Grift (1995) en van Maslowski (2001) als uitgangspunt gebruikt. Aan de hand van de Student Engagement and Family Culture Survey van Leithwood et al. (Leithwood & Jantzi, 2000; Leithwood, Aitken & Jantzi, 2000) werd de ‘Schoolbeleving’ van

leerlingen geoperationaliseerd. De uitkomstvariabelen ‘leerprestaties’ en ‘doorstroompercentage’ zijn bepaald met behulp van de databestanden die door de Inspectie van het Onderwijs worden gebruikt om de opbrengstenkaarten voor het voortgezet onderwijs samen te stellen. Deze gegevens worden beschikbaar gesteld door DANS (Data Archiving and Networked Services). DANS is een instituut van de KNAW (Koninklijke Nederlandse Akademie van Wetenschappen) en NWO (Nederlandse Organisatie voor Wetenschappelijk Onderzoek) dat onder meer sociaal wetenschappelijke databestanden beheert. Voor de leerprestaties is gebruik gemaakt van de cijfers voor het centraal schriftelijk eindexamen. Het doorstroompercentage is gebaseerd op het percentage leerlingen dat onvertraagd doorstroomt van de derde klas naar het diploma. De uitkomstvariabelen worden nader beschreven in hoofdstuk 5. In deze paragraaf wordt uiteengezet hoe de onderzoeksinstrumenten voor het meten van de handelingen van schoolleiders, de schoolcultuur, de schoolorganisatiepraktijk en de schoolbeleving van leerlingen zijn geconstrueerd.

3.2.1 Handelingen van schoolleiders

Het instrument *Onderwijskundige Gerichtheid van Schoolleiders* (OGS), dat is ontwikkeld door Krüger (1994), diende als belangrijkste bron van items voor het in kaart brengen van de handelingen van schoolleiders. Dit instrument is gebaseerd op de *Principal Instructional Management Rating Scale* (PIMRS) van Hallinger (zie Hallinger & Murphy, 1985). De schalen van Hallinger werden aangevuld met schalen en items uit de vragenlijst *Onderwijskundig Leiderschap* van Van de Grift (Van de Grift 1985a en 1985b). De OGS bestaat uit 50 items, verdeeld over zes schalen (zie Krüger, 1994, p. 109):

1. Missiegericht onderwijskundig leiderschap
2. Cultuurgericht onderwijskundig leiderschap
3. Klassegericht onderwijskundig leiderschap
4. Onderwijskundig leiderschap via het belonen van docenten
5. Onderwijskundig leiderschap via het bevorderen van nascholing
6. Onderwijskundig leiderschap via het werken aan een ordelijke en taakgerichte sfeer

Voor het meten van de handelingen van schoolleiders werden de items van de OGS opnieuw ingedeeld volgens de vier kwadranten van het concurrerende waarden model (zie tabel 2.2). Items uit de schaal ‘Missiegericht onderwijskundig leiderschap’ konden worden gekoppeld aan het rationele doel model. Bij het interne proces model pasten items uit de schaal ‘Onderwijskundig leiderschap via het werken aan een ordelijke en taakgerichte sfeer’. Items uit de schalen ‘Cultuurgericht onderwijskundig leiderschap’, ‘Onderwijskundig leiderschap via het belonen van docenten’ en ‘Onderwijskundig leiderschap via het bevorderen van nascholing’ konden worden ondergebracht bij het ‘human relations’ model. De items over klassegericht onderwijskundig leiderschap waren niet goed bij één model van het concurrerende waarden model onder te brengen. Na deze verdeling bleek dat de OGS vooral veel items opleverde voor het ‘human relations’ model, terwijl er ook voldoende items gevonden werden voor het rationele doel model. Voor het interne proces model was slechts een beperkt aantal items beschikbaar terwijl er uiteindelijk één item werd gevonden voor het Open systeem model. Het vrijwel ontbreken van items voor het open systeem model geeft aan dat de OGS nog is gebaseerd op een relatief nauwe opvatting van onderwijskundig leiderschap (zie hoofdstuk 1).

De nieuwe schalen van OGS-items werden op betrouwbaarheid getoetst met behulp van de oorspronkelijke dataset uit het onderzoek naar schoolleiderschap van Krüger (1994). De schalen bleken voldoende betrouwbaar. Om gelijkwaardige meetschalen te creëren voor de

vier organisatiemodellen in het concurrerende waarden model, werden vervolgens de meest geschikte OGS-items geselecteerd en werden nieuwe items ontwikkeld om de schalen aan te vullen waar dat nodig was (met name voor de open systeem schaal en de interne processchaal). De procedure die werd gevolgd was erop gericht om een zo groot mogelijke inhoudsvaliditeit te verkrijgen. Inhoudsvaliditeit wil zeggen dat het instrument alle relevante aspecten van een concept meet (Swanborn, 1994, p. 189). Bij de constructie van de vragenlijsten werd gebruik gemaakt van de uitkomsten van interviews met vijftien schoolleiders. Deze interviews waren erop gericht om actuele thema's voor de beroepsgroep van schoolleiders bloot te leggen en om de onderwerpen voor de schoolleidersvragenlijst te bespreken.

Uiteindelijk is een vragenlijst gerealiseerd met 42 uitspraken over handelingen en gedragingen van schoolleiders, verdeeld over vier schalen. Voorbeelden van items zijn:

- Rationele doel model: “Werken aan het ontwikkelen van doelen die door docenten gemakkelijk vertaald kunnen worden naar operationele doelen” en “Proberen zoveel mogelijk duidelijkheid te scheppen over wat wij met ons onderwijs willen”.
- Interne proces model: “Werken aan het tot stand brengen van een ordelijke sfeer op school” en “Erop toezien dat de activiteiten en procedures vastgelegd worden in verslagen”.
- ‘Human relations’ model: “Op individueel hulp en begeleiding aan docenten geven” en “In docentenvergaderingen proberen om een sfeer van onderling vertrouwen te realiseren”
- Open systeem model: “Over de missie of visie van de school praten met contactpersonen buiten school (bestuur, gemeente, andere scholen, ouders, e.d.)” en “Extra budget creëren zodat de school financiële armslag heeft om eigen keuzes te maken”.

In de vragenlijst kan worden aangegeven in welke mate de schoolleider het specifieke gedrag vertoont in een normaal (gemiddeld) werkjaar. De OGS-items waren oorspronkelijk voorzien van zes antwoordmogelijkheden (Krüger, 1994). Voor de nieuwe vragenlijst is gekozen voor een vierpunts Likert-schaal, om een vlotte beantwoording van de vragenlijst mogelijk te maken en om een neutrale antwoordmogelijkheid (die bestaat bij een oneven aantal antwoordcategorieën) te vermijden. De antwoordmogelijkheden lopen van ‘nooit’ tot ‘heel vaak’.

Er zijn twee versies gemaakt van de vragenlijst over de handelingen van schoolleiders: één voor schoolleiders en één voor docenten (zie bijlagen 1A en 1B). Zowel van de OGS als van de originele *Principal Instructional Management Rating Scale* bestaan er ook twee versies. Het idee is dat door zowel schoolleiders als docenten te bevragen een objectiever beeld verkregen kan worden van het functioneren van de schoolleider (zie Krüger, 1994, p. 83). Krüger refereert aan Blase (1987) die pleit voor onderzoek waarin het docentperspectief op schoolleiderschap wordt meegenomen (Krüger, 1994, p. 31), maar zij geeft ook aan dat het in Nederlands onderzoek in de jaren tachtig en negentig al gebruikelijk was om docentpercepties mee te nemen (zie bijv. Brandsma en Stoel, 1987; Brandsma, 1988; Van de Grift, 1985b, 1986; Van de Grift en Akkermans, 1991; Van Marwijk Kooy, 1984). Van de Grift ging oorspronkelijk uit van zelfpercepties van het onderwijskundig leiderschap van schoolleiders om het verband tussen de handelingen van schoolleiders en leerlingprestaties te onderzoeken (Van de Grift, 1986). Hierbij werden directe relaties gevonden tussen schoolleidersgedragingen en leerlingprestaties. Later uitte Van de Grift echter twijfel over deze resultaten, omdat ze waren gebaseerd op zelfbeoordeling van schoolleiders. Uitgaande

van docentpercepties bleek er geen directe relatie aantoonbaar tussen schoolleiderschap en leerlingprestaties (Van de Grift, 1990). Uit het onderzoek van Krüger (1994) bleek eveneens dat er verschillen waren tussen de percepties van schoolleiderschap van de schoolleiders zelf en van de docenten. In recenter onderzoek naar de effectiviteit van schoolleiderschap wordt veelal gewerkt met docentpercepties (zie bijv. Leithwood & Jantzi, 2000, 2006; De Maeyer & Rymenans, 2004; Mulford & Silins, 2003).

3.2.2 Schoolcultuur en schoolorganisatie

Voor de constructie van de vragenlijsten over schoolcultuur en schoolorganisatie werd voortgebouwd op het werk van Houtveen, Voogt, van der Vegt en van de Grift (1995) en van Maslowski (2001). Houtveen et al. ontwikkelden op basis van het concurrerende waarden model een instrument voor het onderzoeken van de organisatiecultuur op basisscholen. Maslowski gebruikte het concurrerende waarden model om de relatie tussen schoolcultuur en schoolprestaties in het voortgezet onderwijs te onderzoeken (zie ook paragraaf 2.5).

Houtveen et al. (1995) baseerden zich bij de ontwikkeling van het onderzoeksinstrument voor de organisatiecultuur op de effectiviteitscriteria die horen bij de vier kwadranten van het concurrerende waarden model (zie tabel 2.2). Uitgaande van deze criteria werden schalen ontwikkeld, waarbij gebruik werd gemaakt van items uit bestaande instrumenten (Fisher & Fraser, 1990; Glaser & Zamanou, 1987; Hoy & Feldman, 1987; Staessens, 1990; Stoel, 1980), al of niet in aangepaste vorm, terwijl ook nieuwe items werden geformuleerd. De vragenlijst van Houtveen et al. werd niet zonder meer toepasbaar geacht voor het onderzoek van Maslowski (2001), vanwege de grote omvang (vijftien schalen en 123 items) en omdat veel items specifiek waren geformuleerd voor het basisonderwijs.

Om na te gaan of er een bruikbaar instrument bestond voor het meten van de verschillende dimensies van schoolcultuur in het voortgezet onderwijs, analyseerde Maslowski (2001) een zestal schoolcultuurvragenlijsten (Cavanagh & Dellar, 1996; Edwards, Green & Lyons, 1996; Houtveen, Voogt, Van der Vegt & Van de Grift, 1995; Pang, 1996; Saphier & King, 1985; Snyder, 1988; Staessens, 1991). Geen van deze instrumenten werd echter als geheel geschikt bevonden. Voor de ontwikkeling van een nieuw instrument maakte Maslowski (2001) wel gebruik van items en schalen uit de geselecteerde vragenlijsten. De eerste vragenlijst die hij construeerde (Form I), bevat 62 items, verdeeld over acht schalen: twee schalen voor elk model in het concurrerende waarden model. Deze vragenlijst is, net als de meeste vragenlijsten waaraan items werden ontleend, gericht op het meten van het gedrag van de schoolleider en docenten en de praktijk op school. Bij een vooronderzoek bleek echter dat niet alle schalen even relevant waren voor scholen (Maslowski, 2001). Uit de resultaten kon worden opgemaakt dat de items minder gekoppeld zouden moeten zijn aan specifieke activiteiten en meer direct zouden moeten refereren aan onderliggende organisatorische waarden (Maslowski, 2001, p. 75, zie ook Rokeach, 1973; Enz, 1986; Van Muijen, 1994; Quinn & Spreitzer, 1991).

Bij de ontwikkeling van een nieuwe versie van de vragenlijst over schoolcultuur maakte Maslowski gebruik van andere instrumenten die waren gebaseerd op het concurrerende waarden model (Cameron, 1985; Cameron & Quinn, 1999; Denison & Mishra, 1995; Enz, 1986; Quinn, 1988; Van Muijen, 1994; Van Muijen, Koopman & De Witte, 1996). De nieuwe vragenlijst (Form II) bestond uit vier schalen en 40 items. Deze vragenlijst bevatte voor elk kwadrant in het concurrerende waarden model tien begrippen, bijvoorbeeld 'prestatiegerichtheid' en 'effectiviteit' voor het rationele doel model, 'stabiliteit' en 'controle'

voor het interne proces model, ‘loyaliteit’ en ‘wederzijds vertrouwen’ voor het ‘human relations’ model en ‘aanpassingsvermogen’ en ‘gerichtheid op verandering’ voor het open systeem model. De relatieve ‘sterkte’ van de vier organisatie modellen binnen een school werd bepaald door te meten hoeveel waarde door docenten werd toegekend aan deze begrippen. De nieuwe schalen bleken zeer betrouwbaar (Cronbach’s $\alpha > 0.90$).

Ondanks de hoge betrouwbaarheid van de tweede serie schalen van Maslowski werd voor het hier gerapporteerde onderzoek echter besloten om toch aan te sluiten bij de eerste benadering (Form I) van Maslowski (2001) en het instrument van Houtveen et al. (1995). De reden hiervoor was, dat volgens het onderzoeksmodel (zie figuur 2.5) niet alleen de schoolcultuur, maar ook de schoolorganisatiepraktijk in kaart gebracht zou moeten worden. De verschillen tussen deze twee begrippen zouden tot uiting moeten komen in de (uitgebreidere) formulering van de items. Bij de constructie van de vragenlijsten werd uitgegaan van vier dimensies, overeenkomend met de vier organisatie modellen in het concurrerende waarden model en de bijbehorende effectiviteitscriteria (zie boven).

De vragenlijst over *schoolcultuur* die voor dit onderzoek is ontwikkeld, bestaat uit vier schalen met elk tien items (zie bijlage 1C). Ongeveer tweederde van de items is, in gewijzigde vorm, overgenomen uit de vragenlijsten van Maslowski (2001, Form I) en Houtveen et al. (1995) Voor de formulering van de overige items is gebruik gemaakt van het overzicht van schooleffectiviteitsbevorderende factoren en hun componenten van Scheerens en Bosker (1997), het *ZEBO instrument voor zelfevaluatie in het basisonderwijs* van Hendriks en Bosker (2003), de *Leidraad Periodiek KwaliteitsOnderzoek* voor het voorgezet onderwijs (2002-2003) van de Inspectie van het Onderwijs en het onderzoek van Witziers (1992) naar coördinatie binnen scholen voor voortgezet onderwijs. Voorbeelden van items zijn:

- Rationele doel model: “Bij ons op school hebben wij hoge verwachtingen van de leerprestaties van alle leerlingen.” en “Bij ons op school vinden wij dat de leerprestaties centraal moeten staan.”
- Interne proces model: “Bij ons op school hechten wij waarde aan heldere richtlijnen en procedures.” en “Bij ons op school zijn wij gericht op continuïteit in de schoolorganisatie.”
- ‘Human relations’ model: “Bij ons op school vinden wij investeren in het personeel van groot belang.” en “Bij ons op school zijn wij erop gericht om met elkaar samen te werken.”
- Open systeem model: “ Bij ons op school hechten wij aan een duidelijke profilering de school.” en “Bij ons op school vinden wij een positieve houding tegenover onderwijsvernieuwingen van belang.”

De items worden gemeten door middel van een vierpunts Likert-schaal die loopt van ‘niet of nauwelijks’ tot ‘in zeer sterke mate’.

De vragenlijst over de *schoolorganisatiepraktijk* voor dit onderzoek is vooral geïnspireerd op het werk van Houtveen et al. (1995). De meeste items zijn echter speciaal voor deze vragenlijst geconstrueerd. Daarbij werd gebruik gemaakt van dezelfde bronnen als voor de schoolcultuurvragenlijst (zie boven) en daarnaast het onderzoek van Vermeulen (1997) naar de school als arbeidsorganisatie, een overzicht van kortlopend onderwijsonderzoek over de school als professionele organisatie (van der Linden, Teurlings & Vermeulen, 2003) en een onderzoek van Witziers en de Groot (1993) naar de marketingmiddelen die scholen inzetten om hun concurrentiepositie te verbeteren. Verder zijn enkele items ontleend aan de *Principal*

Instructional Management Rating Scale van Hallinger (1989, 1994) en de vragenlijst *Onderwijskundige Gerichtheid van Schoolleiders* van Krüger (1994).

De vragenlijst over de schoolorganisatiepraktijk bevat net als de vragenlijst over de schoolcultuur vier schalen met elk tien items (zie bijlage 1D), bijvoorbeeld:

- Rationele doel model: “Bij ons op school wordt aan leerlingen duidelijk gemaakt dat hoge leerprestaties verwacht worden.” en “Bij ons op school worden de resultaten van leerlingen systematisch geanalyseerd.”
- Interne proces model: “Bij ons op school zijn er procedures om ervoor te zorgen dat iedereen tijdig van de juiste informatie wordt voorzien.” en “Bij ons op school zijn er heldere taak- en functieomschrijvingen voor docenten.”
- ‘Human relations’ model: “Bij ons op school is er een uitgewerkt nascholingsplan.” en “Bij ons op school maken docenten gebruik van intervisie om van elkaar te leren.”
- Open systeem model: “Bij ons op school worden marketingactiviteiten ondernomen om de concurrentiepositie te verbeteren.” en “Bij ons op school worden activiteiten ondernomen om het imago van de school te onderzoeken en te verbeteren.”

De items worden gemeten door middel van een vierpunts Likert-schaal die loopt van ‘niet of nauwelijks’ tot ‘in hoge mate’. Deze antwoordmogelijkheden verschillen van die voor de schoolcultuurvragenlijst vanwege nuanceverschillen in de formulering van de items. Dit verschil is van weinig betekenis en achteraf gezien overbodig.

De verschillen in de formulering van de items in de twee vragenlijsten zouden moeten bijdragen aan het onderscheid tussen de begrippen ‘schoolcultuur’ en ‘schoolorganisatiepraktijk’. Bij de schoolcultuur gaat het om het ‘belang’ van bepaalde zaken en de ‘waarde’ die eraan gehecht wordt, terwijl in de vragenlijst over de schoolorganisatiepraktijk gevraagd wordt naar het daadwerkelijk voorkomen van bepaalde processen en activiteiten. Tijdens de constructie van de vragenlijsten werden de formulering van de items en de samenstelling van de schalen in verschillende rondes uitgewerkt en bijgesteld aan de hand van commentaar en suggesties van de leden van de samenwerkende onderzoeksteams. Bovendien werd gebruik gemaakt van items uit bestaande instrumenten die uitvoerig getest zijn. Hiermee werd een zo groot mogelijke inhoudsvaliditeit nagestreefd (Swanborn, 1994, p. 189). Tussentijds werd een gevorderde versie van de vragenlijsten becommentarieerd door drie schoolleiders, waarbij het ging om ‘face validity’, of validiteit-op-het-eerste-gezicht (Swanborn, 1994, p. 189) van de items. De schoolleiders letten er vooral op of de items duidelijk waren geformuleerd en voldoende van betekenis voor het karakteriseren van de situatie op school. Zowel de vragenlijst over schoolcultuur als die over de schoolorganisatiepraktijk zijn bedoeld voor docenten, daarom werd de voorlaatste versie van de vragenlijsten becommentarieerd door twee docenten. Uit het commentaar van de docenten kon worden opgemaakt dat de meeste items voldoende duidelijk waren. Enkele suggesties voor de verbetering van items zijn verwerkt in de laatste versie van de vragenlijsten.

3.2.3 Schoolbeleving van leerlingen

De vragenlijst over de schoolbeleving van leerlingen is gebaseerd op de *Student Engagement and Family Culture Survey* van Leithwood c.s. (1999a, 1999b, 2000a, 2000b; Leithwood, Aitken & Jantzi, 2000). Dit instrument bevat items over de participatie in schoolactiviteiten, de identificatie met de school, de kwaliteit van de instructie en de ‘family educational culture’ (onderwijsondersteuning van thuis). De items over participatie in onderwijsactiviteiten zijn in dit onderzoek niet gebruikt, omdat de samenhang van schoolleiderschap met de participatie

van leerlingen (nog) veel kleiner is dan de samenhang met de identificatie van leerlingen (zie par. 2.4). Van de 40 items in de *Student Engagement and Family Culture Survey* zijn er 25 overgenomen in de vragenlijst over schoolbeleving in dit onderzoek. Items die niet werden overgenomen werden minder toepasselijk geacht voor leerlingen in het Nederlandse voortgezet onderwijs. Voor de vertaling van items in het Nederlands en voor aanvulling van de schalen met passende items is gebruik gemaakt van een vragenlijst over de schoolbeleving van leerlingen van Krüger (1994), een instrument voor het meten van het onderwijsleerklimaat van Houtveen et al. (Houtveen, Vermeulen & van de Grift, 1993), een vragenlijst over het welbevinden van leerlingen in het secundair onderwijs in Vlaanderen (Engels, Aelterman, Scheppens & Van Petegem, 2003) en de *Leidraad Periodiek KwaliteitsOnderzoek* voor het voortgezet onderwijs (2002-2003) van de Inspectie van het Onderwijs. Een enkel item werd ontleend aan het PISA 2000 onderzoek (OECD, 2003).

Het onderzoeksinstrument voor het meten van de schoolbeleving van leerlingen bestaat uit 42 items, verdeeld over vier schalen (zie bijlage 1E). Voorbeelden van items zijn:

- Zich thuisvoelen op school:
 - Omgang met medeleerlingen: “Op deze school kun je echte vrienden maken”
 - Omgang met leraren: “Ik kan goed opschieten met de meeste leraren”
- Waardering:
 - Waardering van onderwijs: “Ik vind mijn schoolwerk echt belangrijk”
 - Waardering van de school: “De sfeer bij ons op school is goed”
- Kwaliteit van de instructie:
 - Resultaatgerichtheid: “De meeste leraren verwachten dat ik altijd mijn best doe”
 - Ondersteuning door leraren: “De meeste leraren doen erg hun best om leerlingen te helpen”
 - Klassenklimaat: “Tijdens de meeste lessen wordt de tijd goed besteed”
 - Organisatie van het onderwijs: “De meeste lessen hebben een goede structuur”
 - Invulling van de lessen: “De meeste leraren zorgen voor afwisselende activiteiten tijdens de lessen”
- Onderwijsondersteuning van thuis (‘Family educational culture’):
 - “Ik bespreek mijn schoolwerk vaak met mijn ouders/verzorgers”

De items worden gemeten met een vierpunts Likert-schaal die loopt van ‘helemaal mee oneens’ tot ‘helemaal mee eens’.

De vragenlijst over schoolbeleving is gebaseerd op een bestaand meetinstrument dat uitvoerig is getest. Voor de aanpassing aan de Nederlandse situatie en aanvulling van de items is daarnaast gebruik gemaakt van andere bestaande meetinstrumenten. Bij de samenstelling van de schalen waren verschillende onderzoekers betrokken. Een docent controleerde of de items voldoende begrijpelijk zouden zijn voor leerlingen.

3.3 Opzet en uitvoering van de dataverzameling

De leerprestaties van leerlingen worden door vele factoren beïnvloed. Om het effect van de handelingen van schoolleiders, de schoolcultuur en de schoolorganisatiepraktijk op de leerprestaties te kunnen onderzoeken, is het van belang om de invloed van andere factoren op deze variabelen goed in kaart te brengen of te minimaliseren. In dit onderzoek is ervoor gekozen om het onderzoek te richten op de havo omdat het hier gaat om een homogene opleiding met relatief geringe verschillen in aanleg tussen leerlingen (zie ook Houtveen, Vermeulen & van de Grift, 1993, p. 77). Bij de opzet van het onderzoek werd ervan uitgegaan

dat voor het uitvoeren van de benodigde (meerniveau-)analyses op tenminste 100 scholen gegevens verzameld zouden moeten worden (zie bijv. Hox & Maas, 2001). De meeste havo-opleidingen maken deel uit van scholengemeenschappen waarin meer vormen van voortgezet onderwijs (vwo, havo en vmbo) zijn gecombineerd. Scholengemeenschappen hebben vaak verschillende vestigingen of locaties. Voor dit onderzoek zijn VO-vestigingen met een havo-bovenbouw benaderd. Op de scholen die medewerking wilden verlenen aan het onderzoek werden gegevens verzameld onder de leerlingen van havo-5 klassen en docenten in de havo-bovenbouw. De schoolleider die centraal staat in dit onderzoek is de locatieleider, ofwel de eindverantwoordelijke schoolleider van een locatie met een havo-bovenbouw.

Hieronder wordt eerst een nadere beschrijving gegeven van de positie van schoolleider in het Nederlandse voortgezet onderwijs en van de functie van locatieleider. Daarna wordt uiteengezet hoe de werving van scholen heeft plaatsgevonden. Tenslotte wordt besproken hoe de gegevensverzameling is uitgevoerd en hoe de eerste resultaten zijn gerapporteerd aan de deelnemende scholen.

3.3.1 De schoolleider

In het kader van het OECD-project 'Improving School Leadership' is in 2007 een beschrijving gemaakt van het leiderschap op scholen in Nederland, getiteld 'Schoolleiders in Nederland' (EIM, 2007). Hierin wordt de situatie als volgt gekenschetst

Typend voor de Nederlandse situatie is dat er een enorme diversiteit is aan namen voor schoolleiders: er is vrijheid in het benoemen van de functie. Het gaat bij schoolleiders om functionarissen die leiding geven binnen een school. Leiderschap wordt in toenemende mate gedeeld tussen of over verschillende functionarissen, zeker bij grote complexe organisaties waar duizenden leerlingen staan ingeschreven. Hoe de leidinggevende taken over deze functionarissen worden verdeeld, en wat de functies precies inhouden, verschilt per situatie. (p. 37)

In het voorgezet onderwijs zijn de volgende leiderschaps- en managementfuncties te onderscheiden: voorzitter van het College van Bestuur, (centrale) directeur, rector, sector/afdelingsdirecteur, locatiedirecteur en teamleider. Daarnaast spelen leraren een belangrijke rol bij de ondersteuning van het management in het voorgezet onderwijs (EIM, 2007, p. 39).

De grote scholengemeenschappen in Nederland hebben vaak een centrale directie met een algemeen directeur die soms rector wordt genoemd. Meestal heeft de scholengemeenschap verschillende locaties of vestigingen. Een locatie omvat een of meer sectoren (basisvorming, vmbo, tweede fase), of delen van sectoren. De locaties kunnen een eigen directeur hebben (locatiedirecteur of sectordirecteur), maar de schoolleiding kan ook zijn verdeeld over verschillende afdelingsdirecteuren (brugklassen, havo/vwo 2 en 3, bovenbouw vwo, e.d.). Afdelingsdirecteuren hebben in het algemeen meer uitvoerende taken dan locatie- of sectordirecteuren. Naast de locatie- of sectordirecteur zijn er enkele adjunct-directeuren die een speciaal deel van het totale takenpakket van de schoolleiding vervullen (bijvoorbeeld personeelszaken, financiën, kwaliteitszorg) of leiding geven aan verschillende afdelingen. Wanneer een (locatie-)directeur rector wordt genoemd, heten de andere directieleden meestal conrector. Een teamleider coördineert een team van docenten (bijvoorbeeld van het eerste leerjaar). Een scholengemeenschap kan alleen of samen met andere scholengemeenschappen worden overkoepeld door een College van Bestuur.

Volgens de beschrijving van ‘Schoolleiders in Nederland’ (EIM, 2007, p. 40) hebben (centrale) directeuren een breed takenpakket. De directeuren hebben strategische, coördinerende en uitvoerende taken op het gebied van onderwijs en personeel. Verder hebben ze strategische taken op het terrein van financieel en materieel beheer. Adjuncten, sector- en locatiedirecteuren zijn meer belast met uitvoerende taken, terwijl een deel ook strategische en coördinerende taken heeft op het terrein van onderwijs.

In de onderwijswetgeving zijn er geen regels opgenomen voor de taken, functies en bevoegdheden en ook niet voor de opleiding en de kwaliteit van de schoolleiders (EIM, 2007, p. 101). In 2000 is door de Vereniging voor het management in het Voortgezet Onderwijs (VVO, de voorloper van Schoolmanagers_VO), in samenwerking met schoolleiders, ISIS (een projectorganisatie) en de VSWO (Vereniging Samenwerkende Werkgeversorganisaties Onderwijs), een beroepsprofiel ontwikkeld, dat in 2003 werd geactualiseerd. Dit profiel geeft een overzicht van de benodigde competenties per kerntaak voor een aantal niveaus binnen het management. De kerntaken zijn (www.werkeninhetonderwijs.nl, feb. 2007):

- Het ontwikkelen van een onderwijsvisie en een strategie
- Het aansturen van het primaire onderwijsproces
- Het motiveren en aansturen van medewerkers
- Het managen van middelen

Hierbij worden de volgende niveaus onderscheiden in het schoolmanagement:

- Integraal schoolleider op sector-/locatieniveau;
- Integraal schoolleider op schoolniveau
- Bovenschoolse schoolleider.

De locatieleider valt in de eerste categorie, de centrale directeur in de tweede en de voorzitter van het College van Bestuur in de derde.

Omdat het beroepsprofiel door schoolleiders nog te abstract, te breed en te weinig ontwikkelingsgericht werd bevonden, is in 2007 door de Werkgroep Basiscompetenties een beperkte set van basiscompetenties geformuleerd, met een handleiding voor de competentieontwikkeling (ISISQ5/VO-raad, 2007). De volgende basiscompetenties worden onderscheiden voor schoolleiders in het voortgezet onderwijs: ‘visiegerichtheid’, ‘omgevingsbewustzijn’, ‘nieuwe vormen van leiderschap’, ‘organisatiebewustzijn’ en ‘hogere orde denken’. Bij de ontwikkeling van deze basiscompetenties is uitgegaan van hetzelfde beschrijvingsmodel van schoolleiderschap als in dit onderzoek wordt gebruikt (Bossert et al., 1982; Dwyer et al., 1985; Dwyer et al., 1987). Volgens de Werkgroep Basiscompetenties is de schoolleider in feite (ISISQ5/VO-raad, 2007, p. 15):

- Een organisatie-ontwikkelaar: door handelingen die gericht zijn op het structuur/cultuurdomein.
- Een cultuurmanager: via de strategieën die worden gevormd door handelingen gericht op de cultuur van de school.
- Een onderwijskundig leider: door het managen van het onderwijs en het pedagogisch klimaat.
- Een beheersmatig manager: door de gerichtheid op personeel en faciliteiten.

De schoolleider die centraal staat in dit onderzoek, is de schoolleider die op het niveau van de locatie de bovenstaande rollen vervult, meestal is dit de locatiedirecteur of de sectordirecteur. Deze integrale schoolleider staat dicht bij het primaire proces dan de centrale directeur of de bovenschoolse manager. Hierdoor is de kans groter dat een relatie wordt gevonden tussen de handelingen van de schoolleider en de leerprestaties (zie par. 2.7).

Een voorbeeldbeschrijving van de functie van locatiedirecteur is te vinden in het FunctieWaarderingsysteem Voortgezet Onderwijs 2002 (Voorbeeldfuncties FUWA-VO 2002, z.j.). In dit systeem overlappen de functies van sectordirecteur en locatiedirecteur elkaar. De aard en het niveau van de functie hangen af van de aanwezigheid van een onderwijskundige beleidseenheid (tweede fase havo/vwo, basisvorming of tenminste twee leerwegen vmbo) binnen een locatie, de vrijheid die er is om een eigen onderwijskundige koers te varen en de bijdrage die de sectordirecteur/locatiedirecteur levert aan het (strategisch) beleid van de onderwijsinstelling. De sectordirecteur/locatiedirecteur is onder verantwoordelijkheid van de centrale directie belast met het voorbereiden, ontwikkelen en (mede) uitvoeren van het (onderwijskundige) beleid van zijn sector/locatie, het leveren van een bijdrage aan de ontwikkeling van het centrale beleid van de onderwijsinstelling en het leiding geven aan de sector/locatie. Een sector/locatie kent een directeur, adjunct-directeur(en), teamleiders, docenten en onderwijsondersteunend personeel (OOP). De sectordirecteur/locatiedirecteur vormt met collega-sectordirecteuren/locatiedirecteuren en de centrale directie het centraal managementteam van de onderwijsinstelling. De centrale directie zet de strategische beleidslijnen uit en vertaalt deze in het centrale beleid. In de centrale directie leveren de sectordirecteuren/locatiedirecteuren een bijdrage aan de meningsvorming en eventueel ook een bijdrage aan de strategische beleidsontwikkeling van de onderwijsinstelling (hoogste functieniveau).

3.3.2 Werving van scholen

De werving van scholen en de dataverzameling vonden plaats in twee rondes, de eerste in het schooljaar 2003-2004 en de tweede in het schooljaar 2004-2005. In november 2003 zijn de locatieleiders van alle 484 locaties van VO-scholen met een havo-bovenbouw via een brief geïnformeerd over het onderzoek en uitgenodigd om deel te nemen. Er is geen steekproef getrokken uit de populatie van scholen, omdat werd gestreefd naar tenminste 100 deelnemende scholen, terwijl een respons werd verwacht van niet meer dan 25% (op basis van ervaring in de onderzoekteams). Van januari tot en met april 2004 hebben drie onderzoekers een aantal belrondes gehouden, om scholen te werven voor het onderzoek. In deze periode vonden de onderzoekers 80 schoolleiders bereid om mee te doen aan het onderzoek, terwijl op 69 scholen de gegevensverzameling daadwerkelijk werd gerealiseerd. Na de eerste ronde bleven 122 scholen over waarmee nog geen contact was gelegd. Het kostte vaak vele pogingen voordat telefonisch contact met de drukbezette schoolleiders tot stand kwam. De periode voor de dataverzameling was bovendien beperkt, doordat de 5-havo klassen later in het schooljaar niet meer beschikbaar waren voor het invullen van de leerlingvragenlijsten, vanwege het centraal schriftelijke eindexamen. De respons in de eerste ronde (uitgaande van het aantal scholen waar gegevens verzameld zijn) was 19%.

De overblijvende 122 scholen zijn in oktober 2004 benaderd in een tweede wervingsronde. De tweede wervingsronde is uitgevoerd om het streefaantal van 100 scholen te bereiken. Deze keer werd aan schoolleiders gevraagd om door middel van een antwoordformulier aan te geven of er interesse bestond in het onderzoek. Verder werd een aantal scholen uit de eerste ronde opnieuw gebeld, omdat deze scholen hadden aangegeven op een later tijdstip te willen deelnemen aan het onderzoek. Van de 122 scholen die schriftelijk werden benaderd, reageerden er 28 positief, waarna op 25 scholen de dataverzameling daadwerkelijk uitgevoerd kon worden. De respons bij deze procedure is 20%. Van de scholen die opnieuw gebeld werden deden er nog 9 mee. In totaal hebben 103 van de 484 VO-vestigingen met een havo-bovenbouw meegewerkt aan het onderzoek. De totale respons komt daarmee op 21%.

De telefonische werving van schoolleiders (eerste ronde) en de schriftelijke procedure met het antwoordformulier (tweede ronde) leverden een vrijwel gelijke respons op. De schriftelijke procedure was echter veel efficiënter, omdat direct contact gelegd kon worden met geïnteresseerde schoolleiders. Het totale responspercentage (21%) was ongeveer zo groot als werd verwacht. Ter vergelijking: Maslowski (2001) haalde bij zijn onderzoek naar schoolcultuur een respons van 28% van de benaderde scholen; Houtveen, Vermeulen en van de Grift (1993) bereikten bij hun onderzoek naar het onderwijsleerklimaat een respons van 19% van alle scholen met een havo-afdeling. Schoolleiders die niet wilden deelnemen aan het onderzoek gaven aan dat zij heel vaak (soms zelfs wekelijks) benaderd worden met verzoeken om medewerking aan onderzoek, zodat zij moeten selecteren. Verder gaven schoolleiders aan dat zij docenten niet met het onderzoek wilden belasten, vanwege de grote werkdruk in het algemeen, of bijvoorbeeld omdat er op de school gewerkt werd aan een bijzonder (onderwijsvernieuwings-)project. Ook enkele scholen die in eerste instantie hadden toegestemd in deelname aan het onderzoek vielen tussentijds af, bijvoorbeeld omdat de belasting van het onderzoek toch te hoog was voor de schoolleider of de docenten, of vanwege ziekte (in de schoolleiding) of vertrek van de schoolleider.

3.3.3 Gegevensverzameling en schoolrapportage

De onderzoeksinstrumenten voor het totale UvA-UT-onderzoeksproject zijn opgenomen in drie vragenlijsten:

- De vragenlijst schooldirecteuren, met de volgende onderdelen:
 1. Persoonlijke gegevens
 2. Gegevens over de school
 3. Gegevens over externe factoren (voor het UvA onderzoek)
 4. Opvattingen van de schoolleider over de eigen rol (voor het UvA onderzoek)
 5. Handelingen van de schoolleider
- De vragenlijst docenten, met de onderdelen:
 1. Persoonlijke gegevens
 2. Waarneming van de handelingen van de schoolleider
 3. Beleving van de school: schoolorganisatie en schoolcultuur
- De leerlingenvragenlijst:
 1. Beleving van de school: zich thuisvoelen op school, waardering, kwaliteit van de instructie, 'family educational culture' van thuis

De vragenlijsten zijn geproduceerd in de vorm van optisch leesbare formulieren door ELION, de afdeling voor elektronische gegevensverwerking, automatisering en onderzoeks-ondersteuning van het SCO-Kohnstamm Instituut van de Universiteit van Amsterdam.

Op de deelnemende scholen zijn de vragenlijsten voorgelegd aan:

- de eindverantwoordelijke schoolleider van de vestiging
- vijftien docenten van de havo bovenbouw
- de leerlingen van twee havo-5 klassen.

De aantallen docenten en leerlingen zijn zodanig gekozen dat een voldoende betrouwbaarheid van de (geaggregeerde) variabelen bereikt kan worden (zie hoofdstuk 4).

Om de gegevensverzameling op de scholen te ondersteunen, hebben de uitvoerende onderzoekers alle deelnemende scholen bezocht. Tijdens het schoolbezoek vulden de schoolleiders hun vragenlijst direct in, waardoor de respons van schoolleiders was gewaarborgd. Ondertussen selecteerden de onderzoekers willekeurig vijftien docenten uit de

havo-bovenbouw en twee 5-havo-klassen. De geadresseerde docenten- en leerlingenvragenlijsten werden overgedragen aan de schoolleider, die zorg droeg voor de verspreiding. Ter afsluiting van het schoolbezoek vond een kort nagesprek plaats, waarin het organogram van de school werd besproken en de specifieke taken van de schoolleider.

Tabel 3.1 geeft een overzicht van de aantallen geretourneerde vragenlijsten. Op alle 103 deelnemende scholen is de schoolleidersvragenlijst ingevuld. Drie scholen stuurden geen docentenvragenlijsten terug, vijf scholen geen leerlingenvragenlijsten.

Tabel 3.1 Overzicht van aantallen geretourneerde vragenlijsten

Schoolleiders- vragenlijsten	Docentenvragenlijsten			Leerlingenvragenlijsten		
	aantal scholen	gemiddeld per school	totaal	aantal scholen	gemiddeld per school	totaal
103	100	10	998	98	44	4336

Per school vulden gemiddeld tien docenten de docentenvragenlijst in. Omdat vijftien vragenlijsten per school waren uitgezet bedraagt de gemiddelde respons van docenten 67%. De respons van leerlingen is niet nagegaan. De leerlingenvragenlijsten werden meestal tijdens een les door alle aanwezige leerlingen ingevuld. In totaal werkten 998 docenten en 4336 leerlingen mee aan het onderzoek.

De geretourneerde vragenlijsten zijn door ELION gescand en verwerkt in databestanden. Alle deelnemende scholen ontvingen als tegenprestatie een schoolrapport. In dit rapport zijn de scores van de school en de landelijke gemiddelden opgenomen, voor de schoolbeleving van leerlingen, de schoolcultuur, de schoolorganisatie en de handelingen van de schoolleider. Om scholen meer inzicht te bieden in hun positie ten opzichte van andere scholen, is hierbij aangegeven of de school behoort tot de 25% hoogst scorende scholen, de middelste 50% of de 25% laagst scorende scholen op een bepaald gebied.

3.4. Representativiteit van de deelnemende scholen; achtergrondkenmerken

Om de representativiteit van de onderzoeksgroep te bepalen, zijn de gegevens van de deelnemende scholen vergeleken met de gegevens van de totale populatie van havo-bovenbouw. De uitkomsten worden hieronder beschreven. Verder worden achtergrondgegevens gepresenteerd van scholen, schoolleiders en docenten.

3.4.1 Achtergrondkenmerken en representativiteit van de deelnemende scholen

Van de 103 scholen die meewerkten aan het onderzoek stuurden er 97 zowel docenten- als leerlingenvragenlijsten terug. De representativiteit van deze groep van 97 scholen voor de gehele populatie van 484 scholen is nagegaan aan de hand van de verdeling over de provincies, de denominatie en de grootte van de scholen. De vergelijkende gegevens zijn ontleend aan een bestand voor de kwaliteitskaart over het schooljaar 2002-2003.

Tabel 3.2 Verdeling van scholen over de provincies

Provincie	% deelnemende scholen	% populatie
Groningen	1.0	3.9
Friesland	5.2	3.3
Drente	4.1	2.3
Overijssel	10.3	6.4
Flevoland	1.0	2.3
Gelderland	13.4	11.4
Utrecht	3.1	8.1
Noord-Holland	18.6	16.3
Zuid-Holland	12.4	22.9
Zeeland	2.1	2.5
Noord-Brabant	18.6	14.0
Limburg	10.3	6.6

In tabel 3.2 staat een overzicht van de verdeling van scholen over de provincies. Uit de tabel is op te maken dat de verschillen in de verdeling over de provincies in het algemeen klein zijn, maar de onderzoeksgroep bevat relatief weinig scholen uit Zuid-Holland. Een Chi-kwadraat toets³ geeft aan, dat de groep van deelnemende scholen wat betreft de verdeling over de provincies (net) als representatief beschouwd mag worden ($\chi^2=19.334$, $df = 11$, $p=0.0535$).

De onderzoeksgroep is niet representatief wat betreft de denominatie ($\chi^2 = 8.657$, $df = 3$, $p=0.034$). Het percentage katholieke scholen is groter en het percentage protestants-christelijke en overige (algemeen bijzondere en samenwerkende) scholen kleiner dan in de gehele populatie (zie tabel 3.3).

Tabel 3.3 Denominatie van scholen

Denominatie	% deelnemende scholen	% populatie
Openbaar	27.8	28.3
Katholiek	38.1	26.0
Protestants-Christelijk	14.4	21.7
Overig	19.6	24.0

De verdeling in schoolgrootte (het aantal leerlingen op de locatie) is weergegeven in tabel 3.4. De onderzoeksgroep is representatief wat betreft schoolgrootte ($\chi^2 = 2.818$, $df = 3$, $p=0.421$).

Tabel 3.4 Indeling van scholen naar schoolgrootte

Schoolgrootte	% deelnemende scholen	% populatie
minder dan 500 leerlingen	2.1	4.8
500 tot 1000 leerlingen	32.0	35.1
1000 tot 1500 leerlingen	47.4	40.9
Meer dan 1500 leerlingen	18.6	19.2

De verdeling van de havo-scholen over de verschillende schooltypen is niet vergeleken met landelijke gegevens. De helft van de onderzochte scholen biedt naast de havo ook vwo en vmbo-t aan (zie tabel (3.5)). Een kwart van de onderzochte scholen biedt de combinatie havo-vwo aan en ook een kwart havo-vwo-vmbo (breed). In de steekproef komt één categorale havo-school voor.

³ Hier is het gebruik van de χ^2 toets discutabel, vanwege een laag aantal scholen in sommige provincies.

Tabel 3.5 Schooltypen in de onderzochte scholen

Schooltypen	% deelnemende scholen
Alleen havo	1.0
Havo-vwo	24.3
Havo-vwo-vmbo-t	50.5
Havo-vwo-vmbo	24.3

In tabel 3.6 zijn beschrijvende gegevens opgenomen over het aantal leerlingen, het percentage culturele minderheden en het aantal docenten op de onderzochte scholen (locaties).

Tabel 3.6 Beschrijvende kenmerken van de onderzochte scholen

Kenmerk	gemiddeld	std	minimum	maximum	n scholen
Aantal leerlingen	1146.5	456.8	340	2251	103
% culturele minderheden ¹	8.8	15.8	0	95	97
Aantal docenten	93.2	37.7	30	227	95

¹ Op de gehele school, volgens de inschatting van de schoolleider.

Uit de bovenstaande gegevens kan worden opgemaakt dat de scholen in de onderzoeksgroep zeer uiteenlopen wat betreft grootte, schooltype en samenstelling van de leerlingpopulatie. Deze factoren oefenen invloed uit op het werk van de locatieleider. De denominatie speelt daarbij ook een rol. Voor de gemiddelde leerprestaties blijkt vooral het percentage culturele minderheden van belang. Het percentage culturele minderheden heeft een groter effect op de gemiddelde leerprestaties dan andere schoolfactoren (zie bijv. Rekers-Mombarg, Lodewick & Bosker, 2000).

3.4.2 Achtergrondkenmerken van de schoolleiders

De leeftijd van de locatieleiders in de onderzoeksgroep is gemiddeld 52 jaar. De totale groep van 103 schoolleiders bestaat uit 84 mannen (81.6%) en 19 vrouwen (18.4%). Verreweg de meeste schoolleiders hebben een eerstegraads lerarenopleiding gevolgd (78.6%), enkelen een tweede of derdegraads opleiding. Een klein deel van de schoolleiders (8.7%) volgde geen lerarenopleiding. De schoolleiders hebben gemiddeld 13.4 jaar ervaring als manager in het onderwijs en 6.1 jaar als locatieleider op de onderzochte school. Ter vergelijking: in het onderzoek van Houtveen et al. (1993, p. 98) naar het onderwijsleerklimaat op havo-scholen, was de gemiddelde leeftijd van de ondervraagde schoolleiders 45 jaar, was de verdeling mannen-vrouwen ongeveer gelijk aan de huidige verdeling en waren de schoolleiders gemiddeld 8.6 jaar in functie. De gestegen gemiddelde leeftijd van schoolleiders zal verband houden met de vergrijzing in het onderwijs. Omdat schoolleiders in dit onderzoek centraal staan, zijn schoolleiders met relatief weinig ervaring op de huidige school misschien meer geïnteresseerd en daarom wellicht oververtegenwoordigd in de onderzoeksgroep.

3.4.3 Achtergrondkenmerken van de docenten

De gemiddelde leeftijd van de docenten in de onderzoeksgroep is 48.0 jaar. Van de 997 docenten die de vragenlijst invulden is 71.6% man en 28.4% vrouw. In deze groep is 65% voltijds aangesteld en 35% in deeltijd. Ter vergelijking: het aandeel vrouwelijke docenten in het gehele vo is 45.5% en het aandeel deeltijders is 51.3%. Vrouwelijke docenten werken vaker in deeltijd: 72% over alle onderwijssectoren. (MinOCW, 2007, Nota werken in het onderwijs 2007). Het is mogelijk dat er in de havo-bovenbouw minder vrouwen werken dan

mannen, maar het zou ook kunnen dat deeltijders de vragenlijst minder vaak retourneerden. Het grootste deel van de docenten heeft een eerstegraads opleiding (88.3%), 7.6% heeft een tweedegraads opleiding en de rest heeft een andere opleiding. De docenten hebben gemiddeld 17.7 jaar ervaring op de onderzoeksschool.

3.5 Data-analyse

In deze paragraaf wordt een overzicht gegeven van de analysemethoden die zijn gebruikt voor de validering van de instrumenten en de toetsing van het onderzoeksmodel aan de data.

3.5.1 Betrouwbaarheid en validiteit van onderzoeksinstrumenten

Om de kwaliteit van onderzoeksinstrumenten te bepalen, is de *betrouwbaarheid* en de *validiteit* onderzocht. Bij het bepalen van de betrouwbaarheid van een onderzoek (een of meer waarnemingen) gaat het om de vraag of het onderzoek door toevalsfouten wordt verstoord (Swanborn, 1994, p. 175, *Lexicon Methoden en Technieken*, z.j.). Toevalsfouten kunnen worden vastgesteld door herhaling van het onderzoek. Als er toevalsfouten optreden, levert herhaling een ander resultaat op; als toevalsfouten afwezig zijn, levert herhaling hetzelfde resultaat op. Wanneer het niet mogelijk is om een onderzoek te herhalen, of in twee versies uit te voeren, kan de *interne consistentie* van het onderzoeksinstrument worden onderzocht. Hierbij worden delen van het instrument als elkaars herhaling opgevat. Een veelgebruikte maat voor de interne consistentie is *Cronbach's alpha*. Cronbach's alpha geeft de onderlinge samenhang tussen alle items in een instrument aan met een getal tussen 0 (geen samenhang) en 1 (perfecte samenhang). Bij een alpha van .70 of hoger wordt de interne consistentie als voldoende beschouwd. In dit geval worden dus de items van een instrument opgevat als herhaalde metingen van een begrip (Garson, z.j., c., *Lexicon methoden en technieken*, z.j.). Aan Cronbach's alpha kleven wel twee belangrijke bezwaren (Rowe, 2002; Vehkalati, Puntanen & Tarkkonen, 2006): de grootte van alpha is afhankelijk van het aantal items in een schaal en in het algemeen onderschat alpha de betrouwbaarheid, omdat aan bepaalde aannames niet wordt voldaan. De interne consistentie van een schaal kan ook worden bepaald met behulp van factoranalyse (Garson, z.j., d.). Als de indicatoren voor een bepaald construct eenduidig en in voldoende mate op de bijhorende factor laden is de interne consistentie goed.

De validiteit van een onderzoek wordt bepaald door de mate waarin *systematische fouten* voorkomen (Swanborn, 1994, p. 176, *Lexicon methoden en technieken*, z.j.). Systematische fouten werken steeds in dezelfde richting en kunnen niet worden aangetoond door herhaling van het onderzoek. Voor de validiteit van waarnemingen kunnen de volgende vormen worden onderscheiden: *gezichtsvaliditeit* (*face validity*), *inhoudsvaliditeit* (*content validity*), *constructvaliditeit* (*construct validity*) en *criteriumvaliditeit* (*criterion validity*).

Tijdens het construeren van de onderzoeksinstrumenten is aandacht besteed aan de *inhoudsvaliditeit*, door middel van een beoordeling van de geschiktheid van items en schalen door deskundigen (zie par. 3.2). In hoofdstuk 4 komt de *construct- of begripsvaliditeit* aan de orde (Swanborn, 1994, p. 189; Garson, z.j., d.). Hierbij gaat het erom of de begrippen op de juiste wijze geoperationaliseerd zijn. De vraag bij constructvaliditeit is of relaties die volgens de theorie bestaan tussen begrippen, ook worden waargenomen tussen de operationalisering van de begrippen (*Lexicon methoden en technieken*, z.j.). Twee belangrijke onderdelen van constructvaliditeit zijn *convergerende validiteit* en *discriminante validiteit*. Verschillende instrumenten om hetzelfde begrip te meten, moeten hetzelfde resultaat opleveren, ofwel

convergeren. Als met een meetinstrument verschillende begrippen gemeten worden, moet de methode deze begrippen voldoende van elkaar onderscheiden (discrimineren).

Convergerende validiteit kan worden bepaald door na te gaan in hoeverre de meetwaarden correleren met de uitkomsten van andere onderzoeken, of met bepaalde standaardmaten of criteria (dit wordt *criteriumvaliditeit* genoemd), of door de uitkomsten voor verschillende steekproeven of verschillende meetmethoden te vergelijken (Garson, z.j., d.). De interne consistentie van een instrument kan ook worden gezien als een vorm van convergerende validiteit. Een redelijke correlatie tussen de items wijst op een goede convergerende validiteit. Discrimante validiteit kan bijvoorbeeld worden getest door middel van het berekenen van correlaties tussen meetwaarden (schalen) voor verschillende begrippen, of door factoranalyse (Garson, z.j., d.). De discriminante validiteit is in orde als twee schalen niet teveel correleren (vuistregel: $r \leq .85$), of als de indicatoren voor de twee schalen laden op verschillende factoren, of als de passing van het model bij de factoranalyse goed is.

3.5.2 Optimalisering van meetmodellen met behulp van factoranalyse

Voor de validering van de onderzoeksinstrumenten (hoofdstuk 4) is gebruik gemaakt van zowel confirmatieve, als exploratieve factoranalyse (zie bijlage 2A). De confirmatieve factoranalyse is uitgevoerd door middel van *structural equation modeling* (SEM) in LISREL versie 8.8 (zie bijlage 2B). De exploratieve factoranalyse is uitgevoerd in SPSS versie 12.0, in de vorm van een principale componenten analyse. Om de construct-validiteit vast te stellen van de onderzoeksinstrumenten die zijn gebaseerd op het concurrerende waarden model (handelingen van schoolleiders, schoolcultuur en schoolorganisatiepraktijk), werd in een confirmatieve factoranalyse getest in hoeverre een meetmodel met vier factoren (één voor elke organisatie-model, zie bijlagen 1A-D) past op de data. Voor de schoolbelevingsvragenlijst werd ook een vier factorenmodel getest (één voor elke schaal, zie bijlage 1E). Naar aanleiding van de uitkomsten van de confirmatieve factoranalyse werden de modellen bijgesteld en opnieuw getest. Deze aanpak wordt 'model genererend' genoemd (Jöreskog & Sörbom, 1993a, p. 115, zie bijlage 2C). Hoewel er op theoretische gronden wordt uitgegaan van een bepaalde factorstructuur, is het niet zeker dat alle items die zijn geformuleerd, voldoende laden op de bijbehorende factor. Het is bovendien ook mogelijk dat items laden op meer dan één factor of op een andere factor dan verwacht. Om de modellen te kunnen bijstellen, is bij de confirmatieve factoranalyse gekeken naar de hoogte van de factorladingen, de modelpassing en de modificatie indices die LISREL genereert (zie bijlage 2D). De modelpassing is bepaald aan de hand van de chi-kwadraat test (zie bijlage 2E) en een aantal andere *goodness-of-fit indices* (zie bijlage 2F), met name de *root mean square error of approximation* (RMSEA), de *standardized root mean square residual* (SRMR) en *comparative fit index* (CFI). Er is ook gebruik gemaakt van de wat minder betrouwbare (*adjusted*) *goodness of fit index* (GFI of AGFI) omdat dit criterium in de literatuur regelmatig voorkomt.

De ontwikkeling van alternatieve modellen werd ondersteund door exploratieve factoranalyse. Hiermee kon de structuur van de data worden onderzocht. Door middel van principale componenten analyse (hoofdcomponentenanalyse) werd vastgesteld welke factoren het grootste deel verklaren van de variantie, die in de oorspronkelijke geobserveerde variabelen (indicatoren) aanwezig is. De factorstructuur die zichtbaar wordt bij de exploratieve factoranalyse helpt om de uitkomsten van de confirmatieve factoranalyse te verklaren. In de uitkomsten van de exploratieve factoranalyse is op eenvoudige wijze te zien of items laden op meer dan één factor of op een andere factor dan verwacht. Bij een confirmatieve factoranalyse

is dat alleen af te leiden uit een lage factorlading en een lijst van modificatie indices die LISREL genereert.

3.5.3 Meerniveau confirmatieve factoranalyse

De databestanden in dit onderzoek bevatten gegevens van schoolleiders en van groepen docenten en leerlingen binnen verschillende scholen. Voor de docenten- en leerlingengegevens hebben de data een *hiërarchische* structuur: leerlingen en docenten zijn geclusterd (*genest*) in scholen. De Maeyer en Rymenans beschrijven in hun proefschrift (2004, p. 260) waarom er voor deze hiërarchisch gestructureerde data het beste gekozen kan worden voor meerniveau confirmatieve factoranalyse (MCFA of MFA). De eerste reden is statistisch van aard: de gegevens van docenten en leerlingen zijn niet onafhankelijk. Als het bijvoorbeeld gaat om de waarneming van bepaalde school(leiders)kenmerken, zullen twee willekeurige docenten binnen dezelfde school meer gemeenschappelijk hebben, dan twee willekeurige docenten van verschillende scholen. Wanneer de docentgegevens in de confirmatieve factoranalyse als onafhankelijke, individuele gegevens worden behandeld, kunnen er afwijkingen ontstaan in de parameterschattingen. Als vervolgens de uitkomsten geaggregeerd worden naar het schoolniveau, kunnen er grote meetfouten optreden (Snijders & Bosker, 1999; Hox, 2002).

De tweede reden die door De Maeyer en Rymenans (2004, p. 260) wordt aangegeven om te kiezen voor meerniveau confirmatieve factoranalyse, is van inhoudelijke aard. De score van een individuele docent (of een leerling) op een item dat betrekking heeft op een bepaald schoolkenmerk zal aan de ene kant worden bepaald door de feitelijke situatie op school (de gemiddelde score op een item binnen een school, met meetfouten) en aan de andere kant door de latente mening van de docent zelf over dit kenmerk. Daarbij moet er ook op het niveau van de docent nog rekening worden gehouden met meetfouten. Om een zo zuiver mogelijke schatting te krijgen van een schoolkenmerk kunnen de twee concepten –de mening van de docent en de mate waarin het schoolkenmerk effectief aanwezig is op school, het beste van elkaar worden losgekoppeld door middel van meerniveau factoranalyse (De Maeyer & Rymenans, 2004, p. 261). In een meerniveau confirmatieve factoranalyse kan de passing van een meetmodel zowel op schoolniveau en als op docentenniveau (of leerlingenniveau) worden bepaald. In principe is het mogelijk om op schoolniveau en op docentniveau verschillende modellen te valideren (Heck & Thomas, 2000; Hox, 2002). Verondersteld wordt dat een schoolkenmerk optimaal wordt gemeten door uit te gaan van het best passende model op schoolniveau. De Maeyer en Rymenans (2004) gebruikten de uitkomsten van een meerniveau confirmatieve factoranalyse die werd uitgevoerd met behulp van de software Mplus (Muthén & Muthén, 1998-2001), om factorscores op schoolniveau te berekenen voor hun kernconcepten geïntegreerd leiderschap en prestatiegericht leerklimaat (De Maeyer, persoonlijke communicatie, juli 2006).

3.5.4 Aanpak van de meerniveau factoranalyse; intraklasse correlatie en design effect

Meerniveau analyse methoden zijn in de jaren tachtig ontwikkeld en worden vanaf het begin van de jaren negentig toegepast in onderzoek (Snijders & Bosker, 1999). Procedures voor het gebruik van ‘*Structural Equation Modeling*’ (SEM) bij meerniveau data zijn pas vanaf het eind van de jaren negentig beschikbaar. Meerniveau confirmatieve factoranalyse is voor het eerst beschreven door Muthén (1991). Muthén ontwikkelde een meerniveau schattingsmethode die kon worden gebruikt in combinatie met bestaande structural equation modeling software (zie Muthén, 1994). Vanaf 1998 is deze techniek beschikbaar in het

software programma Mplus (Muthén & Muthén, 1998-2001). In datzelfde jaar constateren Rowe en Hill (1998) nog: “However, routines for accommodating meerniveau data within the context of structural equation modeling are still under development and not readily accessible” (p. 309). Heck en Thomas (2000) beschrijven in hun boek over ‘multilevel modeling techniques’, hoe meerniveau factoranalyse kan worden uitgevoerd met Mplus, LISREL (Jöreskog & Sörbom, 1993b) en STREAMS (Gustaffson & Stahl, 1996). De uitvoering in LISREL is dan nog ingewikkeld, omdat de meerniveau benadering van Muthén (1991) nog niet is ingebouwd. In 2001 kwam versie 8.5 van LISREL beschikbaar, waarmee zowel meerniveau regressieanalyse als meerniveau structural equation modeling kan worden uitgevoerd (du Toit & du Toit, 2001; Mels, 2004).

De strategie voor de aanpak van meerniveau factoranalyse, zoals die is beschreven door Heck & Thomas (2000, p. 110), is afkomstig van Muthén (1994). Volgens deze strategie wordt eerst een éénniveau model getest. De éénniveau test geeft een ruwe indicatie van de geschiktheid van het model en van misspecificaties, zoals zwakke items. Pas als het éénniveau model goed genoeg fit, wordt de volgende analysestap uitgevoerd (Heck & Thomas, 2000, p. 124). Hierbij wordt voor de geobserveerde variabelen de intraklasse correlatie bepaald. De *intraklasse correlatie coëfficiënt* geeft aan in welke mate eenheden binnen een groep op elkaar lijken (Snijders & Bosker, 1999). Als de intraklasse correlatie dicht bij nul ligt, is een meerniveau factor analyse niet nodig (Heck & Thomas, 2000, p. 110). Er kan dan worden uitgegaan van de individuele gegevens.

De intraklasse correlatie coëfficiënt wordt berekend door de variantie tussen de groepen te delen door de totale variantie (Snijders & Bosker, 1999). Als er twee niveau's zijn in de onderzoeksgroep, zoals bij waarnemingen van docenten binnen verschillende scholen, is de totale variantie gelijk aan de som van de variantie tussen de scholen (‘between groups variance’) en de variantie binnen de scholen (‘within group variance’). In het hier beschreven onderzoek naar de effecten van handelingen van schoolleiders geeft de intraklasse correlatie bijvoorbeeld aan in hoeverre er binnen de scholen overeenstemming is tussen de waarnemingen van de docenten. Feitelijk is het de gemiddelde correlatie tussen de waarnemingen van docenten binnen dezelfde school (Hox, 2002, p. 5).

Om de intraklasse correlatie coëfficiënt te berekenen kan een meerniveau regressie analyse worden uitgevoerd zonder verklarende variabele (onconditioneel model, zie bijvoorbeeld Norušis, 2003). Deze procedure levert de variantie tussen scholen en de totale variantie. De intraklasse correlatie kan ook worden benaderd met ANOVA, met behulp van de volgende formule (Garson, z.j., c.; Hox, 2002, p. 14; Snijders & Bosker, 1999, p. 20):

$$ICC = \rho_I = (ms_{\text{between}} - ms_{\text{within}}) / (ms_{\text{between}} + [n-1]ms_{\text{within}})$$

waarbij ms_{between} een schatting is voor de variantie tussen groepen, ms_{within} voor de gemiddelde variantie binnen groepen en n staat voor de groepsgrootte.

Als de groepsgrootte niet constant is, wordt het *harmonische gemiddelde* n' voor N groepen berekend met de formule:

$$n' = (1/[N-1])(\text{SUM}n - [\text{SUM}n^2]/\text{SUM}n).$$

De schatting van de intraklasse correlatie via ANOVA is heel gebruikelijk (Hox, 2002, p. 14), maar volgens Muthén (1997) zou deze benadering te lage waarden kunnen opleveren, omdat de variantie binnen de scholen groter wordt geschat door de meetfout op individueel niveau.

Als vuistregel wordt vaak aangehouden dat een meerniveau analyse zinvol is, als de intraklasse correlatie coëfficiënt .04 of meer bedraagt (zie bijvoorbeeld Schönrock-Adema, 2002). Volgens Linda Muthén (Mplus Discussion, 1999) gaat het echter feitelijk niet om de intraklasse correlatie maar om het *design effect* (het effect van de wijze waarop de steekproef is samengesteld: zie ook Snijders & Bosker, 1999, p. 22; Hox & Maas, 2001; Hox, 2002). Voor een steekproef met twee niveau's geldt, bij een gemiddelde groepsgrootte n :

$$\text{design effect} = 1 + (n-1)\rho_1$$

Wanneer het design effect groter is dan 2 betekent dit dat bij analyses rekening gehouden moet worden met het effect van clustering. Deze vuistregel wordt gegeven door Linda Muthén (Mplus Discussion, 1999) op basis van een simulatiestudie van Muthén en Satorra (1995). In het design effect speelt behalve de intraklasse correlatie coëfficiënt ρ_1 ook de groepsgrootte n een rol. Het design effect neemt dus toe als de intraklasse correlatie toeneemt en ook als de groepsgrootte toeneemt.

Muthén (1997) wijst erop dat de *maximum likelihood* schattingsmethode (zie bijlage 2B) in meerniveau analyse bij lage intraklasse correlaties en kleine groepsgroottes problemen zou kunnen opleveren, maar in het algemeen werkt de methode goed als er tenminste 50 groepen zijn.

3.5.5 Berekenen van latente variabele scores; aggregeren van variabelen

In dit onderzoek is meerniveau confirmatieve factoranalyse toegepast om na te gaan in hoeverre meetmodellen voor de handelingen van schoolleiders, schoolcultuur en schoolorganisatie passen op de gegevens die werden verzameld bij docenten. Voor de passing van meetmodellen voor schoolleidersgedrag op schoolleidersdata en voor de schoolbeleving van leerlingen is éénniveau confirmatieve factoranalyse gebruikt. Bij het uitvoeren van de meerniveau confirmatieve factoranalyses bleek dat met behulp van LISREL (versie 8.8) geen factorscores (zie bijlage 2G) konden worden berekend voor meerniveau modellen (John von Briesen & Gerhard Mels, SSI, persoonlijke communicatie, oktober-november 2006). Mels gaf aan dat de methoden om factor scores te produceren voor meerniveau modellen nog niet goed zouden zijn getest en beschreven (persoonlijke communicatie, november 2006).

Aan de hand van de (bijgestelde) meetmodellen werden 'latente variabele' scores bepaald op basis van individuele docenten- en leerlingengegevens (Mels, 2004, p. 22). De uitkomsten werden geaggregeerd, om scores op schoolniveau te verkrijgen (zie hoofdstuk 4). Vervolgens werd de betrouwbaarheid van de geaggregeerde variabelen berekend met de volgende formule (Snijders & Bosker, 1999, hoofdstuk 3):

$$\lambda_j = n_j\rho_1 / (1 + [n_j - 1]\rho_1)$$

λ_j is de betrouwbaarheid van de geaggregeerde variabele voor een groep met groepsgrootte n_j bij een intraklasse correlatie ρ_1 . In deze formule kan n_j worden vervangen door het harmonisch gemiddelde n' , om de betrouwbaarheid voor alle groepen te schatten (Maslowski, 2001, p. 43).

3.5.6 Validering van de onderzoeksinstrumenten: multitrait multimethod benadering

Nadat de meetmodellen waren verbeterd konden de betrouwbaarheid en de construct-validiteit van de schalen nader worden bepaald met behulp van de latente variabele scores. Als indicatie voor de betrouwbaarheid van de meetschalen is Cronbach's alpha berekend (zie ook 3.4.1). Om de construct-validiteit van de schalen te kunnen vaststellen, werd voor de handelingen van schoolleiders, de schoolcultuur en schoolorganisatie een *multitrait-multimethod* benadering gehanteerd (Campbell & Fiske, 1959; zie Trochim, 2006). Hiertoe werden correlaties berekend tussen alle schalen (vier per meetmodel). Aan de hand van een matrix van de correlaties kon worden nagegaan of de discriminante en de convergerende validiteit in orde was. De discriminante validiteit is goed als de correlaties tussen verschillende concepten die met dezelfde methode (bijvoorbeeld de vier verschillende schalen in een meetmodel) worden gemeten, laag zijn. Er is sprake van een goede convergente validiteit als de correlaties tussen dezelfde concepten, gemeten met verschillende methoden (bijvoorbeeld de perceptie van het gedrag van schoolleiders door schoolleiders of docenten).

3.5.7 Toetsing van het onderzoeksmodel

In dit onderzoek wordt uitgegaan van de leerprestaties en het doorstroompercentage van leerlingen in de bovenbouw als indicatoren voor de effectiviteit van de school (zie hoofdstuk 5). Gegevens over de leerprestaties en het doorstroompercentage zijn ontleend aan databestanden die door de Inspectie van het Onderwijs zijn gebruikt voor de samenstelling van de opbrengstenkaarten. Deze databestanden zijn opgevraagd bij DANS (Data Archiving and Networked Services, www.dans.knaw.nl). Als maat voor de leerprestaties is uitgegaan van het gemiddelde cijfer voor alle vakken van het centraal schriftelijk eindexamen. Het gemiddelde eindexamencijfer, zoals dat wordt vermeld op de opbrengstenkaarten, is een gewogen gemiddelde, omdat de deelname van leerlingen aan de verschillende vakken varieert. Voor dit onderzoek is het gemiddelde eindexamencijfer voor alle vakken gemiddeld over drie jaar (2004, 2005 en 2006) gebruikt. Daarnaast is het doorstroompercentage van de derde klas naar het diploma gemiddeld over dezelfde drie jaren. Er wordt uitgegaan van een gemiddelde over drie jaar om de invloed van toevallige schommelingen te verminderen. De eindexamencijfers en het doorstroompercentage blijken redelijk stabiel te zijn van het ene jaar op het andere, maar minder stabiel over een langere periode (zie Maslowski, 2001, p. 104). Om het gemiddelde eindexamencijfer en het gemiddelde doorstroompercentage te kunnen berekenen voor de deelnemende scholen, zijn de bestanden met de gegevens van de opbrengstenkaarten gekoppeld aan het databestand met de latente variabele scores, waarbij een matchingsprocedure is uitgevoerd aan de hand van het BRIN-nummer (Basisregistratie Instellingen).

De toetsing van het causale onderzoeksmodel aan de data is uitgevoerd met behulp van *structural equation modeling* (SEM, zie bijlage 2B) in LISREL versie 8.8. Bij het opstellen van causale modellen zijn drie regels van toepassing (ontleend aan Saris & Stronkhorst, 1984, p.35):

1. Common causes of variables which are related by direct causal effects cannot be left out of the theory;
2. Intervening variables can be omitted from the theory without harming the test procedures of causal hypotheses in the theory;
3. Variables which influence either cause or effect variables but not both can be omitted without harming the test procedures of causal hypotheses in the theory.

Vanwege de eerste regel kunnen contextvariabelen die van invloed zijn op de leerprestaties en mogelijk ook op het gedrag van de schoolleider niet worden genegeerd. De derde regel geeft aan dat (context-) variabelen die van invloed zijn op hetzij de handelingen van schoolleiders, hetzij de uitkomstvariabelen, zonder problemen kunnen worden weggelaten. Uit de tweede regel valt af te leiden dat het niet nodig is om intermediaire variabelen aan het model toe te voegen om aan te tonen dat er een effect bestaat van de handelingen van schoolleiders op de uitkomstvariabelen. Dit is opmerkelijk omdat er in het onderzoek naar schoolleiderschap vanuit wordt gegaan dat indirecte effectmodellen (met ‘antecedente’ factoren) een beter resultaat zullen opleveren dan directe effectmodellen (zie bijv. Witziers, Bosker & Krüger, 2003). Als er geen correlatie is tussen twee variabelen, is er geen (direct of indirect) causaal verband, tenzij er sprake is van een zogenaamde *suppressor variabele* (zie Verschuren, 1991, p. 279), die negatief correleert met de ene variabele en positief met de andere. Uit het voorgaande kan worden opgemaakt, dat als er geen significant direct causaal verband aantoonbaar is tussen schoolleiderschap en leerprestaties, ook geen significante indirecte effecten gevonden zullen worden, tenzij er sprake is van positieve en negatieve effecten die elkaar neutraliseren. Een indirect-effect-model biedt de mogelijkheid om de invloed van handelingen van schoolleiders op verschillende variabelen in het onderzoeksmodel te ‘ontrafelen’.

3.6 Discussie

In dit hoofdstuk zijn achtereenvolgens de constructie van de onderzoeksinstrumenten, de dataverzameling, de representativiteit van de steekproef en de data-analyse methoden beschreven. Hieronder volgt een overzicht met een discussie van enkele resultaten.

Voor de constructie van de vragenlijsten voor het onderzoeken van het gedrag van schoolleiders, de schoolcultuur, de schoolorganisatiepraktijk en de schoolbeleving van leerlingen (zie paragraaf 3.2), werd zoveel mogelijk gebruik gemaakt van bestaande onderzoeksinstrumenten. Voor het meten van de handelingen van schoolleiders werd een groot aantal items ontleend aan het instrument *Onderwijskundige Gerichtheid van Schoolleiders* van Krüger (1994). De vragenlijsten voor het meten van de schoolcultuur en de schoolorganisatiepraktijk werden gebaseerd op instrumenten van Houtveen et al. (1995) en Maslowski (2001). Voor de vragenlijst over schoolbeleving van leerlingen werden vooral items ontleend aan de *Student Engagement and Family Educational Culture Survey* van Leithwood et al. (Leithwood, Aitken & Jantzi, 2000). Waar nodig werden de schalen aangevuld met nieuwe items, waarbij gebruik werd gemaakt van diverse andere bronnen. Door zoveel mogelijk uit te gaan van bestaande en geteste instrumenten, werd ‘inhoudsvaliditeit’ nagestreefd (Swanborn, 1994, p.189). Schoolleiders en docenten controleerden de ‘validiteit-op-het-eerste-gezicht’ (face validity) van de items.

Bij de werving van scholen werd een aantal van 100 deelnemende scholen als doel gesteld (zie paragraaf 3.3). Dit aantal werd noodzakelijk geacht voor het uitvoeren van de benodigde analyses. Het onderzoek was gericht op de eindverantwoordelijke schoolleiders (locatiedirecteuren) van vestigingen voor voortgezet onderwijs met een havo-bovenbouw. Door het onderzoek te richten op de havo, werd gestreefd naar een overeenkomstige ‘context’ voor het onderzoeken van de invloed van de schoolleider. De havo kan worden beschouwd als een homogene opleiding met relatief geringe verschillen in aanleg tussen leerlingen. Het was de bedoeling om op deze manier ook schoolleiders te vinden met een vergelijkbaar takenpakket. De taken en verantwoordelijkheden van locatiedirecteuren blijken echter behoorlijk uiteen te lopen. Vanwege de verwachte respons van 20-30% voor dit soort

onderwijsonderzoek werden alle 484 locaties met een havo-bovenbouw uitgenodigd om deel te nemen. Uiteindelijk werden 103 scholen bereid gevonden om mee te werken aan het onderzoek, dit is een respons van 21%. Deze respons valt binnen de verwachting. Scholen die niet mee wilden doen aan, gaven aan dat zij selectief moeten zijn vanwege de vele verzoeken om mee te werken aan onderzoek. Verder kon ook de werkdruk van docenten, of speciale projecten die werden uitgevoerd binnen de school, aanleiding vormen om niet deel te nemen. Voor de dataverzameling werden op elke school vragenlijsten voorgelegd aan de schoolleider, 15 docenten van de havo-bovenbouw en twee havo-5 klassen. In totaal werkten 103 schoolleiders, 998 docenten en 4336 leerlingen mee aan het onderzoek.

De groep van deelnemende scholen is representatief voor de totale populatie wat betreft de verdeling over de provincies en de schoolgrootte (zie paragraaf 3.4). Het percentage rooms-katholieke scholen onder de deelnemende scholen is wat groter dan in de totale populatie. Katholieke scholen waren kennelijk wat vaker bereid om mee te werken aan het onderzoek dan andere bijzondere scholen. De deelnemende schoolleiders waren gemiddeld wat korter in functie dan in een vergelijkbaar onderzoek van Houtveen et al. (6.1 jaar vs. 8.6 jaar, zie Houtveen et al., 1993). Het is mogelijk dat schoolleiders met relatief weinig ervaring op de huidige school meer geïnteresseerd waren om deel te nemen aan dit onderzoek. Het is echter ook mogelijk dat schoolleiders tegenwoordig sneller wisselen van school. Dit is niet nader onderzocht.

In de voorlaatste paragraaf van dit hoofdstuk (3.5) wordt een overzicht gegeven van de methoden die zijn gebruikt voor de data-analyse. De uitkomsten van de data-analyse zijn beschreven in de volgende hoofdstukken. De kwaliteit van de onderzoeksinstrumenten is nader onderzocht met behulp van (meerniveau) confirmatieve factoranalyse. Hierbij is gebruik gemaakt van de multitrait multimethod benadering. De resultaten worden besproken in hoofdstuk 4 van dit proefschrift. In hoofdstuk 5 worden, aan de hand van de variabelen in het onderzoeksmodel, de relaties verkend tussen de kenmerken van scholen en de uitkomsten. De resultaten van de toetsing van het (causale) onderzoeksmodel worden beschreven in hoofdstuk 6. Hoofdstuk 7 geeft een overzicht van het gehele onderzoek en de conclusies.

Hoofdstuk 4

Validering van de onderzoeksinstrumenten

4.1 Inleiding

In dit hoofdstuk wordt verslag gedaan van de validering van de instrumenten die in hoofdstuk 3 zijn beschreven. Het gaat hierbij om het meten van handelingen van schoolleiders, de schoolcultuur, de schoolorganisatiepraktijk en de schoolbeleving van leerlingen. De vragenlijsten voor het bepalen van schoolleidersgedrag, de schoolcultuur en de schoolorganisatiepraktijk zijn gebaseerd op het *competing values framework* (concurrerende waarden model) van Quinn en Rohrbaugh (1983). Elk instrument bestaat uit vier schalen: één voor ieder kwadrant in het model (zie figuur 2.3). De items zijn voor een deel ontleend aan bestaande vragenlijsten en voor een deel nieuw geformuleerd. Van de vragenlijst over de handelingen van schoolleiders bestaan twee versies: één waarmee schoolleiders zelf kunnen aangeven in welke mate zij bepaalde handelingen uitvoeren en één waarmee docenten kunnen aangeven in hoeverre zij deze handelingen van de schoolleider waarnemen. Bij de schoolcultuur en de schoolorganisatiepraktijk gaat het om de beleving van docenten. Voor de vragenlijst over de schoolbeleving van leerlingen is uitgegaan van de *Student Engagement and Family Culture Survey* van Leithwood c.s. (Leithwood & Jantzi, 2000a; Leithwood, Aitken & Jantzi, 2000). Uit dit instrument zijn vier schalen overgenomen, die zijn aangepast en aangevuld.

In dit hoofdstuk staat de tweede onderzoeksvraag centraal:

Is het concurrerende waarden model een geschikt kader om zowel het schoolleidersgedrag, als de schoolcultuur en de schoolorganisatiepraktijk te karakteriseren en relaties tussen deze variabelen bloot te leggen?

Deze onderzoeksvraag wordt beantwoord, nadat eerst de kwaliteit van de onderzoeksinstrumenten is vastgesteld aan de hand van de volgende vraag:

Voldoen de instrumenten voor het meten van schoolleidersgedrag, de schoolcultuur, de schoolorganisatiepraktijk en de schoolbeleving van leerlingen aan de psychometrische criteria voor validiteit en betrouwbaarheid?

Tijdens de constructie van de instrumenten is aandacht besteed aan de *inhoudsvaliditeit*, dat is de mate waarin de vragenlijst representatief is voor datgene wat gemeten moet worden (zie paragraaf 3.2). Bij het concurrerende waarden model gaat het er bijvoorbeeld om, dat de vier kwadranten worden ingevuld met passende items.

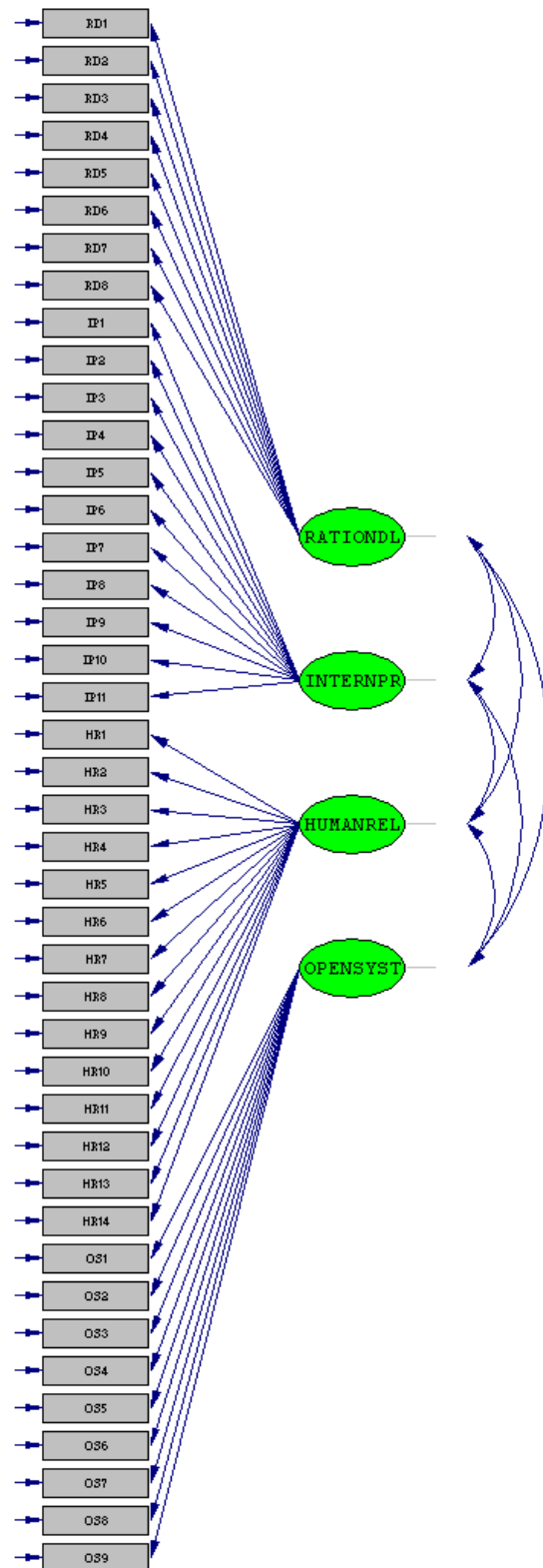
In dit hoofdstuk komen de *betrouwbaarheid* en de *construct- of begripsvaliditeit* van de instrumenten aan de orde (zie 3.5.1). De betrouwbaarheid van de meetinstrumenten kan worden vastgesteld aan de hand van de *interne consistentie*. Constructvaliditeit is onder te verdelen in *convergerende* en *discriminante* validiteit (Swanborn, 1994, p. 191; Garson, z.j., d).

De constructvaliditeit en de betrouwbaarheid van de onderzoeksinstrumenten werden onderzocht door middel van confirmatieve factoranalyse (zie 3.5.2). Daartoe is voor elk onderzoeksinstrument een meetmodel geconstrueerd. In figuur 4.1 is als voorbeeld het meetmodel voor de handelingen van schoolleiders afgebeeld. In dit meetmodel zijn vier factoren opgenomen, die corresponderen met de vier schalen in de vragenlijst (zie bijlage 1A). Iedere factor wordt gemeten door een serie variabelen (de *indicatoren*), waarbij elke variabele slechts aan één factor is gekoppeld. Deze variabelen vertegenwoordigen de items in de vragenlijst. In het meetmodel wordt verondersteld dat alle indicatorvariabelen een meetfout hebben en dat de vier factoren met elkaar correleren.

De mate waarin variabelen ‘laden op’ (voorspeld worden door) de bijbehorende factor zegt iets over de interne consistentie van een schaal. Dit wordt ook wel beschouwd als een vorm van convergerende validiteit. Een andere maat voor de interne consistentie is *Cronbach's alpha*. De discriminante validiteit van het onderzoeksinstrument is in orde als de factoren niet teveel correleren en als de *goodness-of-fit* maten aangeven dat het model goed past bij de data (zie bijlagen 2D-F).

Overigens zou op basis van het concurrerende waarden raamwerk verondersteld kunnen worden, dat de correlaties tussen factoren voor twee naast elkaar gelegen kwadranten hoger zijn dan die tussen de factoren voor twee tegenover elkaar liggende kwadranten (zie figuur 2.3). Dit wordt nader toegelicht in paragraaf 4.2. De correlaties tussen twee kwadranten mogen echter ook weer niet te hoog zijn, omdat het anders niet zeker is dat er inderdaad verschillende aspecten worden gemeten.

In paragraaf 4.3 wordt beschreven hoe het onderzoek naar de betrouwbaarheid en de discriminante validiteit van de vragenlijsten over de handelingen van schoolleiders is uitgevoerd. Bij de confirmatieve factoranalyse is gebruik gemaakt van meerniveau technieken vanwege de hiërarchische structuur van de data (zie 3.5.3 en 3.5.4). De toepassing van meerniveau confirmatieve factoranalyse in LISREL wordt in paragraaf 4.3 geëvalueerd. Aan de hand van de uitkomsten van de confirmatieve factoranalyses, is één nieuw meetmodel geconstrueerd voor zowel de zelfbeoordeling van schoolleiders, als de waarneming van de handelingen van schoolleiders door docenten. De geschiktheid van dit nieuwe meetmodel is weer vastgesteld met behulp van confirmatieve factoranalyse. Het aanpassen van een meetmodel aan de hand van de data kan leiden tot kanskapitalisatie, waardoor het model wel past bij de steekproef, maar niet generaliseerbaar is (zie bijv. MacCallum, Roznowski & Necowitz, 1992; Stevens, 2002). Hier wordt in paragraaf 4.3 ook aandacht aan besteed. In paragraaf 4.4 wordt het onderzoek naar de betrouwbaarheid en de validiteit van de vragenlijsten over de schoolcultuur en de schoolorganisatiepraktijk beschreven. Aan de hand van de uitkomsten zijn de (meet-)modellen voor de schoolcultuur en de schoolorganisatiepraktijk bijgesteld. De betrouwbaarheid en de validiteit van het instrument voor het meten van de schoolbeleving van leerlingen komt aan de orde in paragraaf 4.5. Ook hier is het model weer aangepast op basis van de uitkomsten van de confirmatieve factoranalyse.



Figuur 4.1 Meetmodel voor de validering van de structuur van het onderzoeksinstrument voor schoolleidersgedrag, met vier correlerende factoren, voor rationele doel, interne proces, 'human relations' en open systeem handelingen

Om de constructvaliditeit van de onderzoeksinstrumenten voor de handelingen van schoolleiders, de schoolcultuur en de schoolorganisatiepraktijk nader te bepalen, zijn met behulp van de bijgestelde meetmodellen factorscores (latente variabele scores) berekend met behulp van LISREL en geaggregeerd naar schoolniveau (zie 3.5.5). Vervolgens zijn de correlaties tussen deze scores opgenomen in een *multitrait-multimethod* matrix (Campbell & Fiske, 1959). In de multitrait-multimethod-benadering wordt ervan uitgegaan, dat voor een goede discriminante validiteit de correlaties tussen verschillende concepten die worden gemeten met dezelfde methode, laag moeten zijn, terwijl voor een goede convergerende validiteit de correlaties tussen dezelfde concepten, gemeten met verschillende methoden, voldoende hoog moeten zijn (zie 3.5.6).

De discriminante validiteit is niet alleen van belang voor het onderscheid tussen de verschillende schalen binnen één onderzoeksinstrument (zie boven), maar ook voor het onderscheid tussen de verschillende concepten die zijn geoperationaliseerd aan de hand van het concurrerende waarden model. Het is enerzijds te verwachten dat er verbanden zijn tussen de scores voor de handelingen van schoolleiders, de schoolcultuur en de schoolorganisatiepraktijk per kwadrant van het concurrerende waarden model, maar anderzijds moet ook vastgesteld kunnen worden dat inderdaad onafhankelijke concepten gemeten worden. Voor de zelfbeoordeling door schoolleiders van de door hen uitgevoerde handelingen en de waarneming van docenten van de handelingen van schoolleiders geldt daarentegen dat er juist sprake moet zijn van convergerende validiteit. In paragraaf 4.6 wordt de multitrait-multimethod matrix geanalyseerd. In paragraaf 4.7 komt de betrouwbaarheid van de geaggregeerde variabelen aan de orde. Daarna worden de verbanden tussen de variabelen nader besproken.

De eerste onderzoeksvraag naar de betrouwbaarheid en de validiteit van de onderzoeksinstrumenten kan worden beantwoord aan de hand van de uitkomsten van de confirmatieve factoranalyses. De multitrait-multimethod matrix levert informatie om de tweede onderzoeksvraag te beantwoorden. Hierbij gaat het erom of met behulp van het concurrerende waarden model zowel schoolleidersgedrag als de schoolcultuur en de schoolorganisatiepraktijk kunnen worden gemeten en of deze metingen voldoende onafhankelijk zijn van elkaar. Paragraaf 4.8 bevat een discussie van de resultaten.

4.2 De validering van onderzoeksinstrumenten volgens het concurrerende waarden model

Bij de ontwikkeling van de onderzoeksinstrumenten voor de handelingen van schoolleiders, de schoolcultuur en de schoolorganisatiepraktijk is, zoals eerder aangegeven, uitgegaan van het concurrerende waarden model van Quinn en Rohrbaugh (1983). Dit raamwerk omvat vier organisatiemodellen, die voor een deel tegengesteld zijn aan elkaar (concurreren), maar elkaar ook aanvullen (zie bijv. Quinn, Faermann, Thompson & McGrath, 2003):

- het *rationele doel* (RD) model, gericht op productiviteit en winst,
- het *intern proces* (IP) model, gericht op stabiliteit en continuïteit,
- het *human relations* (HR) model, gericht op inzet, samenhang en moreel,
- het *open systeem* (OS) model, gericht op aanpassingsvermogen en externe ondersteuning.

De validiteit van instrumenten die zijn gebaseerd op het concurrerende waarden model is door diverse onderzoekers getoetst, in verschillende contexten en met verschillende methoden. Hieronder volgt een overzicht (gedeeltelijk ontleend aan Kalliath, Bluedorn & Gillespie, 1999).

Quinn en Rohrbaugh (1983) stelden het concurrerende waarden model op als raamwerk voor de criteria die worden gebruikt om *organisatie effectiviteit* te beoordelen. Zij maakten daarbij gebruik van *meerdimensionale schaalanalyse*. Bij meerdimensionale schaalanalyse worden de correlaties tussen variabelen vertaald in afstanden binnen een geometrische ruimte (Denison, Hooijberg & Quinn, 1995). Meestal gaat het hierbij om een twee- of driedimensionale ruimte. Oorspronkelijk kende het concurrerende waarden model drie dimensies: intern-extern, flexibiliteit-beheersing en doelen-middelen (Quinn & Rohrbaugh, 1983). Later toonde Quinn (1988, 1998) aan dat twee dimensies voldoende waren om het model te beschrijven: intern-extern en flexibiliteit-beheersing. Deze twee dimensies verdelen het model in vier kwadranten (de vier organisatiemodellen, zie ook paragraaf 2.5). Het rationele doel model en het open systeem model zijn extern gericht, in tegenstelling tot het interne proces model en het ‘human relations’ model, waarbij de aandacht uitgaat naar de interne organisatie. Bij het ‘human relations’ en het open systeem model ligt de nadruk op flexibiliteit, terwijl het rationele doel en interne proces model gericht zijn op beheersing. Bij de validering van instrumenten die zijn gebaseerd op het concurrerende waarden model wordt vaak gebruik gemaakt van de (veronderstelde) relaties tussen de vier kwadranten. Dit wordt hieronder toegelicht.

Quinn en Spreitzer (1991; zie ook Kalliath, Bluedorn & Gillespie, 1999) onderzochten de psychometrische eigenschappen van een instrument om *organisatiecultuur* te meten, dat was gebaseerd op het concurrerende waarden model. Zij gebruikten meerdimensionale schaalanalyse en daarnaast de multitrait-multimethod benadering. Met behulp van meerdimensionale schaalanalyse kon het kwadrantenmodel worden bevestigd. In de multitrait-multimethod analyse werden de uitkomsten voor verschillende typen vraagstelling vergeleken. Hierbij bleek dat schaalscores binnen hetzelfde kwadrant behoorlijk met elkaar correleerden (convergerende validiteit) terwijl de schaalscores van verschillende kwadranten juist minder met elkaar correleerden (discriminante validiteit). Denison, Hooijberg en Quinn (1995) gebruikten meerdimensionale schaalanalyse om het concurrerende waarden model voor *leiderschapsrollen* (zie Quinn, 1998) te valideren. Zij constateerden dat effectievere leiders een grotere variëteit aan leiderschapsrollen vervullen. Daarnaast bleken de resultaten het kwadrantenmodel te ondersteunen.

Voortbouwend op het werk van Quinn en Spreitzer (1991), kozen Kalliath, Bluedorn & Gillespie (1999) voor *structural equation modeling* (SEM) als techniek om de validiteit van het concurrerende waarden model voor het meten van de *organisatiecultuur* vast te stellen. In dit onderzoek werd als uitgangspunt gehanteerd dat factoren die naast elkaar liggende kwadranten van het concurrerende waarden model vertegenwoordigen een gematigd positieve correlatie zouden moeten vertonen, omdat deze factoren een van de onderliggende dimensies delen. Dit geldt voor de volgende combinaties van kwadranten: rationele doel en interne proces, rationele doel en open systeem, interne proces en ‘human relations’, ‘human relations’ en open systeem (zie figuur 2.3). Factoren die tegenover elkaar liggende kwadranten representeren zouden daarentegen niet of nauwelijks met elkaar moeten correleren. Hierbij gaat het om twee combinaties: rationele doel en ‘human relations’, interne proces en open systeem.

Voor het onderzoek van Kalliath et al. werden data verzameld onder leidinggevend in een groot ziekenhuis. De uitkomsten van het onderzoek bevestigden de hypothesen, op één na: er werd een hoge correlatie gevonden tussen de interne proces oriëntatie en de open systeem oriëntatie van organisaties. Dit werd verklaard door het feit dat de gezondheidszorg in de Verenigde Staten ten tijde van het onderzoek heftig in beweging was, waardoor er veel aandacht was voor noodzakelijke veranderingen, met als doel om uiteindelijk weer stabiliteit in de interne organisatie te bereiken.

Hooijberg en Choi (2000, 2001) gebruikten eveneens structural equation modeling om de validiteit van het concurrerende waarden model met acht *leiderschapsrollen*, zoals gedefinieerd door Quinn (1998), te onderzoeken bij verschillende groepen waarnemers en in zowel publieke als private organisaties. Zij constateerden dat er voor alle groepen waarnemers (managers, hun collega's, ondergeschikten en superieuren) en voor beide typen organisaties hoge correlaties bestonden tussen twee van de rollen: de rationele doel en de interne proces rol. Zij combineerden deze rollen in één nieuwe rol: doelgerichtheid. Verder toonde het onderzoek van Hooijberg en Choi (2000, 2001) aan dat er verschillen bestonden tussen zowel de groepen waarnemers, als beide typen organisaties, in de mate waarin bepaalde leiderschapsrollen werden geassocieerd met effectiviteit.

Als vervolg op het werk van Denison, Hooijberg en Quinn (1995) is er door Lawrence, Quinn en Lenk (2003) een nieuw instrument ontwikkeld voor het meten van *leiderschapsgedrag*. De validiteit van dit instrument werd vastgesteld met behulp van structural equation modeling. De auteurs gaan uit van *behavioral complexity*, dat wil zeggen dat (goede) managers handelingen in alle vier kwadranten van het concurrerende waarden model zouden moeten uitvoeren. Lawrence et al. vonden een behoorlijke correlatie tussen de naast elkaar liggende kwadranten en lagere, maar nog wel significante, correlaties tussen tegenover elkaar liggende kwadranten. De resultaten ondersteunen het concept van behavioral complexity.

In een ander recent onderzoek werd de validiteit van het concurrerende waarden model getest voor het karakteriseren van de *organisatiecultuur* in Australische organisaties (Lamond, 2003). Hierbij werd opnieuw meerdimensionale schaalanalyse gebruikt. Lamond (2003) gaat ervan uit dat de vier verschillende organisatiemodellen naast elkaar kunnen bestaan in organisaties (zie ook Quinn, 1988). Het gaat er niet om welk model het meest wordt benadrukt in een organisatie, maar om de mate waarin iedere benadering aanwezig is. Uit de uitkomsten van dit onderzoek onder een grote groep managers van verschillende organisaties, blijkt dat alle kwadranten van het concurrerende waarden model matig tot redelijk correleren. Verschillen in de correlaties tussen de kwadranten worden in het artikel niet besproken. Bij nadere beschouwing van de uitkomsten blijkt dat de correlatie tussen de scores voor het rationele doel model en het 'human relations' model (tegenover elkaar liggende kwadranten) hoger is dan de correlatie tussen interne proces model en het 'human relations' model (naast elkaar liggende kwadranten). De andere naast elkaar gelegen kwadranten correleren redelijk. De correlatie tussen de scores voor het interne proces model en het open systeem model (tegenover elkaar liggende kwadranten) is lager dan de overige correlaties. Deze resultaten stemmen slechts gedeeltelijk overeen met de resultaten van Kalliath et al. en Quinn et al., maar ze bevestigen het uitgangspunt van Lamond dat de vier organisatiemodellen naast elkaar kunnen voorkomen binnen een bepaalde organisatie. Lamond (2003) veronderstelt dat de combinatie van de vier schaalcores een indicatie geeft van de totale sterkte van de organisatiecultuur. Volgens de auteur moet er rekening mee worden gehouden dat de beoordeling van de sterkte van de organisatiecultuur beïnvloed kan worden door de specifieke rollen die de respondenten vervullen in de organisatie (Lamond, 2003). Respondenten die zich

bezig houden met human resource management zouden bijvoorbeeld vooral de interne gerichtheid kunnen waarnemen, terwijl marketing managers de externe gerichtheid beter zien.

Uit het bovenstaande kan worden opgemaakt dat de relaties tussen de vier kwadranten (organisatiemodellen) in het concurrerende waarden raamwerk, die op basis van de onderliggende dimensies verondersteld zouden kunnen worden, niet altijd ook daadwerkelijk worden gevonden bij de validering van de onderzoeksinstrumenten. De afwijkende uitkomsten kunnen veroorzaakt zijn door de verschillende rollen en taken die bepaalde groepen respondenten in hun organisatie vervullen, waardoor zij een specifieke kijk op de organisatie hebben, of door (interne of externe) ontwikkelingen waardoor de nadruk op de verschillende organisatiemodellen kan variëren met de tijd (zie ook paragraaf 2.5).

4.3 Betrouwbaarheid en validiteit van de vragenlijsten over schoolleidersgedrag

Er zijn twee versies geconstrueerd van de vragenlijst over de handelingen van schoolleiders: één voor schoolleiders en één voor docenten (zie bijlagen 1A en 1B). Beide versies bevatten dezelfde items, alleen de formulering verschilt voor de twee respondentgroepen. De vragenlijst bestaat uit 42 items, verdeeld over vier schalen die corresponderen met de vier modellen in het concurrerende waarden model. Met behulp van een vierpunts Likertschaal kan de schoolleider aangeven in welke mate een bepaalde handeling wordt uitgevoerd en de docent in welke mate de handeling wordt waargenomen. De antwoordmogelijkheden lopen van ‘nooit’ (1) tot ‘heel vaak’ (4).

De betrouwbaarheid en de validiteit van de onderzoeksinstrumenten werd onderzocht door middel van éénniveau en meerniveau confirmatieve factoranalyse. Daartoe is een meetmodel geconstrueerd met vier factoren die corresponderen met de vier organisatiemodellen in het concurrerende waarden model (zie figuur 4.1). Naast het meetmodel voor het complete instrument werden ook meetmodellen gespecificeerd voor alle schalen afzonderlijk. Deze meetmodellen bevatten telkens één factor met de bijbehorende indicatoren.

In deze paragraaf wordt eerst de *overall fit* oftewel de ‘modelpassing’ besproken van de meetmodellen, zowel voor de schoolleidersgegevens, als voor de docentgegevens. De overall fit geeft aan in welke mate een meetmodel als geheel past bij de data (zie bijlagen 2D-F). De uitkomsten geven een indicatie van de validiteit en de betrouwbaarheid van het onderzoeksinstrument. Daarna worden de parameterschattingen voor alle meetmodellen (compleet en per schaal), respondentengroepen (schoolleiders en docenten) en analysemethoden (éénniveau en meerniveau) vergeleken. Hieruit is af te leiden welke items het meest geschikt zijn als indicator voor een bepaalde factor. Vervolgens wordt beschreven hoe het meetmodel voor de handelingen van schoolleiders is bijgesteld en in hoeverre dit heeft geleid tot een verbetering van de passingswaarden en de geschatte parameters. De paragraaf wordt afgesloten met een technische beschouwing van de uitvoering van de confirmatieve factoranalyses in dit onderzoek.

4.3.1 Passing van de meetmodellen voor schoolleidersgedrag op schoolleidersdata

De vragenlijst over de handelingen van schoolleiders is ingevuld door 103 schoolleiders. Er is een confirmatieve factoranalyse uitgevoerd met behulp van het programma LISREL (versie 8.8) met ruwe data als input, zoals beschreven door Mels (2004). LISREL maakt gebruik van een zogenaemde *maximum likelihood* schatting van de modelparameters (zie bijlage 2B).

Enkele ontbrekende gegevens, in het technische jargon aangeduid als *missing values* zijn in de datafile vervangen door een gemiddelde. In tabel 4.1 staat een overzicht van passingswaarden die zijn verkregen bij de confirmatieve factoranalyse, voor het totale meetmodel en voor de vier schalen afzonderlijk, aan de hand van de schoolleidersdata. Een afbeelding van het volledige meetmodel met de parameterschattingen wordt weergegeven in bijlage 3A.

Tabel 4.1 *Passingsmaten en Cronbach's alpha voor het volledige model van schoolleidersgedrag en afzonderlijke schalen, bij een confirmatieve factoranalyse met schoolleidersdata (n = 103)*

Meetmodel	aantal items	α	X^2 (df)	RMSEA	SRMR	CFI	AGFI
Volledig model	42	.81	1198.29 (813)	.068	.112	.709	.602
Rationele doel	8	.63	48.25 (20)	.118	.094	.771	.810
Interne proces	11	.69	90.09 (44)	.101	.093	.829	.792
Human relations	14	.70	139.83 (77)	.089	.102	.746	.777
Open systeem	9	.77	26.66 (27)	.059	.063	.962	.877

RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean square Residual; CFI = Comparative Fit Index; AGFI = Adjusted Goodness-of-Fit Index

In tabel 4.1 zijn, in aanvulling op de chi-kwadraat toets, de volgende passingsmaten opgenomen: Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean square Residual (SRMR), Comparative Fit Index (CFI) en Adjusted Goodness-of-Fit Index (AGFI). Een toelichting op deze passingsmaten is te vinden in bijlage 2F. Verder is in tabel 4.1 ook Cronbach's alpha (α) opgenomen, ondanks de bezwaren tegen het gebruik van deze coëfficiënt als maatstaf voor betrouwbaarheid. Cronbach's alpha is gevoelig voor het aantal items in het instrument en het onderschat de betrouwbaarheid enigszins, maar geeft wel op een eenvoudige manier een eerste indruk van de interne consistentie van een meetinstrument (zie ook 3.5.1). Voor Cronbach's alpha geldt dat deze tenminste .70 moet zijn om een schaal te beschouwen als een redelijk meetinstrument. Als alpha groter is dan .80, is er sprake van een goede schaal. In exploratief onderzoek wordt een alpha die groter is dan .60 ook nog geaccepteerd (Garson, z.j., c). Uit tabel 4.1 is op te maken dat de interne consistentie, zoals bepaald met Cronbach's alpha, van het totale meetinstrument redelijk is. Voor de rationele doel schaal is de waarde van alpha matig, voor de andere schalen redelijk tot goed.

Chi-kwadraat is voor alle meetmodellen in tabel 4.1 significant. Dit heeft hier echter weinig betekenis omdat chi-kwadraat onder meer gevoelig is voor de steekproefomvang en de complexiteit van het model (zie bijlage 2E). De waarde van χ^2/df is wel steeds kleiner dan 2 of 3 en dit zou volgens Verschuren (1991, p. 470) wijzen op een redelijke modelpassing. Voor RMSEA geldt dat de waarde kleiner moet zijn dan .06-.08, SRMR moet kleiner zijn dan .08 en CFI groter dan of gelijk aan .95 voor een goede modelpassing (zie bijlage 2F). Uit tabel 4.1 blijkt dat de modelpassing voor het volledige model redelijk is (gebaseerd op RMSEA) en voor de rationele doel, interne proces en 'human relations' schalen matig. Alleen voor de open systeem schaal voldoen de belangrijkste passingsmaten (RMSEA, SRMR en CFI) aan de criteria voor een goede fit. De (Adjusted) Goodness-of-Fit Index (-A-GFI) neemt in LISREL output een prominente plaats in. Deze fit index wordt echter tegenwoordig als minder betrouwbaar beschouwd (zie bijlage 2F). Voor (A)GFI geldt dat de waarde groter zou moeten zijn dan, of gelijk aan .95 en aan die maatstaf voldoet dus geen van de modellen aangeduid in tabel 4.1.

Aangezien de meetmodellen in tabel 4.1 niet (helemaal) voldoen aan de criteria voor een goede fit, is het wenselijk om na te gaan of er wellicht nog verbetering mogelijk is van de

interne consistentie en de validiteit van het meetinstrument door veranderingen aan te brengen. Voordat de parameterschattingen voor de verschillende meetmodellen nader worden geanalyseerd, wordt echter eerst onderzocht hoe de meetmodellen passen op de data van docenten

4.3.2 Passing van de meetmodellen voor schoolleidersgedrag op docentenwaarnemingen

In totaal hebben 997 docenten van 100 scholen de vragenlijst over de handelingen van schoolleiders beantwoord. Omdat er sprake is van een hiërarchische structuur van de data zou het gebruik van meerniveau analysetechnieken voordelen kunnen bieden. Om een indruk te krijgen van de passing van het model op docentenwaarnemingen, is eerst een éénniveau confirmatieve factoranalyse uitgevoerd, zoals aanbevolen door Muthén (1994). Bij de analyse zijn ruwe data als input gebruikt (Mels, 2004). Cases met meer dan vijf ontbrekende gegevens (*missing values*) werden vooraf verwijderd, zodat er 896 cases overbleven. De overige ontbrekende waarden werden vervangen door een gemiddelde. Tabel 4.2 geeft een overzicht van de passingswaarden die zijn verkregen bij een éénniveau confirmatieve factoranalyse voor het totale meetmodel en voor de vier schalen afzonderlijk, aan de hand van docentenwaarnemingen.

Tabel 4.2 *Passingsmaten en Cronbach's alpha voor het volledige model van schoolleidersgedrag en afzonderlijke schalen, bij een confirmatieve factoranalyse met docentendata (n = 896)*

Meetmodel	aantal items	α	X^2 (df)	RMSEA	SRMR	CFI	AGFI
Volledig model	42	.94	4660.42 (813)	.073	.065	.950	.779
Rationele doel	8	.84	85.97 (20)	.061	.033	.983	.958
Interne proces	11	.85	628.90 (44)	.122	.068	.906	.830
Human relations	14	.88	723.76 (77)	.097	.059	.938	.859
Open systeem	9	.76	131.36 (27)	.066	.042	.956	.947

RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean square Residual; CFI = Comparative Fit Index; AGFI = Adjusted Goodness-of-Fit Index

De waarden voor Cronbach's alpha (α) die vermeld staan in tabel 4.2, geven aan dat het meetinstrument als geheel en de vier schalen apart een goede interne consistentie vertonen. Voor het totale meetinstrument geldt dat SRMR en CFI voldoen aan de criteria voor een goede fit: SRMR is kleiner dan .08 en CFI is groter dan of gelijk aan .95. RMSEA is weliswaar groter dan .06, maar voldoet nog wel aan het ruimere criterium van .08 als bovengrens voor een redelijk goede modelpassing (Browne & Cudeck, 1993, in du Toit & du Toit, 2001, p. 398). De rationele doel en open systeem schalen vertonen een goede fit. Voor deze schalen is ook de waarde van AGFI voldoende (groter dan, of gelijk aan .95). De interne proces en 'human relations' schalen hebben een iets minder goede passing, hierbij voldoet alleen SRMR aan het criterium.

Voor de data gebaseerd op docentwaarnemingen lijkt de modelpassing van de rationele doel model schaal beter dan voor de schoolleidersdata. De modelpassing van de open systeem schaal is voor beide datasets goed en de interne proces en 'human relations' schaal vertonen zowel bij de schoolleidersdata als bij de docentendata een matige fit. Het volledige meetmodel heeft een redelijke modelpassing bij beide datasets.

Om na te gaan of het nodig is om rekening te houden met de effecten van clustering in de docentendata binnen scholen, is de *intraklasse correlatie coëfficiënt* (ρ_1) berekend voor alle

items in de vragenlijst over de handelingen van schoolleiders (zie 3.5.4). Daarbij werd uitgegaan van een gemiddelde groeps grootte van 10 docenten. Bij een (conservatieve) schatting met behulp van ANOVA bleek dat de intraklasse correlatie voor de afzonderlijke items uiteenloopt van 0.06 tot 0.30, met een gemiddelde van 0.17. Dit komt overeen met een *design effect* ($= 1+(n-1)\rho_i$, zie 3.5.4) van respectievelijk 1.5 voor de laagste waarde van de intraklasse correlatie tot 3.7 voor de hoogste waarde en 2.5 voor de gemiddelde waarde. Aangezien de intraklasse correlatie groter is dan 0.04 en het design effect groter is dan 2, moet dus daadwerkelijk rekening worden gehouden met een effect van de clustering van docenten in scholen (zie 3.5.4). Dit betekent dat het zinvol is om meerniveau analyse technieken toe te passen.

Aan de hand van de beperkte informatie die beschikbaar was over de aanpak van meerniveau confirmatieve factoranalyse met behulp van LISREL (zowel in de handleidingen als in de onderzoeksliteratuur), is ervoor gekozen twee verschillende benaderingen toe te passen en de resultaten te vergelijken. De eerste benadering geeft beperkte informatie over de passing van het meetmodel, maar zou goede parameterschattingen moeten opleveren, de tweede benadering levert meer informatie over de modelpassing, maar hierbij wordt een minder geavanceerde schattingsmethode toegepast.

Bij de eerste benadering werd gebruik gemaakt van de *Full Information Maximum Likelihood* (FIML) methode (zie du Toit & du Toit, 2001, p. 388). Bij de tweede benadering werd *Muthén's approximate Maximum Likelihood solution* (MUML) toegepast (Muthén, 1990). MUML gaat uit van één *between groups* (tussen groepen = schoolniveau) en één *within groups* (binnen groepen = docentenniveau) model (*Mplus discussion* 31-08-2002, zie ook Heck & Thomas, 2000). FIML analyseert daarentegen, naast één *within group*, D *between groups*, waarbij D gelijk is aan het aantal verschillende groeps groottes (Muthén, 1990; zie ook Hox & Maas, 2001). Hoewel de FIML methode geavanceerder is, concluderen Hox & Maas (2001) op basis van simulatiestudies dat de uitkomsten van meerniveau factoranalyse op basis van één *within groep* en één *between groep*, ook goede resultaten oplevert, zolang het aantal groepen groot genoeg is en de groeps grootte niet teveel uiteenloopt.

De meerniveau confirmatieve factoranalyse volgens de eerste benadering (met FIML) is uitgevoerd met behulp van LISREL (versie 8.8), zoals beschreven door du Toit en du Toit (2001; zie ook Mels, 2004). Voor ontbrekende gegevens werd 'paarsgewijze verwijdering' (*pairwise deletion*) toegepast: dit betekent dat correlaties of covarianties tussen twee variabelen alleen worden berekend voor de cases die geldige waarden hebben voor deze variabelen.

In tabel 4.3 staat een overzicht van de waarden voor de passingswaarden die zijn verkregen bij de meerniveau confirmatieve factoranalyse voor het volledige meetinstrument en voor de vier schalen afzonderlijk, aan de hand van docentendata (997 docenten van 100 scholen).

In de tabel wordt naast de waarde voor RMSEA ook het 90 % betrouwbaarheids-interval gegeven. De passing van het model is goed als de bovengrens van het betrouwbaarheidsinterval kleiner is dan .08 (zie ook Sleegers & van Schooten, 2006, p.29). De uitkomsten in tabel 4.3 geven aan dat bij de meerniveau confirmatieve factoranalyse de modelpassing voor het volledige meetinstrument goed is. Voor de rationale doel en open systeem schaal is de modelpassing ook redelijk goed, voor de 'human relations' schaal matig en voor de interne proces schaal slecht. De uitkomsten voor de afzonderlijke schalen stemmen overeen met die van de éénniveau analyse (zie tabel 4.2).

Tabel 4.3 Passingsmaten voor het volledige model van schoolleidersgedrag en afzonderlijke schalen, bij een meerniveau confirmatieve factoranalyse met Full Information Maximum Likelihood (FIML), voor docentendata, ($N = 100$, $n = 997$)

Meetmodel	χ^2 (df)	RMSEA	90% C.I. RMSEA
Volledig model	4067.43 (1626)	.055	.053 - .057
Rationele doel	128.09 (40)	.067	.054 - .080
Interne proces	696.94 (88)	.118	.100 - .126
Human relations	686.84 (154)	.083	.077 - .090
Open systeem	150.12 (54)	.060	.049 - .071

RMSEA = Root Mean Square Error of Approximation; 90% C.I. = 90% betrouwbaarheidsinterval

Bij het gebruik van de Full Information Maximum Likelihood (FIML) methode in LISREL worden geen afzonderlijke passingswaarden gegenereerd voor het meetmodel op schoolniveau en op docentenniveau. Wanneer Muthén's approximate Maximum Likelihood solution (MUML) wordt toegepast in LISREL worden ook passingswaarden geschat per niveau. Deze benadering zou voordelen kunnen bieden voor het optimaliseren van de passing van het meetmodel op schoolniveau en het berekenen van factorscores voor schoolkenmerken (zie bijv. De Maeyer & Rymenans, 2004; Heck & Thomas, 2000; Hox, 2002; Toland & De Ayala, 2005).

Om inzicht te krijgen in de passingswaarden op schoolniveau (between groups) en op docentenniveau (within groups) is een meerniveau confirmatieve factoranalyse uitgevoerd met behulp van LISREL met de schattingsmethode MUML, voor het volledige meetmodel en voor de afzonderlijke schalen, aan de hand van docentendata. De uitkomsten zijn vermeld in tabel 4.4.

Tabel 4.4 Passingsmaten voor het volledige model van schoolleidersgedrag en afzonderlijke schalen, bij een meerniveau confirmatieve factoranalyse met Muthén's approximate Maximum Likelihood solution (MUML), voor docentendata ($N = 100$, $n = 997$)

Meetmodel	χ^2 (df)	RMSEA	CFI	Niveau	RMR	SRMR	GFI
Volledig model	3796.76 (1626)	.052	.950	Between	.020	.201	.280
				Within	.023	.049	.860
Rationele doel	113.70 (40)	.061	.979	Between	.008	.098	.701
				Within	.017	.036	.975
Interne proces	583.84 (88)	.104	.898	Between	.011	.107	.555
				Within	.029	.066	.911
Human relations	576.26 (154)	.074	.953	Between	.013	.127	.519
				Within	.025	.049	.929
Open systeem	123.23 (54)	.051	.967	Between	.011	.123	.747
				Within	.017	.037	.976

RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; (S)RMR = (Standardized) Root Mean square Residual; GFI = Goodness-of-Fit Index

Voor het *between groups* (schoolniveau) en *within groups* (docenten-)niveau kunnen met behulp van LISREL alleen RMR, SRMR en GFI waarden geproduceerd worden (zie tabel 4.4). In het algemeen wordt volgens Maslowski (2001) voor RMR als criterium aangehouden dat waarden kleiner dan .05 wijzen op een goede fit, terwijl waarden kleiner dan .10 acceptabel zijn. SRMR moet kleiner zijn dan .08. GFI is een zwakke maat, maar waarden boven .95 zouden wijzen op een goede fit. Uit tabel 4.4 is op te maken dat voor het volledige meetmodel de modelpassing op schoolniveau slecht is, hoewel RMR wel voldoet aan het

criterium. De modelpassing op docentenniveau voor dit model is wel goed genoeg. Voor de afzonderlijke schalen is de modelpassing op schoolniveau beter dan bij het volledige model, maar niet voldoende, terwijl de modelpassing op docentenniveau ook hier goed is.

Bij een vergelijking van de tabellen 4.3 en 4.4 blijkt dat de waarden voor de totale modelpassing (RMSEA) die zijn gevonden met de MUML methode enigszins afwijken van de waarden die werden gevonden met FIML. De waarden vertonen echter wel hetzelfde patroon. Ook bij de MUML benadering geldt dat de totale modelpassing voor het volledige meetinstrument en voor de rationele doel en de open systeem schaal beter is dan die voor de 'human relations' en de interne proces schaal. Dit patroon is eveneens waarneembaar in de uitkomsten van de éénniveau analyse voor docentendata (zie tabel 4.2).

4.3.3 Parameterschattingen

Om de betrouwbaarheid van de vragenlijst voor schoolleidersgedrag nader te bepalen, zijn de parameterschattingen voor de meetmodellen onderzocht. De factorladingen die worden berekend bij een confirmatieve factoranalyse geven aan in hoeverre de indicator variabelen laden op de (latente) factoren. Een meetinstrument is intern consistent als de indicatoren voor verschillende factoren in voldoende mate en eenduidig laden op de bijbehorende factoren. De factorlading is significant als de kritieke waarde (= de ongestandaardiseerde factorlading gedeeld door de standaardfout, zie bijlage 2D) groter is dan 1.96. Als minimum factorlading voor een indicator wordt meestal .3 of .35 genomen. In het algemeen worden factorladingen kleiner dan .4 beschouwd als onvoldoende en factorladingen groter dan .7 als goed (Garson, z.j., a).

In tabel 4.5 staat een overzicht van de factorladingen van alle items op de bijbehorende factoren, zoals die geschat zijn voor verschillende meetmodellen, aan de hand van schoolleidersdata en docentendata, met éénniveau en meerniveau confirmatieve factoranalyse. De factorladingen zijn afkomstig uit de volledig gestandaardiseerde oplossing (*completely standardized solution*). Hierbij zijn zowel de geobserveerde als de latente variabelen gestandaardiseerd (Jöreskog, 1999). Voor de meerniveau modellen gaat het om een *within group completely standardized solution*. Dit betekent dat de geobserveerde en latente variabelen zijn gestandaardiseerd binnen elke groep.

In de eerste kolom van tabel 4.5 worden de items uit de vragenlijst over de handelingen van schoolleiders (zie bijlage 1A) aangeduid met een code die aangeeft bij welke schaal de items horen: 'rd' staat voor het rationele doel model, 'ip' voor het interne proces model, 'hr' voor het 'human relations' model en 'os' voor het open systeem model. In de tweede en de derde kolom van tabel 4.5 staan de factorladingen die zijn geschat met behulp van éénniveau confirmatieve factoranalyse voor het volledige model van schoolleidersgedrag en voor de afzonderlijke schalen, aan de hand van schoolleidersdata. Voor de meeste items liggen de twee geschatte factorladingen dicht bij elkaar. De uitkomsten zijn dus redelijk consistent. Verschillen tussen de twee schattingen kunnen worden veroorzaakt doordat bij het volledige model correlatie tussen de latente factoren mogelijk is (zie figuur 4.1). Eerder werd al duidelijk dat de passing van de meetmodellen bij de schoolleidersdata matig is (zie tabel 4.1). Uit de factorladingen in tabel 4.5 blijkt nu dat een flink aantal items niet significant of zwak (minder dan .35) laadt op de bijbehorende factoren. Items die zwak laden in het volledige model, doen dat ook in de afzonderlijke schalen.

Tabel 4.5 Factorladingen (volledig gestandaardiseerd) voor het volledige model van schoolleidersgedrag en afzonderlijke schalen, bij éénniveau confirmatieve factoranalyse voor schoolleidersdata (n=103) en docentendata (n=896) en meerniveau confirmatieve factoranalyse met FIML of MUML, voor docentendata (N=100, n = 997)

Item	Schoolleiders		Docenten				Afzonderlijke Schalen		
	Totaal model	Afzonderlijke Schalen	Totaal model	meerniveau		meerniveau			
	éénniveau	éénniveau	éénniveau	FIML between	FIML within	MUML Between	MUML within	FIML between	FIML within
rd1	.32	.31	.67	.24	.62	.89	.62	.94	.63
rd2	.58	.52	.70	.27	.68	.88	.66	.95	.64
rd3	.43	.48	.58	.48	.56	.71	.55	.76	.54
rd4	.60	.65	.59	-.18	.54	.78	.55	.89	.58
rd5	.53	.52	.71	.01	.69	.89	.68	.90	.66
rd6	.48	.51	.71	.17	.67	.93	.67	.99	.68
rd7	.21	.13	.42	-.61	.43	.33	.44	.28	.44
rd8	.35	.30	.63	.47	.57	.77	.57	.80	.60
ip1	.19	.22	.64	.15	.60	.78	.59	.74	.57
ip2	-.09	-.05	.62	.17	.60	.73	.58	.67	.55
ip3	.61	.72	.71	.58	.61	.89	.61	.90	.66
ip4	.54	.57	.62	.89	.53	.81	.56	.87	.62
ip5	.63	.58	.60	.82	.51	.85	.51	.88	.51
ip6	.62	.50	.56	.98	.50	.76	.50	.76	.48
ip7	.12	.21	.50	.10	.53	.56	.51	.54	.48
ip8	.54	.61	.59	.89	.53	.74	.55	.78	.60
ip9	.41	.40	.58	.60	.51	.72	.51	.69	.50
ip10	.19	.32	.46	.49	.41	.71	.40	.70	.38
ip11	.51	.38	.57	.66	.56	.81	.53	.76	.48
hr1	.51	.49	.64	.82	.58	.87	.59	.97	.59
hr2	.75	.82	.70	.61	.64	.89	.65	.91	.67
hr3	.67	.58	.60	.86	.54	.83	.54	.91	.53
hr4	.13	.08	.38	.27	.36	.49	.34	.34	.34
hr5	.60	.65	.60	.77	.56	.84	.56	.87	.56
hr6	.20	.22	.69	.05	.67	.77	.65	.70	.67
hr7	-.03	-.01	.67	.06	.67	.72	.66	.65	.68
hr8	.50	.47	.55	.70	.48	.82	.49	.86	.49
hr9	.32	.30	.51	.04	.52	.53	.51	.31	.63
hr10	.35	.35	.65	.26	.60	.76	.60	.62	.65
hr11	.10	.11	.60	.28	.57	.74	.56	.63	.58
hr12	.38	.40	.48	.60	.46	.64	.46	.57	.47
hr13	.12	.09	.54	-.04	.55	.47	.55	.12	.60
hr14	.15	.10	.48	.01	.49	.45	.49	.13	.53
os1	.80	.79	.64	.94	.56	.87	.59	.93	.64
os2	.29	.27	.44	.51	.38	.35	.42	.30	.47
os3	.61	.64	.50	.93	.50	.63	.50	.67	.50
os4	.22	.22	.38	.04	.41	.47	.39	.47	.34
os5	.60	.58	.53	.19	.51	.58	.49	.48	.44
os6	.70	.73	.66	.69	.60	.88	.61	.92	.62
os7	.47	.49	.42	.95	.43	.64	.40	.69	.37
os8	.34	.37	.53	.42	.50	.63	.50	.60	.52
os9	.49	.48	.52	.32	.49	.52	.49	.49	.49

Schuingedrukte factorladingen zijn niet significant: kritische ratio < 1.96;

Factorladingen >=.35 zijn vetgedrukt;

FIML: Full Information Maximum Likelihood; MUML: Muthén's approximate Maximum Likelihood solution;

'between' = between groups, 'within' = within groups;

'rd' = rationele doel, 'ip' = interne proces, 'hr' = 'human relations' en 'os' = open systeem model.

In de vierde kolom van tabel 4.5 staan de uitkomsten van de éénniveau confirmatieve factoranalyse voor het volledige meetmodel, aan de hand van individuele docentengegevens. Alle factorladingen zijn significant en redelijk hoog. Voor de uitkomsten van de drie benaderingen van meerniveau confirmatieve factoranalyse (vijfde tot en met tiende kolom in tabel 4.5), geldt dat de factorladingen per item op docentenniveau (*within groups*) in de drie series uitkomsten goed overeenkomen. De uitkomsten op docentenniveau liggen dicht bij de uitkomsten van de éénniveau analyse (kolom 4), waarbij is uitgegaan van de gegevens van individuele docenten. Dit ligt voor de hand omdat de intraklasse correlatie voor de docentengegevens relatief klein is: het grootste deel van de variantie ligt op docentenniveau (binnen de scholen).

De uitkomsten op schoolniveau (*between groups*) van de meerniveau analyse met FIML voor het volledige meetmodel (kolom 5 in tabel 4.5), wijken nogal af van die van de andere twee meerniveau analyses (kolommen 7 en 9). Dit zou het gevolg kunnen zijn van het relatief kleine aantal cases op schoolniveau ten opzichte van het aantal te schatten parameters. Hierop wordt nader ingegaan in 4.3.6. De uitkomsten op schoolniveau van de MUML analyse voor het totale meetmodel en de FIML analyse voor de afzonderlijke schalen komen beter overeen in grootte. De kleine verschillen tussen deze parameterschattingen kunnen enerzijds zijn veroorzaakt door de verschillende schattingsmethoden en anderzijds doordat in het volledige meetmodel correlaties tussen de factoren mogelijk zijn. Een afbeelding van het volledige meerniveau meetmodel, met MUML parameterschattingen, is opgenomen in bijlage 3B.

Uit tabel 4.5 kan worden afgeleid welke items onvoldoende laden op de bijbehorende factor. Door verwijdering van items met een niet significante of lage factorlading zou de modelpassing voor en de interne consistentie van de meetschalen kunnen verbeteren. Aan een dergelijke *datagedreven* wijziging van het model kleven bezwaren, er kan sprake zijn van kanskapitalisatie. Dit komt aan de orde in de evaluatie van de gebruikte methoden in 4.3.6. In de volgende paragraaf wordt uiteengezet hoe het meetmodel is bijgesteld aan de hand van de hierboven vermelde uitkomsten van de confirmatieve factoranalyse.

4.3.4 Verbetering van het meetmodel van schoolleidersgedrag

Wanneer de resultaten van de éénniveau confirmatieve factoranalyse voor schoolleidersdata en de drie meerniveau analyses voor de docentendata (zie tabel 4.5) nader worden beschouwd, valt op dat er op schoolniveau een vergelijkbaar patroon is van zwak ladende items (factorlading kleiner dan .4 en/of niet significant). De parameterschattingen van de éénniveau confirmatieve factoranalyse voor individuele docentendata leveren minder informatie op. Ervan uitgaande dat de zwakke items de interne consistentie en daarmee de betrouwbaarheid van het meetinstrument (op schoolniveau) ongunstig beïnvloeden, is het meetmodel bijgesteld.

Voor de aanpassing van het meetmodel werd niet alleen gebruik gemaakt van de parameterschattingen en de modificatie indices die LISREL genereert (zie bijlage 2D), maar ook van de uitkomsten van exploratieve factoranalyses (zie bijlage 2A). Veranderingen in het meetmodel werden alleen doorgevoerd wanneer deze inhoudelijk te verantwoorden waren. Hierbij was het concurrerende waarden model steeds het uitgangspunt. Uit de uitkomsten van de exploratieve factoranalyses bleek dat een model met vier factoren gehandhaafd kon worden, zowel voor schoolleidersdata als voor docentendata. De meeste zwak ladende items bleken inhoudelijk enigszins af te wijken van de overige items in de schaal. Andere items bleken op verschillende factoren te laden omdat ze niet eenduidig waren geformuleerd. De

afwijkende en niet eenduidige items werden verwijderd uit het meetmodel (rd7, rd8, ip1, ip7, ip9, hr4, hr6, hr7, hr9, hr11, hr13, hr14, os2, os4, os5). Twee items bleken beter te passen bij andere factoren (ip2 en ip11). Deze items zijn verplaatst. Tenslotte waren er nog twee items die inhoudelijk veel op elkaar leken (ip5 en ip6), waardoor de passing van het model negatief werd beïnvloed. Hier werd het meest eenduidige item gehandhaafd (ip5), terwijl het andere item werd verwijderd. In totaal zijn 16 (van de 42) items uit het model verwijderd.

Het nieuwe meetmodel dat werd geconstrueerd voor de handelingen van schoolleiders bevat 26 indicatoren. Een overzicht van de schalen (factoren) en de bijbehorende items is te vinden in bijlage 4A. De nieuwe rationele doel schaal (7 items) vertegenwoordigt een gerichtheid op productiviteit, waarbij het vooral gaat om het stellen van doelen. Bij de interne proces schaal (5 items) draait het om controle van de gang van zaken binnen de school. De ‘human relations’ schaal (8 items) gaat vooral over de ondersteuning van docenten bij hun werk. In de open systeem schaal (6 items) komt de externe gerichtheid in de vorm van contacten met partijen buiten de school aan de orde.

Voor het nieuwe meetmodel zijn weer met behulp van LISREL (versie 8.8) dezelfde éénniveau en meerniveau confirmatieve factoranalyses uitgevoerd als voor het oorspronkelijke model. In tabel 4.6 staat een overzicht van Cronbach’s alpha (α) en de passingswaarden die zijn verkregen bij een confirmatieve factoranalyse, voor het totale meetmodel en voor de vier schalen afzonderlijk, aan de hand van de schoolleidersdata.

Tabel 4.6 *Passingsmaten en Cronbach’s alpha voor het nieuwe model van schoolleidersgedrag en afzonderlijke schalen, bij confirmatieve factoranalyse met schoolleidersdata (n = 103)*

Nieuw meetmodel	aantal items	α	X^2 (df)	RMSEA	SRMR	CFI	AGFI
Volledig model	26	.72	359.84 (293)	.047	.094	.851	.744
Rationele doel	7	.67	25.86 (14)	.091	.073	.875	.865
Interne proces	5	.70	5.58 (5)	.034	.042	.996	.936
Human relations	8	.76	31.33 (20)	.075	.065	.940	.872
Open systeem	6	.76	5.35 (9)	.000	.033	1.00	.960

RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean square Residual; CFI = Comparative Fit Index; AGFI = Adjusted Goodness-of-Fit Index

Uit de resultaten in tabel 4.6 kan worden opgemaakt dat Cronbach’s alpha voor het complete nieuwe model ($\alpha = .72$) is gedaald vergeleken met het oorspronkelijke model ($\alpha = .81$, zie tabel 4.1). Het nieuwe model bevat echter aanzienlijk minder items en het is bekend dat Cronbach’s alpha gevoelig is voor het aantal items. Voor de rationele doel, interne proces en ‘human relations’ schaal is Cronbach’s alpha iets gestegen, ondanks het verwijderen van items. Voor de open systeem schaal is alpha min of meer gelijk gebleven. Ten opzichte van het oorspronkelijke model (zie tabel 4.1) is een duidelijke verbetering in de passingswaarden te zien, voor zowel het complete nieuwe model als de afzonderlijke schalen. Voor het complete model voldoet RMSEA aan het criterium (kleiner dan .06) van een goede fit. SRMR en CFI voldoen nog niet aan de criteria voor een goede modelpassing (SRMR kleiner dan of gelijk aan .08, CFI groter dan of gelijk aan .95), maar de waarden komen dichterbij dan bij het oorspronkelijke model. De modelpassing van de nieuwe rationele doel schaal is matig, van de ‘human relations’ schaal redelijk en die van de interne proces schaal en de open systeem schaal goed. Het verschil in modelpassing met het oorspronkelijke model is alleen voor het complete model nagegaan: dit verschil is significant ($\Delta\chi^2 = 868$, $\Delta df = 520$, $p < .001$).

Tabel 4.7 geeft een overzicht van de passingswaarden die zijn verkregen bij een éénniveau confirmatieve factoranalyse voor het totale nieuwe meetmodel en voor de vier schalen afzonderlijk, aan de hand van docentenwaarnemingen. In de tabel zijn ook de waarden van Cronbach's alpha (α) vermeld.

Tabel 4.7 *Passingsmaten en Cronbach's alpha voor het nieuwe model van schoolleidersgedrag en afzonderlijke schalen, bij éénniveau confirmatieve factoranalyse met docentendata (n = 896)*

Nieuw meetmodel	aantal items	α	χ^2 (df)	RMSEA	SRMR	CFI	AGFI
Volledig model	26	.91	1140.80 (293)	.057	.054	.967	.893
Rationele doel	7	.85	49.93 (14)	.054	.025	.990	.969
Interne proces	5	.76	3.22 (5)	.000	.010	1.00	.996
Human relations	8	.84	129.66 (20)	.078	.040	.973	.937
Open systeem	6	.73	30.76 (9)	.052	.028	.984	.974

RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean square Residual; CFI = Comparative Fit Index; AGFI = Adjusted Goodness-of-Fit Index

De waarden voor Cronbach's alpha van het complete nieuwe model en de afzonderlijke schalen zijn, zoals is af te lezen uit tabel 4.7, iets lager dan die van het oorspronkelijke model (zie tabel 4.2), met uitzondering van de waarde voor de rationele doel schaal. Bij vermindering van het aantal items in een schaal is een daling van Cronbach's alpha te verwachten. De waarden voor alpha duiden echter nog steeds op een redelijke tot goede betrouwbaarheid van de schalen.

De passingswaarden van het nieuwe model zijn duidelijk beter dan die voor het oorspronkelijke model. Zowel het complete model als de afzonderlijke schalen voldoen aan de criteria voor een goede modelpassing (RMSEA kleiner dan .06, SRMR kleiner dan of gelijk aan .08, CFI groter dan of gelijk aan .95), alleen voor de human relation schaal schaal is RMSEA nog aan de hoge kant. Voor het complete nieuwe model is het verschil in passing met het oorspronkelijke model, op basis van individuele docentendata, significant ($\Delta\chi^2 = 3519$, $\Delta df = 520$, $p < .001$).

Meerniveau confirmatieve factoranalyse met behulp van de *Full Information Maximum Likelihood* (FIML) methode, zoals beschreven door du Toit en du Toit (2001; zie ook Mels, 2004), leverde voor het complete nieuwe meetmodel in eerste instantie geen goede oplossing op, vanwege een *Heywood case*. Een Heywood case houdt in dat er een negatieve meetfout variantie voorkomt in de parameterschattingen. Heywood cases kunnen bijvoorbeeld worden veroorzaakt door misspecificaties in het model, extreme waarden (*outliers*) in de dataset, weinig indicatoren per factor, of een te kleine steekproef (Chen, Bollen, Paxton, Curran & Kirby, 2001; zie ook Garson, z.j., b). Nadat alle cases met ontbrekende waarden voor de variabele met de negatieve meetfout variantie (item os1 in de open systeem schaal, zie bijlage 4A) uit de dataset waren verwijderd en daarnaast alle cases met meer dan vijf ontbrekende waarden, werd wel een goede oplossing verkregen. In tabel 4.8 staat een overzicht van de passingswaarden die zijn verkregen met de aangepaste dataset van 858 cases (docenten) voor het complete nieuwe meetmodel en met de volledige dataset voor de vier nieuwe schalen afzonderlijk. Uit tabel 4.8 blijkt dat de passingswaarden van het nieuwe model bij een meerniveau confirmatieve factoranalyse beter zijn dan die van het oorspronkelijke model (zie tabel 4.3). Zowel voor het complete model als voor de afzonderlijke schalen voldoet RMSEA aan het criterium voor een goede modelpassing (RMSEA kleiner dan .06 - .08).

Tabel 4.8 Passingsmaten met docentendata, voor het nieuwe model van schoolleidersgedrag ($N=100$, $n = 858$) en afzonderlijke schalen ($N = 100$, $n = 997$), bij een meerniveau confirmatieve factoranalyse met FIML

Nieuw meetmodel	χ^2 (df)	RMSEA	90% C.I. RMSEA
Volledig model	1246.72 (586)	.051	.047 - .055
Rationele doel	73.42 (28)	.057	.041 - .073
Interne proces	8.59 (10)	.000	.000 - .043
Human relations	144.93 (40)	.073	.060 - .085
Open systeem	47.52 (18)	.058	.038 - .078

RMSEA = Root Mean Square Error of Approximation; 90% C.I. = 90% betrouwbaarheidsinterval

Om tenslotte ook nog inzicht te krijgen in de passingswaarden op schoolniveau en op docentenniveau is een meerniveau confirmatieve factoranalyse uitgevoerd met behulp van *Muthén's approximate Maximum Likelihood solution* (MUML). De uitkomsten staan in tabel 4.9. Bij vergelijking van de uitkomsten in tabel 4.9 met die in tabel 4.4 blijkt dat zowel voor het complete nieuwe model, als voor de afzonderlijke schalen, de passingswaarden op schoolniveau (*between groups*) en op docentenniveau (*within groups*) verbeterd zijn. Voor het complete nieuwe model geldt dat de modelpassing op schoolniveau nog matig is, maar voor de afzonderlijke schalen voldoen de passingswaarden op schoolniveau aan de criteria voor een goede modelpassing (RMR kleiner dan .05, SRMR kleiner dan .08). Op docentenniveau voldoen alle passingswaarden aan deze criteria. Voor de afzonderlijke schalen geldt op docentenniveau ook dat GFI groter is dan .95.

Tabel 4.9 Passingsmaten voor het nieuwe model van schoolleidersgedrag en afzonderlijke schalen, bij een meerniveau confirmatieve factoranalyse met MUML, voor docentendata ($N= 100$, $n = 997$)

Nieuw meetmodel	χ^2 (df)	RMSEA	CFI	Niveau	RMR	SRMR	GFI
Volledig model	972.85 (586)	.039	.954	Between	.017	.189	.415
				Within	.021	.044	.930
Rationele doel	73.21 (28)	.057	.985	Between	.003	.046	.809
				Within	.014	.031	.981
Interne proces	3.68 (10)	.000	1.00	Between	.004	.034	.916
				Within	.003	.007	.999
Human relations	121.11 (40)	.064	.974	Between	.007	.073	.753
				Within	.018	.038	.972
Open systeem	52.83 (18)	.062	.970	Between	.007	.072	.898
				Within	.017	.036	.983

RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; (S)RMR = (Standardized) Root Mean square Residual; GFI = Goodness-of-Fit Index

Uit de bovenstaande analyseresultaten is op te maken de passingswaarden van het nieuwe meetmodel van schoolleidersgedrag en van de afzonderlijke schalen beter zijn dan die van het oorspronkelijke meetmodel, zowel voor de schoolleidersdata, als de docentendata. Dit wijst erop dat zowel de discriminante validiteit als de interne consistentie van het meetinstrument verbeterd is.

4.3.5 Betrouwbaarheid en validiteit van het nieuwe model

Uit de passingswaarden van het nieuwe meetmodel voor schoolleidersgedrag kan worden afgeleid dat het model goed past bij de data. Om de validiteit en de betrouwbaarheid van het model nader vast te stellen, zijn de parameterschattingen nader bekeken. In tabel 4.10 staat een overzicht van de parameterschattingen voor het volledige nieuwe meetmodel.

Tabel 4.10 Factorladingen voor het nieuwe meetmodel van schoolleidersgedrag, bij éénniveau confirmatieve factoranalyse aan de hand van schoolleidersdata (n=103) en docentendata (n=896) en meerniveau confirmatieve factoranalyse met FIML of MUML, voor docentendata (N=100, n = 997)

	Schoolleiders		Docenten		meerniveau FIML				meerniveau MUML			
	éénniveau		éénniveau		between		within		between		within	
	css	R ²	css	R ²	css	R ²	css	R ²	css	R ²	css	R ²
Rationele doel												
rd1	.21	.04	.65	.43	.67	.45	.61	.37	.89	.79	.60	.36
rd2	.46	.21	.70	.48	.68	.45	.65	.43	.91	.82	.64	.41
rd3	.43	.18	.60	.36	.76	.58	.56	.32	.74	.55	.56	.32
rd4	.66	.44	.58	.34	.37	.14	.55	.31	.84	.70	.56	.32
rd5	.61	.38	.73	.53	.57	.33	.68	.47	.90	.81	.68	.46
rd6	.52	.27	.69	.48	.50	.25	.67	.44	.88	.78	.67	.45
ip2	.44	.20	.67	.46	.80	.64	.60	.36	.86	.74	.60	.36
Interne process												
ip3	.68	.46	.76	.58	.87	.75	.66	.43	.89	.78	.67	.44
ip4	.54	.29	.72	.52	.98	.97	.59	.35	.95	.91	.63	.39
ip5	.62	.38	.60	.36	.93	.87	.54	.30	.90	.81	.50	.25
ip8	.66	.43	.66	.43	.97	.94	.57	.32	.82	.67	.60	.36
ip10	.34	.12	.43	.19	.67	.44	.42	.18	.64	.41	.37	.14
Human relations												
ip11	.53	.28	.59	.35	.78	.60	.55	.30	.89	.79	.53	.28
hr1	.52	.27	.70	.50	.92	.84	.62	.39	.91	.84	.64	.41
hr2	.77	.60	.71	.51	.82	.66	.64	.41	.82	.67	.65	.42
hr3	.68	.46	.69	.45	.92	.84	.58	.33	.93	.87	.58	.34
hr5	.61	.38	.61	.37	.90	.81	.56	.31	.89	.79	.54	.29
hr8	.48	.23	.58	.33	.89	.80	.48	.23	.80	.64	.49	.24
hr10	.31	.10	.61	.37	.46	.21	.60	.37	.59	.35	.61	.37
hr12	.37	.14	.50	.25	.60	.36	.50	.25	.65	.43	.50	.25
Open systeem												
os1	.79	.62	.67	.45	.90	.80	.58	.34	1.0	1.0	.61	.38
os3	.64	.41	.54	.29	.80	.63	.51	.26	.59	.35	.52	.27
os6	.72	.51	.70	.49	.68	.46	.64	.40	.78	.60	.64	.40
os7	.52	.27	.43	.18	.81	.66	.42	.18	.71	.50	.40	.16
os8	.35	.13	.54	.29	.46	.21	.52	.27	.50	.25	.52	.27
os9	.45	.20	.47	.22	.28	.08	.50	.25	.36	.13	.48	.23
Correlatie												
RD-IP*	.06		.63		.46		.72		.43		.71	
RD-HR	.00		.72		.69		.72		.71		.72	
RD-OS*	.35		.68		-.13		.80		.16		.80	
IP-HR*	.55		.70		.70		.73		.70		.72	
IP-OS	-.09		.46		-.15		.70		-.11		.67	
HR-OS*	-.25		.44		-.51		.68		-.38		.66	

Css = factorlading: volledig gestandaardiseerde oplossing; Factorladingen $\geq .35$ zijn vetgedrukt;

Schuingedrukte factorladingen en correlaties zijn niet significant (kritieke ratio < 1.96 , $p < .05$);

R² = squared multiple correlation = percentage verklaarde variantie in geobserveerde variabele;

FIML en MUML: meerniveau analyse; 'between' = between groups, 'within' = within groups;

RD = rationele doel schaal, IP= interne proces schaal, HR = 'human relations' schaal, OS = open systeem schaal;

* aangrenzende kwadranten in het concurrerende waarden model

Uit de factorladingen in tabel 4.10 is op te maken dat de meeste indicatoren in het nieuwe meetmodel goed laden op de bijbehorende factor. Enkele items laden nog matig (rd1 –alleen voor schoolleiders-, ip10, hr10, hr12, os8 en os9), terwijl aan de waarden voor R^2 is te zien dat bij deze indicatoren een lager percentage van de variantie wordt verklaard, dan bij de andere variabelen. De zwakkere items zijn gehandhaafd om inhoudelijke redenen en vanwege de passing van het meetmodel (op schoolleiders-en/of docentendata). In de uitkomsten van de meerniveau analyses op schoolniveau zijn de waarden voor R^2 in het algemeen hoog. Dit betekent dat de itemscores goed worden voorspeld door de onderliggende factoren. Voor de individuele docenten en de schoolleiders zijn de waarden van R^2 lager.

De goede passing van het nieuwe meetmodel op zowel schoolleidersdata als docentendata geeft aan dat de constructen binnen het meetmodel voldoende verschillen. De correlaties tussen de meetschalen geven echter nog meer inzicht in de discriminante validiteit. Als de correlatie tussen twee factoren groter is dan .85, dan kan de conceptuele overlap tussen de factoren te groot zijn (; Garson, z.j., d; Kenny, 1998). Uit tabel 4.10 is af te lezen dat dit voor de nieuwe meetschalen niet het geval is. De discriminante validiteit is dus voldoende.

Op basis van de onderliggende dimensies in het concurrerende waarden model wordt er door verschillende onderzoekers (Kalliath, Bluedorn & Gillespie, 1999; Lawrence, Quinn & Lenk, 2003) vanuit gegaan dat factoren die naast elkaar liggende kwadranten van het concurrerende waarden model vertegenwoordigen, meer met elkaar zouden moeten correleren, dan factoren die corresponderen met tegenover elkaar liggende kwadranten. Lamond (2003) meent echter dat alle factoren in meer of mindere mate zouden kunnen correleren, omdat de vier leiderschapsmodellen naast elkaar kunnen voorkomen in een organisatie (zie paragraaf 4.2).

De volgende kwadranten grenzen aan elkaar in het concurrerende waarden model: rationele doel en interne proces, interne proces en human relations, ‘human relations’ en open systeem, open systeem en rationele doel . De tegenover elkaar liggende kwadranten zijn: rationele doel en ‘human relations’, interne proces en open systeem (zie figuur 2.3). Uit tabel 4.10 is af te lezen dat de correlaties op docentenniveau (*within groups* bij meerniveau analyse) tussen de factoren allemaal significant, positief en vrij hoog zijn. Op schoolniveau (*between groups* bij meerniveau analyse) zijn er echter duidelijke verschillen te zien in de correlaties. In tabel 4.11 wordt een overzicht gegeven van de relaties tussen de factoren die de vier kwadranten in het concurrerende waarden model vertegenwoordigen, op basis van de gegevens (op schoolniveau) in tabel 4.10.

Tabel 4.11 Overzicht van verbanden tussen de vier factoren voor de handelingen van schoolleiders, aan de hand van schoolleidersdata en docentendata (op basis van gegevens uit tabel 4.9)

Verband tussen handelingen	Schoolleiders	Docenten
rationele doel-interne proces*	n.s.	+ (n.s.) ¹
rationele doel-human relations	n.s.	+
rationele doel-open systeem*	+	n.s.
interne proces-human relations*	+	+
interne proces-open systeem	n.s.	n.s.
human relations-open systeem*	-	-(n.s.) ¹

n.s. = niet significant; + = significant en positief; - = significant en negatief

* aangrenzende kwadranten; ¹ niet significant bij 1 meerniveau analyse

Uit tabel 4.11 blijkt dat de verbanden tussen de factoren niet consistent zijn met de veronderstelling dat naast elkaar liggende kwadranten in het concurrerende waarden model (gemarkeerd met een *) meer met elkaar zouden moeten correleren, dan tegenover elkaar liggende kwadranten. Verder zijn de verbanden tussen de kwadranten ook nog verschillend voor schoolleiders en docenten. Aan de hand van deze uitkomsten kan de validiteit van het onderzoeksinstrument dus worden niet bevestigd. Omdat het in dit onderzoek gaat om één type organisatie, zou het specifieke patroon van correlaties echter kunnen zijn veroorzaakt door actuele maatschappelijke ontwikkelingen, waarop vele schoolleiders met gelijksoortig gedrag reageren (zie paragraaf 4.2). De verschillen tussen schoolleiders en docenten kunnen het gevolg zijn van de uiteenlopende invalshoeken van waaruit deze groepen respondenten naar de schoolorganisatiepraktijk kijken. De correlaties tussen de factoren worden nader besproken in paragraaf 4.6.

4.3.6 Evaluatie van de confirmatieve factoranalyse en de modelverbetering

Het meetmodel voor schoolleidersgedrag is met behulp van LISREL getoetst aan schoolleidersdata en docentendata met behulp van verschillende confirmatieve factoranalyse methoden. Uit tabel 4.5 blijkt dat de parameterschattingen met de *Full Information Maximum Likelihood* (FIML) schattingsmethode op schoolniveau (*between groups*) niet erg consistent zijn (zie 4.3.3). De uitkomsten voor de afzonderlijke schalen lijken beter te kloppen dan die voor het volledige meetmodel. De schattingen met *Muthén's approximate Maximum Likelihood solution* (MUML) op schoolniveau zouden minder nauwkeurig kunnen zijn dan de FIML schattingen, vanwege de vereenvoudigde benadering, maar de uitkomsten lijken acceptabel. De parameterschattingen van de éénniveau confirmatieve factoranalyse voor individuele docentendata liggen dicht bij die van de meerniveau analyses op docentenniveau (*within groups*), omdat de variantie op docentenniveau veel groter is dan op schoolniveau. De éénniveau analyse voor docentendata levert weinig informatie op over de passing van het model op schoolniveau. De uitkomsten van de éénniveau confirmatieve factoranalyses voor schoolleidersdata lijken redelijk consistent te zijn voor het volledige meetmodel en de afzonderlijke schalen. Ook bij uitvoering van deze éénniveau analyses traden echter problemen op, die in het voorgaande nog niet zijn besproken. Zowel de problemen bij de éénniveau analyses voor schoolleidersdata, als die bij de meerniveau analyses met FIML houden waarschijnlijk verband met het relatief beperkte aantal cases op schoolniveau. Dit wordt hieronder toegelicht. De belangrijkste vraag hierbij is hoe accuraat de uitkomsten van de verschillende confirmatieve factoranalyses zijn.

De accuratesse van de parameterschattingen in *structural equation modeling* is gekoppeld aan de steekproefgrootte. In de literatuur zijn verschillende vuistregels te vinden voor de steekproefgrootte (zie Garson, z.j., b). Mitchell (1993) geeft bijvoorbeeld aan dat er tien tot twintig keer zoveel cases zouden moeten zijn als variabelen. Volgens Stevens (1996) zijn er tenminste vijftien cases nodig per geobserveerde variabele. Bij meer dan tien variabelen zou een steekproefgrootte van minder dan 200 cases leiden tot onstabiele parameterschattingen (Garson, z.j., b). Voor multilevel analyse geeft Muthén (1997) aan dat de *maximum likelihood* schattingsmethode al goed werkt bij tenminste 50 groepen. Hox (2002, p.175) stelt dat een steekproef die bestaat uit 100 groepen van tien individuen geschikt is voor complexe multilevel analyses. Bij simulatiestudies (zie bijv. Hox & Maas, 2001) wordt echter meestal gewerkt met relatief eenvoudige modellen. In principe zou een steekproef van 100 scholen met gemiddeld tien docenten per school moeten voldoen voor meerniveau confirmatieve factoranalyse, maar aan de andere kant is in dit onderzoek het aantal cases op schoolniveau

wel erg klein in vergelijking met het aantal variabelen in het meetmodel voor schoolleidersgedrag.

Bij het uitvoeren van éénniveau analyse aan de hand van schoolleidersdata traden problemen op die waarschijnlijk ook een gevolg waren van de steekproefgrootte. LISREL accepteert input in verschillende vormen, bijvoorbeeld een covariantiematrix, een correlatiematrix, of ruwe data. Voor een éénniveau confirmatieve factoranalyse op basis van ordinale data wordt aanbevolen om eerst met PRELIS (een deelprogramma van LISREL) de ‘polychorische correlatie matrix’ en de ‘asymptotische covariantie matrix’ te berekenen, die dan gezamenlijk als input gebruikt worden voor de *Generally Weighted Least Squares* (WLS) schattingsmethode (du Toit & du Toit, 2001). Bij het berekenen van de polychorische correlaties wordt ervan uitgegaan dat ordinale variabelen een onderliggende continue variabele hebben (Jöreskog & Sörbom, 1993a, p. 44). De polychorische correlaties zijn de theoretische correlaties tussen deze onderliggende variabelen. De asymptotische covariantiematrix van de polychorische correlaties is nodig om een z.g. *weight matrix* te berekenen die nodig is voor de WLS schattingsmethode (Jöreskog & Sörbom, 1993a, p. 45).

Voor de schoolleidersgegevens bleken zowel de polychorische correlatie matrix als de asymptotische covariantiematrix niet geschikt te zijn om de WLS schattingsmethode uit te kunnen voeren, vermoedelijk als gevolg van het relatief kleine aantal cases.⁴ Jöreskog en Sörbom (1993b, p. 223) raden af om te werken met een asymptotische covariantiematrix die is gebaseerd op een (te) kleine steekproef. Zij geven aan dat het in zo’n geval beter is om bijvoorbeeld *maximum likelihood* als schattingsmethode te gebruiken in plaats van WLS (zie bijlage 2B).

Vanwege de problemen met de input matrices zijn de éénniveau confirmatieve factoranalyses voor schoolleidersdata en docentendata uitgevoerd met de maximum likelihood schattingsmethode, zoals beschreven door Mels (2004), met ruwe data als input voor LISREL. Hierbij werden de Likert schaal variabelen beschouwd als continue (interval) variabelen. Over het interpreteren van Likert schaal variabelen als interval variabelen is veel discussie (Heiser, z.j.). Jöreskog (2005) geeft aan dat Likert schaal variabelen als ordinale variabelen behandeld zouden moeten worden (zie ook du Toit, du Toit, Mels & Cheng, 2007). In de praktijk worden Likert schaal variabelen echter ook wel geïnterpreteerd als interval variabelen, waarbij bijvoorbeeld wordt gerefereerd aan Tabachnick en Fidell (1983, in Heiser, z.j.). Garson (z.j., b) geeft aan dat voor de weighted least squares schatting een grote steekproef nodig is (meer dan 2000 cases) en dat in diverse onderzoeken is aangetoond dat maximum likelihood en weighted least squares overeenkomstige resultaten opleveren, ook wanneer theoretisch een weighted least squares schatting nodig zou zijn. Op basis van deze overwegingen kan worden aangenomen dat de uitkomsten van de éénniveau confirmatieve factoranalyses in de tabellen 4.5 en 4.9 accuraat zijn.

Bij de meerniveau confirmatieve factoranalyses voor de docentendata werd gebruik gemaakt van twee schattingsmethoden. Volgens Mels en Von Briesen (Scientific Software

⁴ De polychorische correlatie matrix en de asymptotische covariantiematrix waren ‘niet positief definit’. Dat betekent dat de determinant van deze matrices niet positief was. De determinant kan worden geïnterpreteerd als een maat voor ‘algemene’ variantie (Rigdon, z.j.). Negatieve variantie bestaat niet. Volgens Rigdon (z.j.) zijn er vele mogelijke oorzaken voor het niet positief definit zijn van covariantie- of correlatiematrices, zoals lineaire afhankelijkheid tussen variabelen, ontbrekende data, toevallige variaties in de steekproef (vooral bij kleine steekproeven) of een zeer groot aantal variabelen in de matrix.

International⁵, persoonlijke communicate, november 2006) moet de MUMML methode als minder accuraat worden beschouwd dan de FIML methode, omdat MUMML uitgaat van onafhankelijke groepen voor het schoolniveau en het docentenniveau. Hox en Maas (2001) concluderen op basis van simulatiestudies echter dat de uitkomsten van de MUMML methode op basis van één *within* groep en één *between* groep, vergeleken met FIML, goede resultaten oplevert, zolang het aantal groepen groot genoeg is. Deze auteurs wijzen erop dat het *within groups* deel van het model meestal geen problemen oplevert, omdat op dit niveau voldoende observaties zijn. Het aantal groepen speelt vooral een rol op het *between groups* niveau. Bij een klein aantal groepen, gecombineerd met een lage intraklasse correlatie is volgens Hox en Maas (2001) het percentage ‘onaanvaardbare oplossingen’ (*inadmissible solutions*) vrij groot. De onaanvaardbare oplossingen worden gekenmerkt door extreme waarden in de parameter schattingen. Als een acceptabele oplossing wordt bereikt zijn de geschatte factorladingen echter in het algemeen accuraat. Om onaanvaardbare oplossingen en afwijkingen in andere parameterschattingen te voorkomen bevelen Hox en Maas (2001) aan om tenminste honderd groepen te analyseren. Aan de voorwaarde van tenminste honderd groepen wordt in hier beschreven onderzoek naar het effect van schoolleidersgedrag (net) voldaan, maar zoals eerder opgemerkt is het de vraag of deze regel ook geldt voor een complex meetmodel met veel parameters. In principe zouden de MUMML parameterschattingen in de tabellen 4.5 en 4.9 accuraat moeten zijn. De FIML parameterschattingen voor het volledige meetmodel in tabel 4.5 vertonen opvallende afwijkingen ten opzichte van de schattingen voor de afzonderlijke schalen. De uitkomsten voor de afzonderlijke schalen zijn waarschijnlijk beter, omdat het aantal te schatten parameters kleiner is, zodat er meer cases zijn per parameter. De MUMML parameterschattingen voor het volledige meetmodel liggen dicht bij de FIML schattingen voor de afzonderlijke schalen. Wellicht is de FIML methode (in LISREL) gevoeliger voor het aantal cases op schoolniveau en zijn de MUMML schattingen in dit geval toch beter.

Voor het optimaliseren van het meetmodel voor schoolleidersgedrag zijn de uitkomsten van verschillende confirmatieve factoranalyses vergeleken. De FIML schattingen voor het volledige model werden hierbij als het minst robuust beschouwd. Door het verwijderen van minder betekenisvolle en niet eenduidige items, zijn de passingswaarden van het volledige meetmodel en van de afzonderlijke schalen verbeterd. De vraag is nu in hoeverre hier sprake is van kanskapitalisatie. Door de ‘datagedreven’ aanpassingen in het meetmodel zou het resultaat specifiek kunnen zijn voor de steekproef en niet generaliseerbaar naar de hele populatie (zie bijv. Kelloway, 1998, p. 22; MacCallum, Roznowski & Necowitz, 1992; Stevens, 2002). Een veelgebruikte strategie om de geldigheid van een model te testen, is kruisvalidering aan de hand van twee onafhankelijke datasets. Als er geen onafhankelijke datasets beschikbaar zijn, kan de dataset in tweeën gesplitst worden. De ene helft van de data dient dan om het model vast te stellen en de andere helft voor de validering (Jöreskog en Sörbom, 1993a, p.129). Vanwege de complexiteit van het meetmodel voor schoolleidersgedrag is het splitsen van de dataset hier echter geen goede optie, omdat het aantal cases op schoolniveau te klein zou worden om acceptabele schattingen op te leveren. Een andere aanbevolen strategie is om vooraf op basis van de theorie verschillende modellen op te stellen en die te testen (MacCallum, Roznowski & Necowitz, 1992). Door Schmidt (2009, zie ook Sleegers & van Schooten, 2006, p.31) is voor de handelingen van schoolleiders, als alternatief voor het vier factoren model, een ‘facet-model’ getest waarbij elk item laadt op twee van de vier oriëntaties (kwadranten) in het concurrerende waarden model. Het facet-model bleek echter minder goede resultaten op te leveren dan het model waarbij elk

⁵ SSI is de leverancier van LISREL

item laadt op één factor. Vanwege de koppeling aan het concurrerende waarden model is het vier factoren model voor het meten van de handelingen van schoolleiders verder gehandhaafd.

Uit simulatiestudies van MacCallum, Roznowski en Necowitz (1992) kan worden opgemaakt dat modelverbetering voor grote steekproeven van 800-1200 cases stabiele resultaten oplevert bij kruisvalidering, terwijl bij kleinere steekproeven de resultaten uiteenlopen. In dit onderzoek is de omvang van de steekproef op docentenniveau groot genoeg, maar het aantal cases op schoolniveau niet. Naar aanleiding van hun simulatiestudies raden MacCallum, Roznowski en Necowitz (1992) aan om een *parallel specification search* uit te voeren. Dat betekent dat het modelverbeteringsproces voor twee onafhankelijke steekproeven wordt doorlopen, waarna de uitkomsten worden vergeleken. Daarbij zou ook een dubbele kruisvalidering uitgevoerd moeten worden. In dit onderzoek is de parallelle procedure benaderd doordat bij de modelverbetering voor schoolleidersgedrag werd uitgegaan van zowel schoolleidersdata als docentendata. Er heeft echter geen kruisvalidering plaatsgevonden en bovendien waren schoolleiders en docenten afkomstig van dezelfde 100 scholen. Desondanks is het risico van kanskapitalisatie wellicht beperkt door de wijze waarop het model voor de handelingen van schoolleiders is verbeterd. Bij een *specification search* worden volgens MacCallum et al. (1992) in het algemeen parameters *toegevoegd* om de modelpassing te verbeteren (bijvoorbeeld covarianties tussen indicatoren). Bij de verbetering van het meetmodel voor de handelingen van schoolleiders werden alleen indicatoren *verwijderd* en enkele verplaatst. Daarbij werd steeds nagegaan in hoeverre de vier schalen inhoudelijk in overeenstemming waren met het concurrerende waarden model. Feitelijk werden telkens alternatieve vier factoren modellen getest (zie ook bijlage 2C).

Samenvattend kan worden gesteld dat voor de (geclusterde) docentendata meerniveau confirmatieve factoranalyse een beter inzicht gaf in de constructvaliditeit van het onderzoeksinstrument dan éénniveau analyse. Meerniveau analyse in LISREL met FIML leverde echter geen stabiele resultaten op vanwege het complexe meetmodel van schoolleidersgedrag. Waarschijnlijk was de steekproef van 100 scholen niet toereikend voor deze analysemethode. De lage intraklasse correlatie voor de docentendata zal hierbij ook van invloed zijn geweest (Hox & Maas, 2001). De MUML methode was wel robuust. De beperkte steekproefgrootte op schoolniveau speelde ook een rol bij de (éénniveau) analyse van schoolleidersdata. Hoewel niet is uit te sluiten dat er bij de verbetering van het meetmodel aan de hand van schoolleidersdata en docentendata kanskapitalisatie optrad, is het effect waarschijnlijk beperkt vanwege de parallelle verbeteringsprocedure, zowel aan de hand van schoolleidersdata, als aan de hand van docentendata en de behoudende aanpak, waarbij alleen indicatoren werden verwijderd of verplaatst en geen parameters (paden) aan het model werden toegevoegd.

De zwakke items die aan het licht kwamen bij meerniveau confirmatieve factoranalyse werden ook bij exploratieve factoranalyses gedetecteerd (zie 4.3.4). In principe zou het model dus verbeterd kunnen worden aan de hand van de uitkomsten van exploratieve factoranalyse. Daarbij kan echter geen vooraf opgesteld model worden getest, zodat de modelverbetering nog meer datagedreven zou zijn. Verder levert confirmatieve factoranalyse betere parameterschattingen op dan exploratieve factoranalyse, omdat rekening wordt gehouden met meetfouten. Het was dan ook de bedoeling om aan de hand van het verbeterde meetmodel factorscores (latente variabele scores) te berekenen op schoolniveau, die gebruikt zouden kunnen worden bij de causale analyse van het onderzoeksmodel (zie ook De Maeyer en Rymenans, 2004, p.285). Nadat de meerniveau confirmatieve factoranalyses waren uitgevoerd, bleek echter dat het in LISREL niet mogelijk was om bij FIML factorscores op

schoolniveau te genereren⁶ (Mels & Von Briesen, SSI, persoonlijke communicatie, november 2006). Dat betekent dat de latente variabele scores voor de schoolleidersdata met behulp van éénniveau confirmatieve factoranalyse berekend zouden moeten worden en geaggregeerd. In principe is het niet nodig om factorscores te berekenen voor een causale analyse, omdat meetmodellen kunnen worden opgenomen in het structurele model. Vanwege de omvang van het meetmodel voor de handelingen van schoolleiders en de relatief beperkte steekproefgrootte was dit voor de validering van het onderzoeksmodel (zie hoofdstuk 6) geen goede optie. De bruikbaarheid van meerniveau confirmatieve factoranalyse bleef dus beperkt tot het bepalen van de constructvaliditeit van het onderzoeksinstrument.

4.4 Betrouwbaarheid en validiteit van de vragenlijsten over schoolcultuur en schoolorganisatie

Voor het meten van de schoolcultuur en de schoolorganisatiepraktijk zijn twee vragenlijsten ontwikkeld voor docenten, op basis van het concurrerende waarden model, met elk in totaal veertig items, verdeeld over vier schalen (zie bijlagen 1C en 1D). Met behulp van een vierpunts Likertschaal kan worden aangegeven in hoeverre de uitspraken over de schoolorganisatiepraktijk of de schoolcultuur geldig zijn voor de praktijk op school. De Likertschaal loopt van ‘niet of nauwelijks’ (1), tot ‘in zeer sterke mate’ of ‘in hoge mate’ (4) voor respectievelijk schoolcultuur en schoolorganisatie. De docentenvragenlijsten zijn ingevuld door 997 docenten van 100 scholen. De betrouwbaarheid en de validiteit van de instrumenten is onderzocht met behulp van confirmatieve factoranalyse. Naar aanleiding van de uitkomsten van deze analyses zijn de meetmodellen verbeterd. Hieronder worden eerst de uitkomsten van het onderzoek naar de betrouwbaarheid en validiteit van de vragenlijst over schoolcultuur beschreven en daarna die voor de vragenlijst over de schoolorganisatiepraktijk.

4.4.1 Betrouwbaarheid en validiteit van de vragenlijst over schoolcultuur

Voor de 40 items over schoolcultuur (zie bijlage 1C) bedraagt de intraklasse correlatie gemiddeld 0.10. Bij deze intraklasse correlatie en een groepsgrootte van gemiddeld tien docenten, is het design effect gelijk aan 1.9. Dit zou betekenen dat er bij de analyse van de gegevens geen rekening hoeft te worden gehouden met het effect van clustering (het design effect is kleiner dan 2). Kennelijk is er binnen de scholen tussen docenten slechts beperkte overeenstemming over de schoolcultuur (zoals gemeten door het onderzoeksinstrument). Toch zijn behalve éénniveau ook meerniveau confirmatieve factoranalyses uitgevoerd voor schoolcultuur, om enig inzicht te krijgen in de passing van het meetmodel op schoolniveau en om op die manier het model eventueel te kunnen verbeteren. Voor schoolcultuur bestond het meetmodel uit vier factoren met elk tien indicatoren.

Voor de éénniveau confirmatieve factoranalyse, op basis van ruwe data, zijn eerst alle cases met meer dan vijf ontbrekende waarden voor schoolcultuur uit de datafile verwijderd. Hierna bleven er 972 cases over. De overige ontbrekende waarden werden vervangen door de gemiddelde score. De meerniveau confirmatieve factoranalyse is uitgevoerd met behulp van *Muthén's approximate Maximum Likelihood solution* (MUML), omdat deze methode in voorgaande analyses robuuster bleek dan *Full Information Maximum Likelihood* (FIML) voor

⁶ Voor MUML is het in principe wel mogelijk om factorscores te genereren, omdat hier wordt uitgegaan van onafhankelijke groepen op schoolniveau en docentenniveau. De parameterschattingen en factorscores zijn echter minder accuraat dan bij éénniveau analyse (Mels & Von Briesen, SSI, persoonlijke communicatie, november 2006).

het passen van een complex meetmodel op de beschikbare docentendata (zie 4.3.6). Bij de éénniveau confirmatieve factoranalyse bleken alle items voor schoolcultuur significant te laden op de bijbehorende factoren. Bij de meerniveau confirmatieve factoranalyse bleken op schoolniveau veertien items niet significant te laden⁷. Om nader inzicht te krijgen in de factorstructuur werd een principale componenten analyse uitgevoerd⁸. Op basis van de uitkomsten van confirmatieve en exploratieve factoranalyse, gecombineerd met inhoudelijke overwegingen aan de hand van het concurrerende waarden raamwerk, zijn elf items verwijderd uit het model en twee indicatoren verplaatst naar een andere factor⁹. Bij de inhoudelijke overwegingen werd nagegaan in hoeverre een item tezamen met de andere items in een schaal één geheel vormt en in hoeverre het item helder en eenduidig is geformuleerd. Verder werd met behulp van Cronbach's alpha bepaald in hoeverre het verwijderen van een item gevolgen zou hebben voor de interne consistentie van de schaal.

Het nieuwe meetmodel dat werd geconstrueerd voor schoolcultuur bevat 29 indicatoren (items). Een overzicht van de schalen (factoren) en de bijbehorende items is te vinden in bijlage 4B. In de nieuwe rationele doel schaal (6 items) draait het om bevordering van de productiviteit door een gerichtheid op leerresultaten en een ordelijke werksfeer. De interne proces schaal (7 items) is gericht op zorgvuldigheid in de interne organisatie. In de 'human relations' schaal (9 items) gaat het om ontwikkeling van het personeel en samenwerking binnen de school. De open systeem schaal meet vooral een gerichtheid op profilering en vernieuwing van de school.

Om te bepalen of het nieuwe meetmodel betere passingswaarden heeft dan het oorspronkelijke model, zijn beide modellen vergeleken. In tabel 4.12 staat een overzicht van de passingswaarden die zijn verkregen voor het oorspronkelijke meetmodel van schoolcultuur (40 items) en voor het nieuwe model (29 items), bij zowel éénniveau als meerniveau confirmatieve factoranalyse. Daarnaast is ook de correlatie tussen de schalen in het oorspronkelijke en het nieuwe model vermeld en Cronbach's alpha voor de oorspronkelijke en nieuwe schalen.

Uit tabel 4.12 is af te lezen dat de passingsmaten SRMR en CFI wijzen op een goede modelpassing voor het oorspronkelijke éénniveau meetmodel van schoolcultuur (SRMR is kleiner dan of gelijk aan .08; CFI is groter dan of gelijk aan .95). De waarde voor RMSEA is redelijk goed, de bovengrens van het 90% betrouwbaarheidsinterval ligt beneden .08, maar de ondergrens zou lager moeten zijn dan .06. Het meerniveau meetmodel heeft betere waarden voor RMSEA. Het nieuwe model past significant beter dan het oorspronkelijke, zowel bij éénniveau als meerniveau confirmatieve factoranalyse (gebaseerd op $\Delta\chi^2$ en Δdf , $p < .001$). De waarden voor RMSEA, SRMR en CFI voldoen voor het nieuwe model allemaal aan de criteria voor een goede modelpassing. Ook het verschil in AIC duidt op een verbetering in de modelpassing (zie bijlage 2F). Aan de waarden voor Cronbach's alpha is te zien dat de interne consistentie van het volledige model en de afzonderlijke schalen in het nieuwe model goed is en vrijwel gelijk aan die in het oorspronkelijke model.

⁷ Niet significant op schoolniveau voor schoolcultuur (bijlage 1C): rd2, rd7, rd8, ip8, ip9, ip10, hr3, hr8, hr10, os1, os2, os3, os4, os5.

⁸ Vier factoren voor schoolcultuur uit de principale componenten analyse (bijlage 1C): r1 t/m rd5 en ip1; rd6 t/m rd10, ip2 t/m ip8 en ip10; ip9 en hr1 t/m hr10; os1 t/m os10.

⁹ De items rd6 t/m rd10, ip8, hr3, hr8, os2, os3 en os5 zijn verwijderd; ip1 en ip9 zijn verplaatst (bijlage 1C).

Tabel 4.12 *Passingsmaten, correlaties tussen de schalen en Cronbach's alpha voor het oorspronkelijke en het nieuwe model van schoolcultuur, bij éénniveau confirmatieve factoranalyse (n = 972) en meerniveau confirmatieve factoranalyse (N=100, n = 997) met MUML*

Meetmodel	Oorspronkelijk		Nieuw	
	éénniveau	meerniveau	éénniveau	meerniveau
Passingswaarden				
χ^2 (df)	4976.78 (734)	4196.97 (1468)	1824.13 (371)	1632.07 (742)
RMSEA	.077	.061	.064	.049
90% C.I. RMSEA	.075 - .079	.059 - .063	.061 - .066	.046 - .052
SRMR	.070	-	.064	-
CFI	.95	.94	.96	.95
Model AIC	5149	4540	1952	1888
Correlatie¹		between schools		between schools
RD-IP*	.96	.81	.63	.57
RD-HR	.75	.07	.43	-.35
RD-OS*	.62	-.08	.29	-.45
IP-HR*	.81	.34	.73	.34
IP-OS	.68	.14	.59	.15
HR-OS*	.77	.84	.74	.79
Cronbach's alpha				
Totaal model	.93 (40 items)		.91 (29 items)	
Rationele doel (RD)	.81 (10 items)		.80 (6 items)	
Interne proces (IP)	.83 (10 items)		.80 (7 items)	
Human relations (HR)	.87 (10 items)		.86 (9 items)	
Open systeem (OS)	.80 (10 items)		.78 (7 items)	

RMSEA = Root Mean Square Error of Approximation; 90% C.I. = 90% betrouwbaarheidsinterval;

CFI = Comparative Fit Index; SRMR = Standardized Root Mean square Residual;

AIC = Akaike Information Criterion;

¹ Schuingedrukte correlaties zijn niet significant.

* Aangrenzende kwadranten in het concurrerende waarden model

In het nieuwe meetmodel zijn de (significante) correlaties tussen de schalen duidelijk lager dan die in het oorspronkelijke model (zie tabel 4.12). Vooral de correlaties tussen de rationele doel en interne proces schaal en de interne proces en 'human relations' schaal (op docentenniveau) en de 'human relations' en open systeem schaal (op schoolniveau) in het oorspronkelijke model duiden op een mogelijke overlap tussen de concepten. In het nieuwe model is de discriminante validiteit voldoende, omdat de modelpassing goed is en de correlatie tussen de schalen acceptabel (oftewel laag genoeg). Als vuistregel wordt wel aangehouden dat de correlatie tussen twee schalen kleiner moet zijn dan .85 voor voldoende discriminante validiteit (Garson, z.j., d; Kenny, 1998). De correlatie tussen de 'human relations' en open systeem schaal in het nieuwe meetmodel is op schoolniveau echter nog wel hoog (.79).

Bij de uitkomsten voor de individuele docenten (éénniveau analyse) stemmen de relatief hoge correlaties tussen rationele doel en interne proces, interne proces en 'human relations' en tussen 'human relations' en open systeem overeen met de verwachting dat factoren die aangrenzende kwadranten in het concurrerende waarden model vertegenwoordigen, meer zouden moeten correleren dan factoren van tegenover liggende kwadranten (Kalliath, Bluedorn & Gillespie, 1999; Lawrence, Quinn & Lenk, 2003). De correlatie tussen interne proces en open systeem (tegenover liggende kwadranten) is echter ook behoorlijk hoog, terwijl de correlatie tussen rationele doel en open systeem (aanliggende kwadranten) juist vrij laag is. Op schoolniveau (meerniveau analyse) zijn alleen de correlaties tussen rationele doel

en interne proces en tussen ‘human relations’ en open systeem significant. Net als bij de handelingen van schoolleiders wordt de hypothese dat aanliggende kwadranten van het concurrerende waarden model meer met elkaar zouden moeten correleren dan tegenover liggende kwadranten, niet ondersteund. De correlaties tussen de factoren voor schoolcultuur worden nader besproken paragraaf 4.6.

Omdat het nieuwe model voor schoolcultuur een betere modelpassing heeft dan het oorspronkelijke model en tevens een goede convergente en discriminante validiteit heeft, wordt het nieuwe model beschouwd als een geschikter model voor het meten van de schoolcultuur dan het oorspronkelijke model. Parameterschattingen voor het nieuwe model, die zijn verkregen met behulp van éénniveau en meerniveau confirmatieve factoranalyse, zijn opgenomen in bijlage 5A. De factorladingen bij éénniveau analyse zijn allemaal significant en redelijk hoog. De verklaarde varianties (R^2) in de indicator variabelen zijn acceptabel. Het gemiddelde is ongeveer 0.4, dit komt overeen met een betrouwbaarheid (alpha) van 0.6. Bij meerniveau analyse zijn de factorladingen en de verklaarde varianties op schoolniveau hoger dan bij éénniveau analyse, behalve voor vijf items die ook in het oorspronkelijke model op schoolniveau niet significant laden op de bijbehorende factor (zie bijlage 5A). Deze items zijn gehandhaafd vanwege inhoudelijke redenen. Bij het bepalen van de significantie van de factorladingen op schoolniveau vormt de steekproefgrootte een beperkende factor.

4.4.2 Betrouwbaarheid en validiteit van de vragenlijst over de schoolorganisatiepraktijk

De intraklasse correlatie coëfficiënt voor de 40 items betreffende de schoolorganisatiepraktijk (zie bijlage 1D) bedraagt gemiddeld 0.14. Het design effect is dan voor een groepsgrootte van tien docenten gelijk aan 2.3. Bij de analyse van de docentengegevens dient dus rekening gehouden te worden met het effect van clustering. Om de betrouwbaarheid en de validiteit van de vragenlijst te bepalen, zijn zowel éénniveau als meerniveau confirmatieve factoranalyses uitgevoerd. Voor de éénniveau confirmatieve factoranalyse, op basis van ruwe data, zijn alle cases met meer dan vijf ontbrekende waarden voor schoolorganisatie verwijderd. De overige ontbrekende waarden werden vervangen door de gemiddelde score. De nieuwe ruwe data file bevatte 955 cases.

Het meetmodel dat werd getest, was opgebouwd uit vier factoren met elk tien indicatoren. Uit de parameterschattingen van de éénniveau confirmatieve factoranalyse bleek dat alle items significant laden op de bijbehorende factoren, op één na¹⁰. Bij de meerniveau confirmatieve factoranalyse zijn de factorladingen op schoolniveau voor twintig items niet significant¹¹. Uit een principale componenten analyse werd duidelijk dat de items die wel significant laden op schoolniveau, bijna allemaal op eenzelfde factor laden. Dit wijst op een grote overlap tussen de oorspronkelijk geconstrueerde schalen voor het bepalen van de schoolorganisatiepraktijk, mogelijk als gevolg van het feit dat de meeste items over de schoolorganisatiepraktijk eigenlijk gaan over interne processen. Deze items laden allemaal op een ‘interne proces factor’. Om toch te komen tot vier min of meer onafhankelijke schalen voor schoolorganisatie, die passen bij de vier modellen binnen het concurrerende waarden model, zijn op basis van de uitkomsten van confirmatieve en exploratieve factoranalyse, inhoudelijke overwegingen en een analyse van Cronbach’s alpha, zoveel mogelijk zwakke en niet

¹⁰ Niet significant op docentenniveau voor schoolorganisatie (bijlage 1D): hr10 (zie bijlage 1D).

¹¹ Niet significant op schoolniveau voor schoolorganisatie (bijlage 1D): rd1 t/m rd4, ip1 t/m ip4, ip8, ip9, hr3, hr7, hr10 en os1 t/m os7.

eenduidige items uit het model verwijderd¹². Uit de overblijvende items werden vier nieuwe schalen geconstrueerd. De vraag of het gerechtvaardigd was om vast te houden aan een vier factoren model voor schoolorganisatie, ondanks de overeenkomst tussen de items die werd geconstateerd, komt later in dit hoofdstuk nog aan de orde.

Het nieuwe meetmodel voor de schoolorganisatiepraktijk bevat 26 indicatoren. Een overzicht van de schalen (factoren) en de bijbehorende items is te vinden in bijlage 4C. In de nieuwe rationele doel schaal gaat het zowel om een nadruk op leerprestaties, als om een ordelijke en op productie gerichte atmosfeer. De interne proces schaal draait om zorgvuldigheid in de interne organisatie, ook met betrekking tot vernieuwingen. De ‘human relations’ schaal ligt inhoudelijk dicht tegen de interne proces schaal aan, maar in deze schaal gaan de items vooral over ontwikkeling, zowel van docenten als van de school als geheel. De open systeem schaal vertegenwoordigt een gerichtheid op externe contacten. Om het oorspronkelijke en het nieuwe meetmodel voor de schoolorganisatiepraktijk te kunnen vergelijken zijn tabel 4.13 de passingswaarden vermeld van beide modellen bij confirmatieve factoranalyse, en daarnaast de correlaties tussen de schalen en Cronbach’s alpha voor alle schalen.

Tabel 4.13 Passingsmaten, correlaties tussen de schalen en Cronbach’s alpha voor het oorspronkelijke en het nieuwe model van schoolorganisatie, bij éénniveau confirmatieve factoranalyse ($n = 955$) en meerniveau confirmatieve factoranalyse ($N=100$, $n = 997$) met MUML.

Meetmodel	Oorspronkelijk		Nieuw		
	Passingswaarden	éénniveau	meerniveau	éénniveau	meerniveau
χ^2 (df)		3593.13 (734)	3263.93 (1468)	1153.24 (293)	1046.21 (586)
RMSEA		.064	.050	.055	.040
90% C.I. RMSEA		.062 - .066	.047 - .052	.052 - .059	.034 - .044
SRMR		.054	-	.056	-
CFI		.94	.93	.95	.97
Model AIC		3765	3608	1269	1278
Correlatie¹		between schools		between schools	
RD-IP*		.95	1.0	.63	.31
RD-HR		.83	.88	.56	.28
RD-OS*		.87	.78	.53	.13
IP-HR*		.92	.88	.78	.86
IP-OS		.86	.73	.55	.16
HR-OS*		.90	.82	.52	.14
Cronbach’s alpha					
Totaal model		.92 (40 items)		.89 (26 items)	
Rationele doel (RD)		.78 (10 items)		.74 (7 items)	
Interne proces (IP)		.77 (10 items)		.80 (8 items)	
Human relations (HR)		.74 (10 items)		.78 (7 items)	
Open systeem (OS)		.71 (10 items)		.71 (4 items)	

RMSEA = Root Mean Square Error of Approximation; 90% C.I. = 90% betrouwbaarheidsinterval;

CFI = Comparative Fit Index; SRMR = Standardized Root Mean square Residual;

AIC = Akaike Information Criterion;

¹ Schuingedrukte correlaties zijn niet significant.

* Aangrenzende kwadranten in het concurrerende waarden model

¹² Verwijderd uit het meetmodel voor schoolorganisatie (bijlage 1D): rd1, rd6, rd7, rd10, ip4, ip5, ip8, ip9, hr4, hr5, hr6, hr8, hr10, os6; verplaatst: rd8, rd9, ip1, ip2, ip3, ip7, ip10, hr7, os4, os7 t/m os10.

Uit tabel 4.13 is af te lezen dat het oorspronkelijke model van de schoolorganisatiepraktijk een behoorlijk goede modelpassing vertoont, zowel bij de éénniveau als de meerniveau confirmatieve factoranalyse. Alleen de waarde voor CFI is nog iets te laag vergeleken met het criterium voor een goede fit. Het nieuwe model past echter nog (significant) beter dan het oorspronkelijke model (RMSEA kleiner dan .06 - .08; SRMR kleiner dan of gelijk aan .08; CFI groter dan of gelijk aan .95).

In het oorspronkelijke model zijn de correlaties tussen de schalen erg hoog (zie tabel 4.13), dit wijst op een conceptuele overlap tussen de schalen. In het nieuwe model zijn de correlaties tussen de schalen aanzienlijk minder hoog. De correlatie tussen de interne proces en 'human relations' schaal is het hoogst (.78 bij éénniveau analyse), de overige correlaties zijn iets minder hoog en verschillen onderling weinig van elkaar. De uitkomsten van de éénniveau confirmatieve factoranalyse wijzen erop dat de discriminante validiteit van het nieuwe model voldoende is. Bij meerniveau analyse is alleen de correlatie op schoolniveau tussen interne proces en 'human relations' nog significant en groter dan .85 (het afkapcriterium voor discriminante validiteit). De hypothese dat naast elkaar liggende kwadranten van het concurrerende waarden model meer met elkaar zouden moeten correleren dan tegenover elkaar liggende (zie paragraaf 4.2), wordt ook in het model voor de schoolorganisatiepraktijk niet ondersteund. De waarden voor Cronbach's alpha in tabel 4.13 geven aan dat de interne consistentie van het volledige nieuwe model en de afzonderlijke schalen redelijk tot goed is. De waarden voor het nieuwe model zijn vrijwel gelijk aan die voor het oorspronkelijke model, ondanks het feit dat het aantal items verminderd is.

Omdat het nieuwe model voor de schoolorganisatiepraktijk een betere discriminante validiteit heeft, wordt het beschouwd als een geschikter model voor het meten van de schoolorganisatie dan het oorspronkelijke model. Parameterschattingen voor het nieuwe model, die zijn verkregen met behulp van éénniveau en meerniveau confirmatieve factoranalyse, zijn opgenomen in bijlage 5B. Uit de resultaten van de éénniveau confirmatieve factoranalyse voor het nieuwe model blijkt dat alle indicatoren significant en redelijk hoog laden op de bijbehorende factoren. De verklaarde varianties (R^2) in de indicator variabelen zijn acceptabel. Op schoolniveau zijn de factorladingen en de verklaarde varianties hoger dan op het niveau van individuele docenten. Ook in het nieuwe meetmodel zijn er enkele variabelen die op schoolniveau niet significant laden op de bijbehorende factoren. Deze items zijn gehandhaafd om inhoudelijke redenen.

4.5 Betrouwbaarheid en validiteit van de vragenlijst over schoolbeleving

De schoolbeleving van leerlingen is gemeten met een vragenlijst bestaande uit vier schalen, die voor een deel zijn ontleend aan de *Student Participation and Identification Survey* van Leithwood c.s. (Leithwood & Jantzi, 2000a; Leithwood, Aitken & Jantzi, 2000). De vragenlijst over schoolbeleving bevat in totaal 42 items, verdeeld over vier schalen: 'Zich thuisvoelen op school' (10 items), 'Waardering' (10 items), 'Kwaliteit van de Instructie' (17 items) en 'Family Educational Culture' (5 items). Een overzicht van deze vier schalen, met subschalen, is te vinden in bijlage 1E. Met behulp van een vierpunts Likertschaal die loopt van 'helemaal mee oneens' (1) tot 'helemaal mee eens' (4) kunnen leerlingen aangeven in hoeverre zij het eens zijn met de uitspraken in de vragenlijst.

De gemiddelde intraklasse correlatie coëfficiënt (ρ_1) voor de schoolbelevingsitems is 0.04. Deze intraklasse correlatie is kleiner dan die voor de items in de docentenvragenlijsten, omdat het bij de schoolbeleving van leerlingen meer gaat om persoonlijke kenmerken en ervaringen,

dan om schoolkenmerken. De intraklasse correlatie van 0.04 is precies de grenswaarde voor meerniveau analyse, als er geen rekening wordt gehouden met de groepsgrootte. Omdat gemiddeld 44 leerlingen per school de vragenlijst hebben ingevuld is het design effect $(1+(n-1)\rho_1)$ gelijk aan 2.73. Dat betekent dat er bij de analyse van de leerlingengegevens wel rekening zou moeten worden gehouden met de hiërarchische structuur van de data. Vanwege de haken en ogen aan meerniveau confirmatieve factoranalyse in LISREL (zie 4.3.6), is ervoor gekozen om voor de schoolbelevingsvragenlijst alleen éénniveau confirmatieve factoranalyse toe te passen.

Omdat er zoveel leerlingengegevens beschikbaar waren (4336 cases), zijn voor éénniveau confirmatieve factoranalyse alle cases met ontbrekende waarden verwijderd. De nieuwe ruwe data file bevatte nog 3760 cases. Uit de uitkomsten van de confirmatieve factoranalyse voor een meetmodel met vier factoren, bleek dat alle items voor schoolbeleving significant laden op de bijbehorende factoren. Voor enkele items is de factorlading vrij laag¹³. Verder is er een hoge correlatie tussen de schalen 'Zich thuisvoelen op school' en 'Waardering' en de schalen 'Zich thuisvoelen op school' en 'Kwaliteit van de Instructie' (zie tabel 4.14). Dit wijst op een mogelijk conceptuele overlap.

Aan de hand van de uitkomsten van exploratieve factoranalyse is een nieuw meetmodel geconstrueerd met drie factoren voor de schoolbeleving van leerlingen. Vijf items met een lage factorlading werden uit het meetmodel verwijderd¹⁴. In het drie factoren model (bijlage 4D) zijn de subschalen anders gegroepeerd dan in het oorspronkelijke model (bijlage 1E). De schaal 'Waardering van de school' met de subschalen 'Omgang met medeleerlingen' en 'Waardering van de school' geeft aan hoe leerlingen de sfeer op school ervaren. De schaal 'Waardering van onderwijs' met de subschalen 'Waardering van onderwijs' en 'Family Educational Culture' zegt iets over de achtergrond van de leerling. Het 'Werkklimaat' voor leerlingen wordt bepaald door zowel de 'Omgang met leraren' en de 'Ondersteuning door leraren', als het 'Klassenklimaat', de 'Organisatie van het onderwijs' en de 'Invulling van de lessen'. In tabel 4.14 staan de uitkomsten van confirmatieve factoranalyses voor het oorspronkelijke vier factoren model en het nieuwe drie factoren model. In de tabel zijn ook de waarden van Cronbach's alpha opgenomen voor de oorspronkelijke en de nieuwe schalen.

Uit de passingsmaten in tabel 4.14 blijkt dat het oorspronkelijke vier factoren model een redelijke modelpassing vertoont. Het nieuwe model past echter significant beter (gebaseerd op $\Delta\chi^2$ en Δdf , $p < .001$) en voldoet volledig aan de criteria voor een goede modelpassing (RMSEA kleiner dan .06 - .08; SRMR kleiner dan of gelijk aan .08; CFI groter dan of gelijk aan .95). Ook AIC geeft een verbetering aan in de modelpassing voor het nieuwe model. In het vier factoren model zijn de correlaties tussen een aantal schalen erg hoog. In het drie factoren model zijn de correlaties redelijk, zodat een verband tussen de begrippen verondersteld mag worden, maar niet te hoog. De waarden voor Cronbach's alpha geven aan dat de interne consistentie van de nieuwe schalen redelijk tot goed is.

¹³ Items over schoolbeleving (bijlage 1E) met een lage factorlading (<.35): llomlln3, llwaond1, llwaond4, llresult1, llresult2, llinvul3 en llinvul4.

¹⁴ Verwijderd uit het meetmodel voor schoolbeleving (zie bijlage 1E): llresult1, llresult2, llresult3, llinvul3 en llinvul4.

Tabel 4.14 Passingsmaten, correlaties tussen de schalen en Cronbach's alpha voor het oorspronkelijke vier factoren meetmodel van schoolbeleving en voor het nieuwe modellen met drie factoren bij een confirmatieve factoranalyse (n=3760)

Meetmodel	Oorspronkelijk	Nieuw
Passingswaarden	Vier factoren model	Drie factoren model
χ^2 (df)	13183.18 (696)	8388.99 (626)
RMSEA	.069	.057
90% C.I. RMSEA	.068 - .070	.056 - .059
SRMR	.059	.048
CFI	.92	.95
Model AIC	13351	8542
Correlatie		
Tuisvoelen-Waardering	.80	Waardering school-Waardering onderwijs .45
Tuisvoelen-Kwaliteit instructie	.87	
Tuisvoelen-Family Educational Culture	.38	Waardering school-Werkklimaat .62
Waardering-Kwaliteit instructie	.66	
Waardering-Family Educational Culture	.44	Waardering onderwijs-Werkklimaat .54
Kwaliteit instructie-Family Educ. Culture	.43	
Cronbach's alpha		
Totaal model (42 items)	.90	Totaal model (37 items) .89
Tuisvoelen (10 items)	.78	Waardering school (10 items) .83
Waardering (10 items)	.77	Waardering onderwijs (10 items) .70
Kwaliteit instructie (17 items)	.81	Werkklimaat (17 items) .82
Family Educational Culture (5 items)	.59	

RMSEA = Root Mean Square Error of Approximation; 90% C.I. = 90% betrouwbaarheidsinterval; CFI = Comparative Fit Index; SRMR = Standardized Root Mean square Residual; AIC = Akaike Information Criterion.

Parameterschattingen voor het nieuwe meetmodel zijn opgenomen in bijlage 5C. In het model laden alle indicatoren significant en redelijk hoog op de bijbehorende factoren, met uitzondering van de items in de subschaal 'Family Educational Culture', die matig laden. De betekenis van de factorladingen hangt af van de steekproefgrootte (Stevens, 1992, in Field, 2000, p. 440). In grote steekproeven zijn ook kleinere factorladingen nog van betekenis. Volgens Stevens (1992) dient bij een steekproef van 100 de factorlading groter te zijn dan .512 en bij een steekproef van 1000 groter dan .162. Aan dit criterium voldoen alle factorladingen ruimschoots. Uit de uitkomsten van de confirmatieve factor analyse kan worden opgemaakt dat het drie factoren model bruikbaar is voor vervolganalyses.

4.6 Analyse van de multitrait-multimethod matrix

De vragenlijsten voor het meten van de handelingen van schoolleiders, schoolcultuur en schoolorganisatie bevatten elk vier schalen die corresponderen met de vier kwadranten in het concurrerende waarden model: rationele doel (RD), interne proces (IP), 'human relations' (HR) en open systeem (OS). Om de convergerende en discriminante validiteit van de onderzoeksinstrumenten nader vast te stellen is een *multitrait-multimethod* matrix opgesteld van de correlaties tussen de schalen uit de verschillende onderzoeksinstrumenten. Daartoe werden eerst met behulp van LISREL latente variabele scores berekend voor elk meetmodel (Mels, 2004, p.22). De latente variabele scores (factorscores) voor de docenten werden geaggregeerd, zodat scores op schoolniveau werden verkregen (zie ook 4.3.6). Op basis van

deze geaggregeerde scores zijn de correlaties berekend tussen de verschillende schalen. Voor de handelingen van schoolleiders werden de correlaties berekend aan de hand van schoolleidersgegevens en docentengegevens, voor de schoolcultuur en de schoolorganisatiepraktijk waren alleen docentengegevens beschikbaar. De multitrait-multimethod matrix is afgebeeld in tabel 4.15.

Tabel 4.15 *Multitrait-multimethod matrix van significante correlaties ($\alpha = .01$) tussen de scores op de meetschalen voor schoolleidersgedrag, schoolcultuur en schoolorganisatiepraktijk, met op de diagonaal waarden van Cronbach's alpha*

Schalen	Schoolleiders-handelingen				Schoolleiders-handelingen				Schoolcultuur				Schoolorganisatiepraktijk			
	<i>schoolleiders</i>				<i>docenten</i>				<i>docenten</i>				<i>docenten</i>			
	RD	IP	HR	OS	RD	IP	HR	OS	RD	IP	HR	OS	RD	IP	HR	OS
<i>Schoolleiders</i>																
RDhandelingen	.67															
IPhandelingen		.70														
HRhandelingen			0.56	.77												
OShandelingen	0.34			.76												
<i>Docenten</i>																
RDhandelingen	0.40				.85											
IPhandelingen		0.40			0.61	.76										
HRhandelingen			0.27	0.37	0.77	0.74	.84									
OShandelingen	0.26			0.61	0.56	0.28	.73									
RDcultuur									.80							
IPcultuur					0.44	0.34		0.36	0.61	.80						
HRcultuur				0.26	0.47		0.29	0.51		0.58	.86					
OScultuur				0.26	0.39			0.52		0.42	0.80	.78				
RDorganisatie					0.42	0.50	0.30	0.37	0.68	0.72	0.37		.74			
IPorganisatie	0.31			0.28	0.67	0.29	0.40	0.55		0.65	0.70	0.67	0.52	.80		
HRorganisatie	0.28			0.27	0.67	0.33	0.42	0.54		0.53	0.70	0.62	0.49	0.83	.78	
OSorganisatie				0.29	0.50	0.31	0.31	0.58		0.41	0.40	0.45	0.47	0.51	0.47	.71

RD = rationele doel schaal, IP= interne proces schaal, HR = 'human relations' schaal, OS = open systeem schaal;

In tabel 4.15 is te zien dat er een groot aantal significante correlaties ($\alpha = .01$) bestaat tussen de scores op de schalen in de verschillende meetmodellen. De cursieve getallen op de diagonaal van de matrix zijn de waarden van Cronbach's alpha voor de schalen. Cronbach's alpha is een maat voor de betrouwbaarheid, die is gebaseerd op de gemiddelde correlatie tussen de items. De waarden op de diagonaal zouden hoger moeten zijn dan de andere waarden in de matrix, omdat elke schaal meer met zichzelf zou moeten correleren dan met andere schalen (Trochim, 2006). Dit klopt niet voor alle schalen in tabel 4.15, omdat enkele schalen een vrij lage alpha hebben, terwijl er buiten de diagonalen hoge correlaties voorkomen. De hoge correlaties tussen sommige schaalcores komen hieronder nog aan de orde.

Het gedrag van de schoolleiders werd zowel door de schoolleiders zelf beoordeeld, als door de docenten. Dit maakt het mogelijk om de convergerende validiteit te bepalen voor de onderzoeksinstrument. In het grijze blok linksboven in tabel 4.15 is op de diagonaal te zien dat er voor rationele doel, interne proces, 'human relations' en open systeem handelingen een significante correlatie bestaat tussen de zelfbeoordeling van de schoolleiders en de waarneming van de docenten. De correlaties op de diagonaal zijn bovendien hoger dan de andere correlaties in het blok. Dit wijst op convergerende validiteit. In de driehoeken boven

(zelfbeoordeling van schoolleiders) en naast het blok (waarneming van docenten) komen echter nog hogere correlaties voor. Verschillende handelingen die worden gemeten met dezelfde methode vertonen dus een groter verband, dan dezelfde handelingen die met verschillende methoden worden gemeten. Dat betekent dat de meetmethode een belangrijke invloed heeft op de uitkomsten (Trochim, 2006).

Uit de bovenste twee driehoeken onder de diagonaal is op te maken dat er voor de docentengegevens een behoorlijke correlatie is tussen de interne proces schaal en de 'human relations' schaal. Bij de docenten zijn er tevens vrij hoge correlaties tussen de rationele doel schaal en de andere drie schalen en tussen de interne proces en 'human relations' schaal. De correlaties in deze driehoeken stemmen in orde van grootte redelijk overeen met de schattingen van de (éénniveau) confirmatieve factoranalyse in tabel 4.10. Het patroon van de correlaties voor de schoolleiders wijkt af van het patroon voor de docenten. Dat wijst op een gebrek aan convergerende validiteit. Alleen tussen rationele doel en open systeem handelingen en tussen 'human relations' en interne proces handelingen is de correlatie min of meer consistent voor docenten en schoolleiders.

De onderste twee driehoeken onder de diagonaal geven de correlaties voor de cultuurschalen onderling en de organisatieschalen onderling. Bij de cultuurschalen correleren de interne proces, 'human relations' en open systeem schaal met elkaar, terwijl de rationele doel schaal alleen correleert met de interne proces schaal. De correlatie tussen 'human relations' en open systeem is hoog. De organisatieschalen correleren allemaal met elkaar, waarbij de correlatie tussen interne proces en 'human relations' het hoogst is.

Bij de analyse van de uitkomsten van de confirmatieve factoranalyses in de voorgaande paragrafen werd al geconstateerd dat de correlaties tussen de vier factoren binnen de meetmodellen voor schoolleidersgedrag, de schoolcultuur en de schoolorganisatie, niet in overeenstemming zijn met de hypothese, dat factoren die aangrenzende kwadranten in het concurrerende waarden model vertegenwoordigen, meer zouden moeten correleren, dan factoren die tegenoverliggende kwadranten representeren (zie paragraaf 4.2). Binnen het concurrerende waarden model is er tussen de kwadranten wel een zekere conceptuele overlap, maar de correlaties die optreden tussen de schalen kunnen bijvoorbeeld ook afhangen van de respondentengroep (Lamond, 2003), van de ontwikkelingsfase van de organisatie (Quinn, 1988) of van actuele maatschappelijke omstandigheden (Kalliath, Bluedorn & Gillespie, 1999). De correlaties die zijn gevonden tussen de schalen onderling voor respectievelijk schoolleidersgedrag, schoolcultuur en schoolorganisatie wijzen dus niet per se op een gebrekkige discriminante validiteit, behalve voor de hoogste waarden.

Nu is het nog de vraag in hoeverre de schalen voor schoolleidersgedrag, de schoolcultuur en de schoolorganisatie van elkaar verschillen. Worden deze concepten, die allemaal zijn gebaseerd op het concurrerende waarden model, voldoende van elkaar onderscheiden? In tabel 4.14 is te zien (in de twee blokken linksonder) dat er slechts lage correlaties zijn tussen de factorscores voor schoolleidersgedrag die zijn gebaseerd op de zelfbeoordeling van schoolleiders en de cultuur- en organisatieschalen. Tussen de factorscores voor schoolleidersgedrag die zijn gebaseerd op de waarnemingen van docenten en de cultuur- en organisatieschalen zijn er meer en hogere correlaties, zowel voor schalen die hetzelfde kwadrant vertegenwoordigen, als voor andere. De volgende vraag die dan gesteld kan worden, is in hoeverre de correlaties tussen de schalen zouden kunnen wijzen op een oorzakelijk verband tussen schoolleidersgedrag en de schoolcultuur en –organisatie of op conceptuele overlap ('begripsmatige inclusie'). Het afkapcriterium voor discriminante validiteit is vrij

hoog: de correlatie tussen twee schalen moet kleiner zijn dan .85 (Garson, z.j., d; Kenny, 1998). Op basis van dit criterium is er voldoende onderscheid tussen de schalen voor de verschillende concepten. Het risico van begripsmatige inclusie is wellicht het kleinst wanneer voor het meten van schoolleidersgedrag wordt uitgegaan van de zelfbeoordeling van schoolleiders, omdat de factorscores voor schoolcultuur en schoolorganisatie zijn gebaseerd op docentengegevens. Daarbij moet worden opgemerkt dat de perceptie van de schoolleider van bepaalde activiteiten kan verschillen van die van docenten. Zowel schoolleiders als docenten kunnen sociaal wenselijke antwoorden gegeven hebben. Verder zullen niet alle docenten evenveel zicht hebben op de activiteiten van de schoolleider. In paragraaf 4.7 wordt besproken welke verbanden bestaan tussen de schoolleiderscores en de docentenscores voor de handelingen van schoolleiders.

Tenslotte dient te worden vastgesteld of er voldoende onderscheid is tussen de metingen van de schoolcultuur en de schoolorganisatiepraktijk. Dit is te zien in het grijze blok rechtsonder in tabel 4.15. Er is een aanzienlijke correlatie tussen de rationele doel schaal voor schoolorganisatie en de rationele doel en interne proces schaal voor schoolcultuur. De interne proces schaal en de 'human relations' schaal voor schoolorganisatie en in minder mate ook de open systeem schaal, vertonen een duidelijk verband met de interne proces, 'human relations' en open systeem schaal voor schoolcultuur. Hoewel de correlaties lager zijn dan het afkapcriterium voor discriminante validiteit (.85), is het hier vanwege de relatief hoge waarden toch de vraag of de concepten voldoende van elkaar worden onderscheiden. Daarbij zijn ook de correlaties tussen de schoolcultuurschalen onderling en de schoolorganisatieschalen onderling van belang. De verbanden tussen de schoolorganisatie- en schoolcultuurvariabelen zijn daarom nader onderzocht. De uitkomsten worden beschreven in de volgende paragraaf.

In paragraaf 4.7 komt eerst de betrouwbaarheid van de geaggregeerde variabelen aan de orde. Daarna worden de verbanden tussen de zelfbeoordeling van schoolleiders en de waarneming van docenten van schoolleidersgedrag besproken en tenslotte de verbanden tussen de schoolcultuur- en schoolorganisatievariabelen.

4.7 Betrouwbaarheid van de geaggregeerde variabelen en verbanden tussen de variabelen

Voor de handelingen van schoolleiders, de schoolcultuur, de schoolorganisatiepraktijk en de schoolbeleving van leerlingen werden latente variabele scores bepaald aan de hand van docenten- of leerlingengegevens, met behulp van éénniveau confirmatieve factoranalyse. Deze scores werden vervolgens geaggregeerd naar schoolniveau (zie 4.3.6). De betrouwbaarheid λ_j van de geaggregeerde variabelen op schoolniveau kan worden vastgesteld aan de hand van de gemiddelde groepsgrootte n_j (het gemiddelde aantal docenten of leerlingen per school) en de intraklasse correlatie ρ_I . Hiervoor geldt de formule (zie 3.5.4):

$$\lambda_j = n_j \rho_I / (1 + [n_j - 1] \rho_I)$$

In tabel 4.16 wordt wordt een overzicht gegeven van de intraklassecorrelatie en de betrouwbaarheid voor de geaggregeerde variabelen.

Tabel 4.16 Intraklassecorrelatie, gemiddelde groeps grootte en betrouwbaarheid van de geaggregeerde variabelen op schoolniveau

Geaggregeerde variabele	ρ_I	n_i^1	Reliability (λ_i)
Rationele doel handelingen	.22	9.0	0.72
Interne proces handelingen	.34	9.0	0.82
Human relations handelingen	.31	9.0	0.80
Open systeem handelingen	.22	9.0	0.72
Rationele doel schoolcultuur	.17	9.7	0.67
Interne proces schoolcultuur	.11	9.7	0.55
Human relations schoolcultuur	.15	9.7	0.63
Open systeem schoolcultuur	.22	9.7	0.73
Rationele doel organisatiepraktijk	.18	9.5	0.68
Interne proces organisatiepraktijk	.21	9.5	0.72
Human relations organisatiepraktijk	.27	9.5	0.78
Open systeem organisatiepraktijk	.17	9.5	0.66
Waardering school	.07	38	0.74
Waardering onderwijs	.02	38	0.44
Werkklimaat	.09	38	0.79

¹ (harmonisch) gemiddelde van het aantal respondenten per school, gebaseerd op het aantal geldige waarnemingen

De betrouwbaarheid van de geaggregeerde variabelen is acceptabel, alleen bij de interne proces schaal voor schoolcultuur en de waardering van onderwijs door leerlingen is de betrouwbaarheid laag. De lage waarde voor de gemiddelde waardering van onderwijs in het algemeen ligt voor de hand, omdat het hier meer gaat om persoonlijke kenmerken van leerlingen dan om schoolkenmerken.

Uit de multitrait-multimethod matrix in tabel 4.15 kan worden opgemaakt dat de correlaties tussen de zelfbeoordeling van schoolleiders en de waarnemingen van docenten van schoolleidersgedrag significant zijn. Dit wijst op convergerende validiteit. Aan de hand van de correlaties kan worden geschat in hoeverre de docentscores voor de handelingen van schoolleiders zouden kunnen worden voorspeld uit de schoolleiderscores. Het kwadraat van de correlatiecoëfficiënt is een maat voor het gedeelte van de variantie in een variabele dat wordt verklaard door een andere variabele (Field, 2000, p.90). Met behulp van meerniveau regressie analyse kan echter nog nauwkeuriger worden bepaald in hoeverre de variantie op schoolniveau in de docentenscores wordt verklaard door de schoolleiderscores. In tabel 4.17 wordt een overzicht gegeven van de verklaarde variantie in de docentenscores op basis van de correlaties en op basis van meerniveau analyse.

Tabel 4.17 Correlaties tussen schoolleiderscores en docentenscores en verklaarde variantie in de docentenscores voor de handelingen van schoolleiders

Schaal	Correlatie r	Verklaarde variantie ¹ R ²	Verklaarde variantie ² schoolniveau	Verklaarde variantie ² totaal
Rationele doel handelingen	0.40	16%	20%	4%
Interne proces handelingen	0.40	16%	20%	7%
Human relations handelingen	0.37	14%	20%	6%
Open systeem handelingen	0.61	37%	53%	12%

¹ Op basis van correlaties (op schoolniveau)

² Op basis van meerniveau regressie analyse

Uit tabel 4.17 blijkt dat de verklaarde variantie in de docentscores voor de handelingen van schoolleiders wordt onderschat als wordt uitgegaan van de correlatie tussen de schoolleiderscores en de (geaggregeerde) docentscores. Het aggregeren van variabelen betekent een verlies aan variantie waardoor minder samenhang kan worden gevonden (zie Meelissen, 2005, p.88). Op basis van meerniveau regressie analyse worden hogere percentages verklaarde variantie gevonden op schoolniveau. Het effect van de schoolleiderscores is voor alle handelingen significant. Voor de rationele doel, interne proces en 'human relations' handelingen is het percentage van de variantie op schoolniveau dat wordt verklaard door de schoolleiderscores vrijwel gelijk. Het percentage verklaarde variantie voor de open systeem handelingen is erg hoog. De schoolleiderscores verklaren slechts een klein deel van de totale variantie in de docentscores voor de handelingen van schoolleiders. Dat komt doordat slechts een beperkt deel van de variantie in de docentscores is toe te schrijven aan verschillen tussen scholen. In hoofdstuk 5 wordt nader ingegaan op de verschillen tussen schoolleiderscores en docentscores voor schoolleidersgedrag.

De correlaties tussen de schoolcultuur schalen die werden gevonden in dit onderzoek (zie tabel 4.15) kunnen worden vergeleken met de uitkomsten van ander onderzoek. Maslowski (2001, p.81) vond bij de validering van de vragenlijst die hij ontwikkelde voor het meten van de schoolcultuur, met behulp van confirmatieve factoranalyse een redelijke correlatie (.54) tussen rationele doel en interne proces cultuur en daarnaast een bescheiden correlatie (.35) tussen de interne proces en 'human relations' cultuur. In het hier gerapporteerde onderzoek werden hogere waarden gevonden voor de correlaties tussen rationele doel en interne proces cultuur (.61) en tussen interne proces en 'human relations' cultuur (.58), terwijl er ook substantiële verbanden bestaan tussen open systeem cultuur en interne proces cultuur (.42) en tussen open systeem cultuur en 'human relations' cultuur (.80). Een verband tussen de rationele doel en interne proces oriëntaties wordt vaker gevonden bij onderzoek naar het concurrerende waarden model (Quinn & Rohrbaugh, 1983; Van Muijen, 1994). Tussen de andere oriëntaties worden wisselende correlaties gevonden (zie paragraaf 4.2, zie ook Maslowski, 2001, p.79). Zoals ook in paragraaf 4.6 werd geconcludeerd is er enige conceptuele overlap tussen de kwadranten in het concurrerende waarden model en kunnen de correlaties tussen factoren die de kwadranten vertegenwoordigen bijvoorbeeld afhangen van de ontwikkelingsfase van de organisatie, de respondentengroep of actuele maatschappelijke omstandigheden.

Vanwege de relatief hoge correlaties die voorkomen tussen de schoolcultuur- en schoolorganisatieschalen (zie paragraaf 4.6) is het niet zeker of de schalen binnen de meetinstrumenten voldoende discrimineren en of de schoolcultuur en de schoolorganisatiepraktijk in dit onderzoek als onafhankelijke concepten kunnen worden beschouwd. Bij de samenstelling van de vragenlijst over de schoolorganisatiepraktijk werd al geconstateerd dat de schoolorganisatie, zoals die werd geoperationaliseerd, kon worden opgevat als een neerslag van de schoolcultuur in de schoolpraktijk (zie paragraaf 3.2). Om na te gaan in hoeverre de schalen voor schoolorganisatie en schoolcultuur inderdaad verschillende factoren vertegenwoordigen, werd een principale componenten analyse uitgevoerd (met varimax rotatie). Uit deze analyse komen twee factoren naar voren die samen 82% van de variantie verklaren. De uitkomsten staan in tabel 4.18.

Tabel 4.18 *Uitkomsten van een principale componenten analyse (met varimax rotatie) voor schoolcultuur en schoolorganisatie schalen*

Schaal	Factor 1*	Factor 2*
Rationele doel schoolcultuur		.93
Interne proces schoolcultuur	.50	.75
Human relations schoolcultuur	.87	
Open systeem schoolcultuur	.88	
Rationele doel organisatiepraktijk	.34	.86
Interne proces organisatiepraktijk	.87	
Human relations organisatiepraktijk	.85	
Open systeem organisatiepraktijk	.58	.31

*Waarden kleiner dan .03 zijn niet vermeld.

Op basis van de uitkomsten van de exploratieve factoranalyse in tabel 4.18, werd besloten om voor schoolcultuur en schoolorganisatie, in plaats van acht variabelen, twee factoren te onderscheiden: één factor die wordt bepaald door de som van de scores voor rationele doel schoolcultuur, interne proces schoolcultuur en rationele doel organisatiepraktijk en één factor die wordt bepaald door de som van de scores op 'human relations' schoolcultuur, open systeem schoolcultuur, interne proces organisatiepraktijk en 'human relations' organisatiepraktijk. De open systeem organisatiepraktijk is niet meegenomen, omdat deze variabele minder hoog laadt op de eerste factor dan de andere. De eerste factor wordt 'prestatiegerichtheid' genoemd. Deze factor vertegenwoordigt een gerichtheid op prestaties en een ordelijke werksfeer. De tweede factor wordt 'ontwikkelingsgerichtheid' genoemd. Deze factor is een maat voor de gerichtheid op samenwerking, professionele ontwikkeling, aanpassing en vernieuwing. In de tweede factor zijn kenmerken van de 'lerende organisatie' te herkennen (zie paragraaf 2.3). De correlatie tussen deze twee factoren is .41, dit is een acceptabele waarde in verband met discriminante validiteit. Tabel 4.19 geeft een overzicht van de intraklassecorrelatie en de betrouwbaarheid van de nieuwe variabelen.

Tabel 4.19 *Intraklassecorrelatie, gemiddelde groeps grootte en betrouwbaarheid van de geaggregeerde variabelen op schoolniveau*

variabele	ρ_i	n_j	reliability (λ_j)
Prestatiegerichtheid	.16	9.5	0.64
Ontwikkelingsgerichtheid	.24	9.5	0.75

De betrouwbaarheid van de variabele 'prestatiegerichtheid' is nog aan de lage kant, die voor 'ontwikkelingsgerichtheid' is goed. De nieuwe variabelen 'prestatiegerichtheid' en 'ontwikkelingsgerichtheid' worden in hoofdstuk 5 nader beschreven.

4.8 Discussie

De eerste onderzoeksvraag bij dit hoofdstuk heeft betrekking op de psychometrische kwaliteit van de onderzoeksinstrumenten voor het meten van schoolleidersgedrag, de schoolcultuur, de schoolorganisatiepraktijk en de schoolbeleving van leerlingen. De validiteit en de betrouwbaarheid van de onderzoeksinstrumenten werden onderzocht met behulp van éénniveau en meerniveau confirmatieve factoranalyse (zie paragrafen 4.3 t/m 4.5).

Bij de uitvoering van zowel éénniveau (voor schoolleidersdata) als meerniveau confirmatieve factoranalyse met behulp van LISREL (versie 8.8) bleken problemen op te treden door de

omvang van de meetmodellen (zie 4.3.6). Bij de opzet van het onderzoek werd uitgegaan van een steekproef van 100 scholen met gemiddeld tien docenten per school, omdat dit zou voldoen voor meerniveau analyse (zie bijv. Muthén, 1997, Hox, 2002, p.175). Voor *structural equation modeling* gelden echter andere vuistregels. Garson (z.j., b) stelt bijvoorbeeld dat bij een model met meer dan tien variabelen een steekproefgrootte van minder dan 200 cases leidt tot onstabiele resultaten. Ander auteurs gaan uit van tien tot twintig keer zoveel cases als variabelen (Mitchell, 1993, Stevens, 1996). Aangezien alle onderzoeksinstrumenten tenminste 40 items bevatten, was de steekproefgrootte van 97 scholen dus eigenlijk ontoereikend om de opgestelde meetmodellen te kunnen toetsen.

Voor de meerniveau confirmatieve factoranalyses werden twee verschillende schattingsmethoden gebruikt: *Muthén's approximate Maximum Likelihood* (MUML, Muthén, 1990) en *Full Information Maximum Likelihood* (FIML, du Toit, & du Toit, 2001) methode. MUML gaat uit van twee (onafhankelijke) groepen op schoolniveau en docentenniveau, terwijl FIML op schoolniveau meer groepen analyseert (zoveel als het aantal verschillende groeps groottes, zie Muthén, 1990). Hoewel FIML accuratere schattingen zou moeten opleveren, bleek MUML robuuster te zijn en betere resultaten op te leveren voor de omvangrijke meermodellen die werden getest. Het is niet mogelijk om met behulp van LISREL (versie 8.8) latente variabele scores te berekenen voor meerniveau meetmodellen, daarom werden deze scores berekend aan de hand van éénniveau modellen en geaggregeerd naar schoolniveau.

Een voordeel van de toepassing van meerniveau confirmatieve factoranalyse was dat een beter inzicht werd verkregen in de passing van het model op schoolniveau, terwijl ook zwakke items op schoolniveau konden worden gedetecteerd. Hierdoor konden de meetmodellen beter worden toegesneden op de verschillen tussen scholen. Vanwege de omvang van de meetmodellen waren de databestanden te klein om kruisvalidering te kunnen toepassen. Door de 'datagedreven' verbetering van de meetmodellen is kanskapitalisatie niet uit te sluiten (zie MacCallum, Roznowski en Necowitz, 1992). Voor de handelingen van schoolleiders is de modelverbetering echter uitgevoerd aan de hand van twee verschillende databestanden (van schoolleiders en docenten). Verder is bij de verbetering van de meetmodellen een behoudende aanpak gehanteerd, waarbij alleen slecht passende items (op schoolniveau) werden verwijderd of verplaatst. Hierdoor is het risico van kanskapitalisatie waarschijnlijk toch beperkt.

Door het aanpassen van de meetmodellen werd de *constructvaliditeit* verbeterd. Constructvaliditeit is onder te verdelen in convergerende en discriminante validiteit. Verschillende instrumenten om hetzelfde begrip te meten, moeten hetzelfde resultaat opleveren, ofwel convergeren. Als met een meetinstrument verschillende begrippen worden gemeten, moet de methode deze begrippen voldoende van elkaar onderscheiden (discrimineren). Uit de correlaties tussen de factoren in het meetmodel voor schoolbeleving blijkt dat de discriminante validiteit van dit instrument voldoende is (zie paragraaf 4.5). De convergerende en discriminante validiteit van de meetinstrumenten voor de handelingen van schoolleiders, de schoolcultuur en de schoolorganisatie werd nader geanalyseerd aan de hand van een multitrait-multimethod matrix van correlaties tussen factorscores (zie paragraaf 4.6). Hiertoe werden eerst latente variabele scores bepaald voor alle schalen, die (voor docenten- en leerlingengegevens) werden geaggregeerd naar schoolniveau. De betrouwbaarheid van de geaggregeerde scores voor schoolleidersgedrag, de schoolcultuur de schoolorganisatie en de schoolbeleving van leerlingen op schoolniveau is acceptabel (zie paragraaf 4.7).

Uit de multitrait-multimethod matrix kon worden opgemaakt dat de convergerende validiteit voor schoolleidersgedrag redelijk is, maar ook dat schoolleiders en docenten (deels) verschillende percepties hebben van de uitgevoerde handelingen (zie paragraaf 4.6). De scores van docenten worden slechts voor een beperkt deel voorspeld door de scores van schoolleiders (zie ook paragraaf 4.7). Verder correleren de handelingen zoals die door docenten werden waargenomen, sterker met schoolcultuur- en schoolorganisatievariabelen, dan de handelingen zoals die door schoolleiders zelf werden opgegeven. Vanwege de verschillen tussen de schoolleidersdata en de docentendata voor schoolleidersgedrag, is ervoor gekozen om bij de toetsing van het onderzoeksmodel de gegevens van beide groepen respondenten mee te nemen (zie hoofdstuk 6).

Tussen de schoolcultuur- en de schoolorganisatievariabelen onderling blijken aanzienlijke relaties te bestaan. Dit zou kunnen wijzen op een gebrek aan discriminante validiteit, hoewel werd voldaan aan het criterium dat de correlatie tussen twee variabelen niet hoger mag zijn dan .85 (Garson, z.j., d; Kenny, 1998). Uit een principale componenten analyse van de acht variabelen voor schoolcultuur en schoolorganisatiepraktijk bleek echter, dat twee componenten kunnen worden onderscheiden, die samen 82% van de variantie verklaren (zie paragraaf 4.7). Aan de hand van deze exploratieve factoranalyse zijn twee nieuwe variabelen geconstrueerd: 'prestatiegerichtheid' en 'ontwikkelingsgerichtheid'. De variabele 'prestatiegerichtheid' geeft aan in hoeverre een school is gericht op prestaties en een ordelijke werksfeer. De variabele 'ontwikkelingsgerichtheid' geeft aan in hoeverre de school is gericht op samenwerking binnen de organisatie, professionele ontwikkeling, aanpassing en vernieuwing.

Met de vaststelling van de twee nieuwe variabelen is de tweede onderzoeksvraag voor dit hoofdstuk beantwoord. Het concurrerende waarden model vormt een geschikt kader om de handelingen van schoolleiders te karakteriseren, maar voor de schoolcultuur en de schoolorganisatiepraktijk is er teveel overlap tussen de factoren die de vier kwadranten in het model vertegenwoordigen. In het onderzoeksmodel (zie figuur 2.5) worden daarom de variabelen schoolcultuur en schoolorganisatiepraktijk vervangen door de twee nieuwe variabelen 'prestatiegerichtheid' en 'ontwikkelingsgerichtheid'.

In het volgende hoofdstuk worden de kenmerken van scholen voor voortgezet onderwijs in Nederland beschreven aan de hand van de variabelen in het onderzoeksmodel. Vervolgens worden de relaties tussen deze kenmerken en de uitkomstvariabelen geëxploreerd. De toetsing van het gehele causale model wordt besproken in hoofdstuk 6.

Hoofdstuk 5

Kenmerken van scholen in relatie tot de opbrengsten

5.1 Inleiding

Het concurrerende waarden model is ontwikkeld voor organisatieonderzoek, om de cultuur en de effectiviteit van organisaties te bepalen (Quinn, 1998). Daarnaast wordt het ook gebruikt voor het verbeteren van managementvaardigheden. Bij het model zijn acht rollen gedefinieerd die een manager moet kunnen vervullen: bij het rationele doel model horen de rollen van bestuurder en producent, bij het interne proces model die van controleur en coördinator, bij het ‘human relations’ model de rollen van mentor en stimulator en bij het open systeem model de rollen van bemiddelaar en innovator (zie tabel 2.2). Door Quinn, Faerman, Thompson en McGrath (2003) is een methode opgezet, waarmee managers de vaardigheden die bij deze rollen horen, kunnen ontwikkelen. Hoewel het concurrerende waarden model in bedrijfskundig onderzoek en in managementopleidingen veel wordt toegepast, is het nog maar weinig ingezet in onderzoek naar effectief schoolleiderschap. In paragraaf 5.2 worden enkele voorbeelden van dit type onderzoek beschreven.

Het concurrerende waarden model is in dit onderzoek als uitgangspunt gebruikt bij de constructie van instrumenten voor het meten van de handelingen van schoolleiders, de schoolcultuur en de schoolorganisatiepraktijk (zie paragraaf 3.2). Om het verband tussen schoolleiderschap en de leerprestaties in kaart te kunnen brengen, is daarnaast ook nog een instrument ontwikkeld voor het meten van de schoolbeleving van leerlingen. In het voorgaande hoofdstuk is vastgesteld welke items en schalen in principe geschikt zijn om verschillen tussen scholen aan het licht te brengen. Aan de hand van deze schalen worden in dit hoofdstuk enige beschrijvende gegevens van scholen voor voortgezet onderwijs in Nederland gepresenteerd. Daarbij staat de volgende vraag centraal:

Wat zijn de kenmerken van scholen voor voortgezet onderwijs in Nederland met betrekking tot de handelingen die de schoolleider verricht, de schoolcultuur, de schoolorganisatiepraktijk en de schoolbeleving van leerlingen?

Nadat deze vraag is beantwoord, komt de derde onderzoeksvraag (zie hoofdstuk 1) aan de orde:

Zijn er (directe) relaties tussen de kenmerken van scholen voor voortgezet onderwijs in Nederland en de opbrengsten?

In paragraaf 5.3 worden eerst de schalen voor het meten van de handelingen van schoolleiders, de schoolcultuur, de schoolorganisatiepraktijk en de schoolbeleving van leerlingen kort toegelicht. Daarna volgt een overzicht van de kenmerken van scholen voor

voortgezet onderwijs in Nederland. De variabelen *prestatiegerichtheid* en *ontwikkelingsgerichtheid*, die zijn voortgekomen uit een factoranalyse van de schoolcultuur en schoolorganisatiepraktijk schalen, worden in paragraaf 5.4 nader onderzocht. In paragraaf 5.5 worden de ‘directe’ relaties tussen de schoolkenmerken en de uitkomstvariabelen geëxploreerd met behulp van Pearson correlatiecoëfficiënten. Paragraaf 5.6 bevat een discussie van de resultaten in dit hoofdstuk.

5.2 Het concurrerende waarden model in onderwijsonderzoek

In onderwijsonderzoek is het concurrerende waarden model nog niet op uitgebreide schaal toegepast, maar de bruikbaarheid wordt wel onderkend. Houtveen, Voogt, van der Vegt en van de Grift (1996; zie ook Van Wieringen, Ax, Karstanje & Voogt, 2000) laten zien, hoe de vier organisatiemodellen van Quinn en Rohrbaugh (1983) terug zijn te vinden in vier categorieën van effectiviteitsopvattingen in het onderwijs, die achtereenvolgens zijn opgekomen (vgl. Schaveling, 1998; Scheerens, 1989, 1992; Voogt, 1989, 1994; Wijnen & Wolfsen, 1987). De onderstaande beschrijving van deze vier categorieën is ontleend aan het overzicht van Houtveen e.a. (1996, vanaf p.38).

De eerste categorie effectiviteitsopvattingen, die sinds de jaren vijftig en zestig waargenomen kan worden, is gericht op beheersing van de organisatie via de toepassing van regelgeving. De ‘traditionele school’ kent een min of meer losse eilanden structuur met autonome docenten. Regelgeving zorgt in deze situatie voor een stabiele context: de regels zijn bekend (rooster, leerstof, bevoegdheden, randvoorwaarden) en onderwijsinhoudelijk is de docent autonoom. Nieuwe docenten leren het vak van ervaren docenten, waardoor de continuïteit is gewaarborgd. Het goed beheren van de school wordt gezien als de belangrijkste taak van de schoolleider. In deze opvatting van organisatie effectiviteit is het *interne proces* model van Quinn en Rohrbaugh te herkennen, met de schoolleider als coördinator en controleur. Vanwege de autonomie van docenten, zijn er binnen de school verschillende subculturen, waardoor het moeilijk is voor de school om zich te profileren of te veranderen.

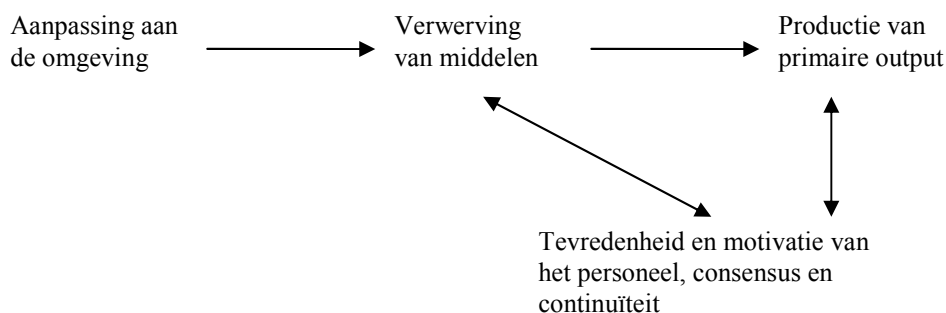
In de zeventiger jaren kwam in het onderwijs een trend op gang, die is gericht op een ‘gezonde organisatie’. Daarbij staat het welbevinden van docenten en leerlingen centraal. Er wordt gesproken van een ‘collegiale school’. Door sociale interactie en informele contacten zou de saamhorigheid bevorderd moeten worden. *Bottom up* processen moeten de betrokkenheid en de inzet van docenten verhogen. Voor het welbevinden en de motivatie van leerlingen wordt ernaar gestreefd om de leerstof en de werkvormen aan te passen aan het niveau, de interesses of de omgeving van de leerling. De docent is een professional, die zelf zorg moet dragen voor zijn professionele ontwikkeling. De schoolleider verleent ondersteuning als ‘primus inter pares’. Het doel van de organisatie is om de capaciteiten van het personeel zo ver mogelijk te ontwikkelen en zo adequaat mogelijk in te zetten. Deze opvatting van organisatie effectiviteit is terug te vinden in het ‘*human relations*’ model van Quinn en Rohrbaugh. In dit model gaat het om inzet, samenhang en moreel. De cultuur is homogeen: collectiviteit en cohesie worden gestimuleerd. De schoolleider vervult de rollen van mentor en stimulator.

In de jaren tachtig kwam een volgende trend op gang, waarbij het gaat om planmatig handelen en het stellen van doelen, om zo hoog mogelijke leerprestaties te bereiken. Het organisatiemodel dat hierbij past is het *rationele doel* model. In dit model staan productiviteit en winst centraal. In het onderwijs heeft zich dit vertaald in aandacht voor ‘effectieve instructie’ en ‘maximale leertijd’. Ook het toetsen van leerprestaties, opgevat als gerealiseerde

onderwijsdoelstellingen, bijvoorbeeld door een eindtoets of een leerlingvolgsysteem, past in deze opvatting van organisatie effectiviteit. Het schoolklimaat moet ordelijk en veilig zijn en gericht op leren. In de ‘effectieve school’ vervult de schoolleider de rollen van ‘bestuurder’ en ‘producent’ en houdt zich niet alleen bezig met het secundaire proces, maar ook met het primaire proces. De cultuur is homogeen en zakelijk.

De vierde categorie effectiviteitsopvattingen kwam op in de negentiger jaren. Deze categorie sluit aan bij het *open systeem* model van Quinn en Rohrbaugh. De ‘bedrijvige school’ is gericht op flexibilisering, om adequaat te kunnen reageren op de omgeving en uiteindelijk een zo groot mogelijk ‘marktaandeel’ te verwerven. Het marktaandeel kan voor scholen worden uitgedrukt in het aantal leerlingen of het aandeel in de beschikbare voorzieningen. Een dergelijk type school maakt bijvoorbeeld gebruik van sponsoring en contractonderwijs en is gericht op profilering en publiciteit. Vaak is er een ruim aanbod aan buitenschoolse activiteiten. De cultuur is innovatief, de schoolleider vervult de rollen van innovator en bemiddelaar. In deze situatie is een voortdurende professionele ontwikkeling nodig.

De vier categorieën van effectiviteitsopvattingen leiden tot verschillende criteria om de kwaliteit van het onderwijs te beoordelen. In het rationele doel model gaat het bijvoorbeeld om hoge leerprestaties, bij het interne proces model om lage lesuitval, bij het ‘human relations’ model om gemotiveerde docenten en bij het open systeem model om een hoog leerlingenaantal. De keuze van de kwaliteitscriteria kan afhangen van de beoordelaar (schoolleider, ouder, inspecteur van onderwijs), of van de situatie waarin een instelling zich bevindt. Scheerens (1992, 1999) stelt voor om de effectiviteitscriteria te ordenen in een volgorde van middelen naar doelen, waarbij productiviteit wordt gezien als het uiteindelijke doel en de andere criteria als voorwaarden (zie Scheerens & Bosker, 1997, p.10). De relaties tussen de effectiviteitscriteria zijn weergegeven in figuur 5.1. Deze opvatting is niet in tegenspraak met het concurrerende waarden perspectief van Quinn c.s. Ook in dit model geldt bijvoorbeeld dat teveel nadruk op interne processen en consensus ten koste kan gaan van de externe gerichtheid en omgekeerd (Scheerens, 1999, p.49). Quinn (1998) beschrijft hoe jonge organisaties vooral extern zijn georiënteerd om middelen te verwerven, terwijl in meer gevestigde organisaties de aandacht verschuift naar consensus en interne stabiliteit. Uiteindelijk zou er een evenwicht moeten ontstaan, waarbij de productiviteit en het verwerven van middelen wat meer nadruk krijgen dan de interne ondersteunende waarden (zie ook paragraaf 2.5). Houtveen, Voogt, van der Vegt en van de Grift (1996, p.53) wijzen erop dat een integrale benadering van de vier organisatie modellen is te vinden in de ‘lerende organisatie’ (Senge, 1990). Een lerende school zou de vier modellen met de bijbehorende waarden en effectiviteitscriteria evenwichtig moeten hanteren.



Figuur 5.1 Doel-middel relaties tussen effectiviteitscriteria (Scheerens, 1999, p.48)

Om na te gaan in hoeverre de vier modellen van Quinn en Rohrbaugh empirisch zijn te vinden in het onderwijs, werd door Houtveen en haar collega's op basis van het concurrerende waarden model een instrument ontwikkeld waarmee de organisatiecultuur van basisscholen onderzocht kan worden. Van de 20 schalen die werden geconstrueerd aan de hand van de effectiviteitscriteria die horen bij de vier kwadranten in het model (zie figuur 2.3), bleken er acht geschikt om verschillen tussen schoolculturen te detecteren. Uit factoranalyses kwam vervolgens een drie factorenmodel naar voren (Houtveen, Voogt, van der Vegt & van de Grift, 1995, p.93). In het driefactoren model zijn twee van de vier modellen van Quinn en Rohrbaugh direct terug te vinden, namelijk het 'human relations' model en het open systeem model. De derde factor wordt gevormd door het rationele doel model en het interne proces model samen. Deze uitkomsten komen overeen met de resultaten van Quinn en Rohrbaugh (1983), die aangeven dat het rationele doel model en het interne procesmodel conceptueel het dichtst bij elkaar liggen (zie ook Quinn & Cameron, 1983). Met behulp van de scores op de drie schoolcultuurtypen constateerden Houtveen et al. (1995, p.99) dat ongeveer tweevijfde van de basisscholen een 'zuiver'schoolcultuurtype heeft. De meeste scholen hebben een gemêleerde schoolcultuur. De combinatie tussen een teamgerichte ('human relations') en een groeigerichte (open systeem) schoolcultuur komt daarbij het meeste voor (éénvijfde van de scholen).

Voortbouwend op het werk van Houtveen et al. onderzocht Maslowski (2001) het verband tussen de schoolcultuur en de prestaties van scholen in het voortgezet onderwijs. Maslowski vond met behulp van clusteranalyse vijf schoolcultuurtypes: 28% van de scholen had een intern gerichte cultuur (interne proces en 'human relations'), 16% een cultuur met de nadruk op doelen en prestaties (rationele doel), 33% een flexibele cultuur ('human relations' en open systeem), 15% een op beheersing gerichte cultuur (rationele doel en interne proces) en 5% een gemengde cultuur (vooral rationele doel en 'human relations', maar ook open systeem en interne proces). Er bleek een negatief effect te zijn van een flexibele cultuur op de leerprestaties, vergeleken met een intern georiënteerde cultuur. Maslowski concludeert hieruit dat samenwerking effectief is wanneer deze is gericht op de aanpak van problemen in het primaire proces, maar niet zozeer wanneer het alleen gaat om innovatie. Verder bleek er een positieve relatie te bestaan tussen een op beheersing gerichte schoolcultuur en de doorstroomcijfers in de bovenbouw (Maslowski, 2001, p.115).

Vermeulen (1997) maakte gebruik van het concurrerende waarden model bij haar onderzoek naar de relatie tussen personeelsbeleid, schoolorganisatie en de arbeidsbeleving van leraren in het voortgezet onderwijs in Nederland. Zij geeft aan dat de uitgangspunten van de *Human Resource Management* theorie vooral terug zijn te vinden in het rationele doel kwadrant en het 'human relations' kwadrant van het model (Vermeulen, 1997, p.31). In de doelstellingen van human resource management zitten namelijk twee elementen: het realiseren van het organisatiedoel en de ontwikkeling van het menselijk potentieel. Uit haar onderzoek concludeert Vermeulen (1997, p.189) dat er nogal wat overeenkomsten zijn tussen de kenmerken van scholen met betere leerlingprestaties (effectiviteitsmaat voor rationele doel model) en scholen met een positievere arbeidsbeleving van leraren (effectiviteitsmaat voor 'human relations' model). De managementrol van mentor blijkt zowel een sterk verband te hebben met de arbeidsbeleving van leraren als met leerlingprestaties. Andere criteria die overeenkomen voor de twee effectiviteitsmodellen zijn gezamenlijke uitgangspunten, hoge verwachtingen van leerlingen en een ordelijk en plezierig schoolklimaat. Vermeulen (1997, p.190) stelt dat voor de implementatie van Human Resource Management een "meer dan gemiddelde uitvoering" nodig is van de managementrollen die horen bij het 'human relations' en het rationele doel model. Er blijken echter ook sterke relaties te bestaan met de uitvoering

van de rollen uit het interne proces model. Bovendien wijst Vermeulen erop dat “een synthese met de waarden van het open systeem kwadrant” een volgende stap zou kunnen zijn in de ontwikkeling van personeelsmanagement.

Het werk van Vermeulen biedt empirische ondersteuning aan de toepassing van het concurrerende waarden model bij de integrale benadering van schoolleiderschap, zoals onder meer door Krüger en Witziers (2003) wordt voorgesteld (zie paragraaf 2.2). Krüger en Witziers benadrukken daarbij de mogelijkheden van het model voor de ontwikkeling van managementvaardigheden. De Engelse onderzoekers Davies en Coates (2005) wijzen erop dat het concurrerende waarden model kan worden gebruikt om een onderscheid te maken tussen verschillende (concurrerende) opvattingen van strategisch leiderschap in scholen. In het artikel van Davies en Coates gaat het om de bruikbaarheid van een door de Engelse overheid gestimuleerd rationeel planningsmodel voor scholen. De onderzoekers stellen dat het rationele planningsmodel tekort schiet, omdat nog niet voldoende duidelijk is welke schoolfactoren precies van invloed zijn op de uitkomsten. Zij benadrukken dat schoolleiders verschillende strategische benaderingen moeten kunnen hanteren, zodat enerzijds kan worden omgegaan met onzekerheden als gevolg van het overheidsbeleid en de concurrentie tussen scholen, terwijl anderzijds het interne proces en de professionele ontwikkeling van docenten steeds in het oog worden gehouden.

In de Verenigde Staten onderzocht Thompson (2000) de effectiviteit van schoolleiders aan de hand van de acht leiderschapsrollen in het concurrerende waarden model. Daarbij ging hij uit van drie leiderschapstypen die waren vastgesteld met behulp van de leiderschapstheorie van Bolman en Deal (1991). Bolman en Deal onderscheiden vier kaders voor leiderschap: het *structuur* kader, waarin de nadruk ligt op efficiëntie en effectiviteit, het *human resource* kader, waarbij de motivatie van de individuele werknemer centraal staat, het *politieke* kader, waarin het gaat om competitie en onderhandelen en het *symbolische* kader dat een sterke cultuur en missie benadrukt. Leiders met een bovengemiddelde score op alle vier kaders behoren volgens Thompson tot het ‘volledig gebalanceerde’ leiderschapstype. Leiders van het ‘matig gebalanceerde’ type scoorden bovengemiddeld op drie van de vier leiderschapskaders van Bolman en Deal. Het ‘ongebalanceerde’ leiderschapstype scoorde slechts op één of twee van de kaders bovengemiddeld. Thompson (2000) vond dat leiders in de volledige en matig gebalanceerde groepen significant hoger scoorden op alle acht leiderschapsrollen in het concurrerende waarden model, dan leiders in de ongebalanceerde groep. Verder bleek dat leiders in de volledig gebalanceerde groep effectiever waren in de vervulling van de rollen die horen bij het interne proces model, dan de leiders in de matig gebalanceerde groep. Thompson concludeert dat leiders effectiever zijn in het uitvoeren van de leiderschapsrollen bij een hogere *cognitieve complexiteit*, ofwel wanneer zij verschillende leiderschapskaders kunnen toepassen. Hij gaat ervan uit, dat alleen leiders die in staat zijn om alle leiderschapskaders in voldoende mate te hanteren, het interne proces voldoende kunnen controleren.

De Amerikaanse onderzoeker Smart (2003) stelt dat:

... the CVF [competing values framework] not only integrates what formerly were to be distinct sets of organizational effectiveness criteria, but further integrates into a single theory the interdependent nature of effectiveness criteria, organizational culture, and leadership roles in a manner that stresses the need for cognitive and behavioral complexity. (p. 679)

Smart gebruikte het concurrerende waarden model om de organisatie effectiviteit van tweejarige *colleges* te onderzoeken. Smart geeft aan dat onderzoek vaak gericht is op het vaststellen van één dominante cultuur, in plaats van de relatieve sterkte van verschillende cultuurtypen. De resultaten van dit type onderzoek wijzen er, volgens hem, in het algemeen op dat organisaties met een ‘human relations’ of open systeem cultuur het meest effectief zijn, gevolgd door organisaties met een rationele doel cultuur, terwijl organisaties met een overwegende interne proces cultuur het minst effectief blijken te zijn. Uit zijn onderzoek concludeert Smart dat er sterke relaties bestaan tussen percepties van de organisatie effectiviteit van de instituten en zowel de complexiteit van de campus cultuur als de complexiteit van het leiderschapsgedrag van *senior campus officials*. Hoe hoger de ‘culturele complexiteit’ en hoe hoger de ‘leiderschapscomplexiteit’, gemeten aan de hand van het concurrerende waarden model, des te hoger de score op acht dimensies van organisatie-effectiviteit (Smart, 2003).

Een laatste voorbeeld dat hier vermeld moet worden, is het onderzoek van Griffith (2003). Griffith gebruikte het concurrerende waarden model als raamwerk voor het analyseren van schooleffectiviteit in het basisonderwijs in de Verenigde Staten. Hierbij werd voor elk organisatie-model afzonderlijk, door middel van *structural equation modeling* getoetst in hoeverre het geschikt was om een vooruitgang in de leerprestaties te verklaren. De uitkomsten van dit onderzoek worden behandeld in hoofdstuk 6.

5.3 Kenmerken van scholen voor voortgezet onderwijs

In hoofdstuk 4 is beschreven op welke wijze de schalen voor het meten van de handelingen van schoolleiders, de schoolcultuur, de schoolorganisatiepraktijk en de schoolbeleving van leerlingen zijn bijgesteld om verschillen tussen scholen zo goed mogelijk naar voren te brengen. Een overzicht van definitieve schalen met de bijbehorende items is te vinden in bijlage 4. De schalen worden in 5.3.1 kort beschreven. De gegevens van scholen die met behulp van deze schalen zijn verkregen, worden gepresenteerd in 5.3.2.

5.3.1 Beschrijving van de schalen

De *rationele doel* schaal voor schoolleidersgedrag meet vooral de gerichtheid van de schoolleider op het stellen van doelen. In de rationele doel schalen voor de schoolcultuur en de schoolorganisatiepraktijk ligt de nadruk echter op het streven naar hoge leerprestaties. In alle drie schalen komt een ordelijke/op leren gerichte sfeer aan de orde. De *interne proces* schaal voor schoolleidersgedrag meet het handhaven van orde en regels en het vastleggen van procedures. In de interne proces schalen voor de schoolcultuur en de schoolorganisatiepraktijk gaat het met name om zorgvuldige procedures, planning en informatievoorziening. Voor de schoolorganisatiepraktijk ligt daarbij de nadruk op de mate waarin docenten worden ondersteund bij vernieuwingen. Het handhaven van orde en regels komt bij de schoolorganisatiepraktijk meer naar voren in de rationele doel schaal dan in de interne proces schaal. Het is echter bekend dat het rationele doel model en het interne proces model conceptueel dicht bij elkaar liggen (zie Quinn & Rohrbaugh, 1983). De *human relations* schaal voor schoolleidersgedrag meet de ondersteuning van de schoolleider bij het werk van docenten. Hierbij gaat het bijvoorbeeld om individuele begeleiding van docenten en het scheppen van een sfeer van onderling vertrouwen tussen docenten. Voor de schoolcultuur ligt in de ‘human relations’ schaal de nadruk samenwerking en gedeelde verantwoordelijkheid en daarnaast op professionele ontwikkeling. In de ‘human relations’ schaal voor schoolorganisatie gaat het om professionele ontwikkeling en (zelf)evaluatie. De *open systeem*

schaal voor schoolleidersgedrag meet vooral de gerichtheid van schoolleiders op externe contacten en het verwerven van middelen. Voor de schoolcultuur ligt de nadruk in de open systeem schaal op zowel profilering als vernieuwing van de school. Voor de schoolorganisatiepraktijk gaat het voornamelijk om *marketing*.

Uit het bovenstaande kan worden opgemaakt, dat de vier schalen waarmee de schoolleiders-handelingen in kaart worden gebracht, elk voornamelijk één van de twee leiderschapsrollen operationaliseren uit het bijbehorende organisatie-model: voor het rationele doel model de rol van *bestuurder*, voor het interne proces model de rol van *controleur*, voor het ‘human relations’ model de rol van *mentor* en voor het open systeem model de rol van *bemiddelaar* (zie tabel 2.2). Dit is hoofdzakelijk het gevolg van accenten die al bij de constructie van de oorspronkelijke vragenlijst werden gelegd (zie bijlagen 1A en 1B). Alleen voor de ‘human relations’ schaal geldt dat in de oorspronkelijke opzet de tweede rol (stimulator) duidelijk was geoperationaliseerd in een aantal items (‘hecht schoolteam’, ‘onderlinge contacten’, ‘invloed op de besluitvorming’). Deze items zijn echter verwijderd omdat ze kennelijk minder geschikt waren om verschillen tussen scholen in kaart te brengen (zie 4.3.4).

Aangezien er conceptuele overlap bleek te bestaan tussen de schalen voor het meten van de schoolcultuur en de schoolorganisatiepraktijk, zijn op basis van factoranalyse twee nieuwe variabelen samengesteld: *prestatiegerichtheid* en *ontwikkelingsgerichtheid* (zie paragraaf 4.7). De variabele prestatiegerichtheid geeft aan in welke mate belang wordt gehecht aan hoge leerprestaties en een ordelijke werksfeer. Deze variabele vertegenwoordigt vooral de dimensie beheersing in het concurrerende waarden model (zie figuur 2.3). De variabele ontwikkelingsgerichtheid meet in hoeverre belang wordt gehecht aan samenwerking, ontwikkeling en vernieuwing. Deze variabele weerspiegelt de flexibiliteit van de schoolorganisatie.

In de vragenlijst over de schoolbeleving meet de schaal ‘waardering van de school’ hoe leerlingen de sfeer op school en de omgang met medeleerlingen ervaren. De schaal ‘waardering van onderwijs’ is gericht op de motivatie en de achtergrond van de leerling. Het ‘werkklimaat’ wordt bepaald aan de hand van de mening van leerlingen over de omgang met leraren, de ondersteuning door leraren bij het leerproces, het klassenklimaat en de organisatie van de vakken en de lessen.

5.3.2 Kenmerken van scholen

Met behulp van de schalen voor het meten van de handelingen van schoolleiders, de schoolcultuur, de schoolorganisatiepraktijk en de schoolbeleving van leerlingen, wordt in deze paragraaf een typering gegeven van scholen voor voortgezet onderwijs in Nederland. Tabel 5.1 bevat een overzicht van de gemiddelde scores op alle schalen.

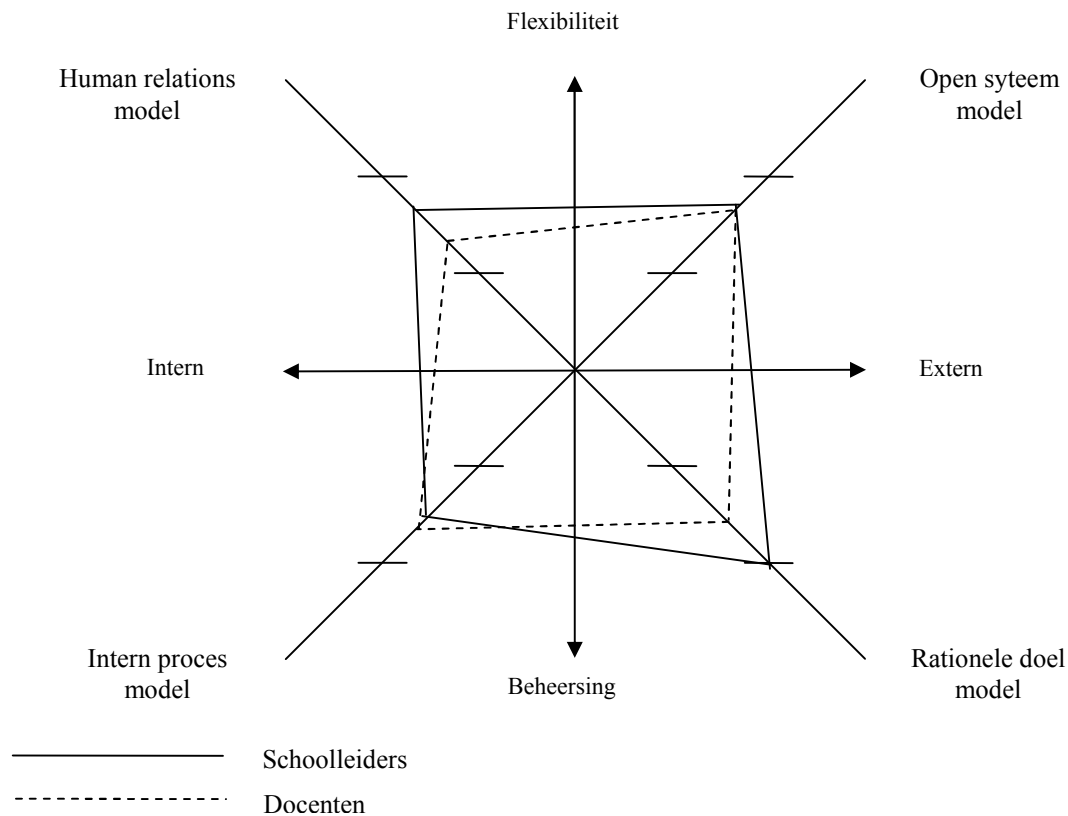
De waarde van de variabele prestatiegerichtheid in tabel 5.1 is berekend door het gemiddelde te nemen van de scores voor rationele doel cultuur, interne proces cultuur en rationele doel organisatiepraktijk. De waarde van de variabele ontwikkelingsgerichtheid is berekend door het gemiddelde te nemen van de scores voor ‘human relations’ cultuur, open systeem cultuur, interne proces organisatiepraktijk en ‘human relations’ organisatiepraktijk.

Tabel 5.1 Beschrijvende gegevens voor bij het onderzoek betrokken scholen (n = 97)

	Minimum	Maximum	Gemiddelde	Std. Deviatie
<i>Schoolleiders</i>				
Rationele doel handelingen	2.00	4.00	3.07	.35
Interne proces handelingen	1.60	3.80	2.56	.44
Human relations handelingen	1.63	3.63	2.64	.41
Open systeem handelingen	1.50	3.50	2.66	.50
<i>Docenten</i>				
Rationele doel handelingen	1.75	3.10	2.61	.30
Interne proces handelingen	1.77	3.43	2.65	.35
Human relations handelingen	1.44	3.13	2.32	.32
Open systeem handelingen	1.76	3.42	2.66	.29
Rationele doel cultuur	1.80	3.52	2.78	.28
Interne proces cultuur	1.98	3.18	2.68	.23
Human relations cultuur	1.88	3.13	2.57	.26
Open systeem cultuur	1.94	3.22	2.65	.27
Rationele doel organisatiepraktijk	2.02	3.39	2.76	.25
Interne proces organisatiepraktijk	2.03	3.14	2.67	.25
Human relations organisatiepraktijk	1.31	2.75	2.06	.31
Open systeem organisatiepraktijk	1.47	3.40	2.42	.29
<i>Docenten: samengestelde variabelen</i>				
Prestatiegerichtheid	1.98	3.23	2.74	.22
Ontwikkelingsgerichtheid	1.83	3.02	2.49	.23
<i>Leerlingen</i>				
Waardering school	2.52	3.31	3.00	.15
Waardering onderwijs	2.71	3.04	2.86	.08
Werkklimaat	2.27	2.94	2.59	.14

Over de handelingen van de schoolleider zijn gegevens verzameld bij zowel schoolleiders als docenten. Uit tabel 5.1 is op te maken dat volgens schoolleiders de nadruk ligt op rationele doel handelingen, gevolgd door ‘human relations’ en open systeem handelingen, terwijl interne proces handelingen minder vaak worden uitgevoerd. Docenten zien geen nadruk op rationele doel handelingen bij schoolleiders. Zij nemen rationele doel, interne proces en open systeem handelingen ongeveer even vaak waar en ‘human relations’ handelingen minder vaak. Kennelijk wordt er met name over de gerichtheid op doelen (rationele doel) en de ondersteuning van het werk van docenten (‘human relations’) verschillend gedacht door schoolleiders en docenten. Houtveen et al. geven aan dat een verschil in perceptie niet verwonderlijk is, omdat schoolleiders hun eigen gedrag moeten beoordelen, terwijl docenten het gedrag van iemand anders beoordelen (Houtveen e.a., 1996, p.90). Van Muijen, Koopman en De Witte (1998, p.109) toonden aan dat er significante verschillen bestaan in de perceptie van de managementcultuur (gebaseerd op het concurrerende waarden model) tussen subgroepen in een organisatie.

In figuur 5.2 zijn de scores van de schoolleiders en de docenten op de vier kwadranten van het concurrerende waarden raamwerk afgebeeld als profielen. Hierin zijn de verschillen in perceptie goed waarneembaar.

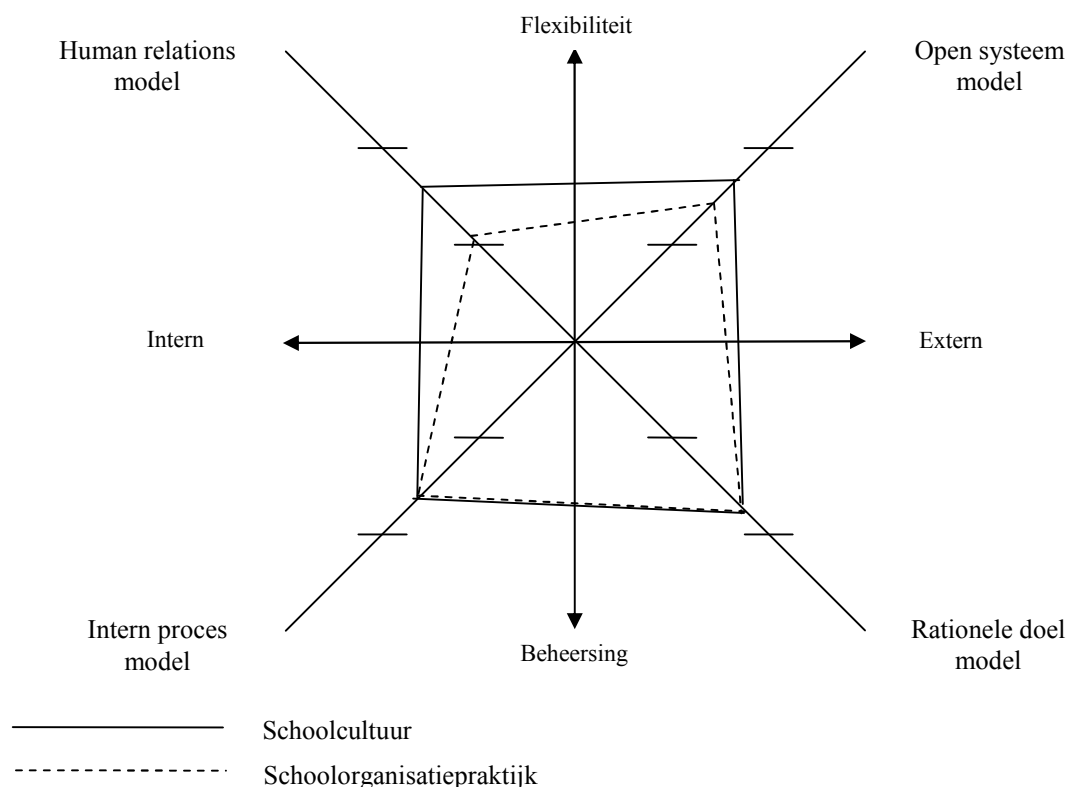


Figuur 5.2 Profielen van schoolleidersgedrag, gebaseerd op de gemiddelde schaalscores van schoolleiders en docenten

De gemiddelde scores op de schoolcultuurschalen lopen niet ver uiteen (zie tabel 5.1). De score voor rationele doel cultuur is wat hoger dan de andere scores. De scores voor interne proces en open systeem cultuur liggen dicht bij elkaar en de score voor ‘human relations’ cultuur is iets lager. Maslowski vond in zijn onderzoek naar de cultuur van scholen in het voortgezet onderwijs ook een hogere gemiddelde score voor de rationele doel oriëntatie dan voor de andere cultuur oriëntaties (Maslowski, 2001, p.85). De score voor ‘human relations’ cultuur was in het onderzoek van Maslowski echter vrijwel even hoog als de score voor rationele doel cultuur, de score voor interne proces cultuur iets lager en de score voor open systeem cultuur veel lager dan de andere scores. Tussen de dataverzameling voor dit onderzoek en die voor het onderzoek van Maslowski ligt een periode van ongeveer zes jaar. Hieruit zou de conclusie getrokken kunnen worden dat in de loop van de tijd de open systeem oriëntatie belangrijker is geworden op scholen voor voortgezet onderwijs in Nederland en de ‘human relations’ oriëntatie minder belangrijk. De verschillen in de gemiddelde scores kunnen echter ook zijn ontstaan door verschillen in de operationalisering van de vier organisatiemodellen. In het onderzoeksinstrument van Maslowski meet de ‘human relations’ schaal vooral het belang van eensgezindheid en samenwerking, terwijl de nieuwe schaal in dit onderzoek ook de professionele ontwikkeling meeneemt. De open systeem schaal van Maslowski meet vooral de gerichtheid op aanpassing en verandering, terwijl in het nieuwe instrument het belang van onderwijsvernieuwing expliciet aan de orde komt, naast profilering van de school. Het is mogelijk dat docenten de aandacht voor professionele ontwikkeling binnen de scholen minder hoog inschatten dan de gerichtheid op consensus en samenwerking. Daarmee zou de relatief lage score op de ‘human relations’ schaal in dit onderzoek, vergeleken met het onderzoek van Maslowski, wellicht verklaard kunnen worden. De relatief

hoge score op de open systeem schaal in dit onderzoek kan veroorzaakt zijn doordat docenten meer belang hechten aan onderwijsvernieuwing, dan aan flexibiliteit van de organisatie in het algemeen. Voor de rationele doel schaal en de interne proces schaal lijken de verschillen in operationalisering tussen dit onderzoek en het onderzoek van Maslowski minder prominent, hoewel de items in het onderzoeksinstrument van Maslowski algemener zijn geformuleerd en meer zijn gericht op de onderliggende waarden (zie Maslowski, 2001, p.75).

De gemiddelde scores voor de schoolorganisatiepraktijk vertonen grotere verschillen dan de scores op de schoolcultuurschalen (zie tabel 5.1). De scores op de rationele doel en interne proces schalen zijn voor de schoolorganisatiepraktijk vrijwel even hoog als voor de schoolcultuur. De scores op de open systeem schaal en met name de ‘human relations’ schaal liggen voor de schoolorganisatiepraktijk veel lager dan voor de schoolcultuur. Dit is goed te zien in figuur 5.3, waar de profielen voor de schoolcultuur en de schoolorganisatiepraktijk zijn afgebeeld.



Figuur 5.3 Profielen van schoolcultuur en schoolorganisatiepraktijk op de gemiddelde schaalscores van docenten

Uit deze resultaten zou geconcludeerd kunnen worden dat de schoolorganisatiepraktijk voor de ‘human relations’ en open systeem oriëntaties niet in overeenstemming is met het belang dat wordt gehecht aan de onderliggende waarden. Een verklaring van de verschillen in de profielen zou echter ook weer gezocht kunnen worden in verschillende operationalisering van de organisatiemodellen. De ‘human relations’ schaal voor de schoolorganisatiepraktijk is voornamelijk gericht op professionele ontwikkeling, terwijl de corresponderende schaal voor de schoolcultuur ook het belang meet van consensus en samenwerking. De open systeem schaal voor de schoolorganisatiepraktijk gaat alleen over externe profilering van de school,

terwijl de overeenkomende schaal voor de schoolcultuur ook onderwijsvernieuwing meeneemt. De gemiddelde waarden voor de samengestelde variabelen (zie tabel 5.1) weerspiegelen het accent dat met name in de schoolorganisatiepraktijk ligt op prestatiegerichtheid (beheersing) vergeleken met ontwikkelingsgerichtheid (flexibiliteit).

Uit de gemiddelde scores op de schoolbelevingsschalen (zie tabel 5.1) kan worden opgemaakt dat havo-leerlingen hun school vrij hoog waarderen. Verder zijn zij redelijk positief over het belang van onderwijs in het algemeen. De gemiddelde score voor het werkklimaat in de klassen is neutraal.

5.4 Prestatiegerichtheid en ontwikkelingsgerichtheid

Vanwege de conceptuele overlap tussen de schalen voor het meten van de schoolcultuur en de schoolorganisatiepraktijk (zie paragraaf 4.7), zijn twee nieuwe variabelen geconstrueerd: *prestatiegerichtheid* en *ontwikkelingsgerichtheid*. De variabele prestatiegerichtheid geeft aan in welke mate belang wordt gehecht aan hoge leerprestaties en een ordelijke werksfeer. De waarde van deze variabele wordt berekend door het gemiddelde te bepalen van de scores voor rationele doel cultuur, interne proces cultuur en rationele doel organisatiepraktijk ($\alpha = .84$). De variabele ontwikkelingsgerichtheid meet de gerichtheid op samenwerking, ontwikkeling en vernieuwing in de schoolorganisatie. De waarde van deze variabele is het gemiddelde van de scores op 'human relations' cultuur, open systeem cultuur, interne proces organisatiepraktijk en 'human relations' organisatiepraktijk ($\alpha = .88$). De twee samengestelde variabelen kunnen worden opgevat als twee 'cultuurtypen': prestatiegerichtheid vertegenwoordigt daarbij een gecombineerde rationele doel en interne proces cultuur en ontwikkelingsgerichtheid een combinatie van 'human relations' en open systeem cultuur. Aan de hand van deze twee cultuurtypen werd een indeling gemaakt van scholen. Hierbij werden vier categorieën onderscheiden: scholen die alleen boven het gemiddelde scores op prestatiegerichtheid (14%) of alleen op ontwikkelingsgerichtheid (20%), scholen die op beide typen bovengemiddeld scores (35%) en scholen die op geen van beide typen bovengemiddeld scores (31%). Hierdoor is het mogelijk om de uitkomsten van dit onderzoek te vergelijken met de resultaten van Houtveen e.a. (1996) en Maslowski (2001).

Houtveen, Voogt, van der Vegt en van de Grift (1996) vonden in hun onderzoek naar de schoolcultuur op basisscholen met behulp van factoranalyse drie schoolcultuurtypen: prestatiegericht (rationele doel en interne proces model), teamgericht ('human relations' model) en vernieuwingsgericht (open systeem model). Voor het vaststellen van de schoolcultuur werd gebruik gemaakt van de mediaan. Om tot een bepaald schoolcultuurtype te behoren diende een school tenminste in het vijftigste percentiel te behoren (Houtveen e.a., 1996, p.98). Houtveen en haar collega's constateerden dat de combinatie van een teamgerichte en een vernieuwingsgerichte cultuur relatief vaak voorkomt in Nederland (21% van de basisscholen). Een 'zuivere' prestatiegerichte cultuur komt minder vaak voor (9% van de basisscholen). Maslowski (2001) onderscheidde met behulp van clusteranalyse vijf schoolcultuurtypen in het voortgezet onderwijs in Nederland. Hij vond op 33% van de scholen een flexibele cultuur (overwegend 'human relations' en open systeem model) en op 15% van de scholen een op beheersing gerichte cultuur (overwegend rationele doel en interne proces model).

In tabel 5.2 worden de uitkomsten voor de schoolcultuurtypen die in dit onderzoek zijn gevonden vergeleken met de uitkomsten van Houtveen et al. (1996) voor het basisonderwijs en van Maslowski (2001) voor het voortgezet onderwijs in Nederland. Ondanks verschillende

onderzoekinstrumenten, verschillende indelingscriteria en verschillende schoolsoorten, stemmen de uitkomsten redelijk overeen.

Tabel 5.2 *Verdeling van scholen over schoolcultuurtypen volgens de uitkomsten van verschillende onderzoeken*

Cultuurtype	Dit onderzoek %	Houtveen e.a. (1996) %	Maslowski (2001) %
Rationele doel en interne proces ¹	14	9	15
'human relations' en open systeem ²	20	21	33
Combinaties	35	30	35
Rest ³	31	40	16

¹ Prestatiegerichtheid

² Ontwikkelingsgerichtheid

³ Hoge score op één model of geen van de vier modellen.

Het opvallendste verschil komt voor bij de gecombineerde 'human relations' en open systeem cultuur: in het onderzoek van Maslowski werd bij deze combinatie een hoger percentage scholen gevonden dan in beide andere onderzoeken. Dit kan natuurlijk een gevolg zijn van verschillende indelingsmethoden, maar ook van verschillen in de operationalisering van organisatiemodellen die vooral naar voren komen bij de 'human relations' en open systeem schalen (zie boven). De schalen in dit onderzoek komen beter overeen met de schalen van Houtveen et al.(1996) dan met die van Maslowski (2001).

Om de bruikbaarheid van de twee schoolcultuurtypen nader te exploreren werden de relaties onderzocht met de achtergrondvariabelen denominatie en schoolgrootte. Aangezien denominatie van invloed is op de missie van een school en de achtergrond van docenten, zou er een verband kunnen bestaan tussen denominatie en schoolcultuur. In Nederland was er volgens Dronkers (2004) van eind jaren zestig tot in de negentiger jaren sprake van een grotere effectiviteit van bijzonder onderwijs (katholiek of protestant) ten opzichte van het openbaar onderwijs, maar deze verschillen zouden inmiddels zijn verdwenen (zie ook Driessen & van der Slik, 2001). Door het ontstaan van grote scholengemeenschappen zouden de voordelen van de kleinschaligheid van godsdienstige scholen in Nederland teniet zijn gedaan (Dronkers, 2004). Opdenakker en Van Damme (2006) vonden een klein effect van de katholieke denominatie op de leerresultaten in het Vlaamse voortgezet onderwijs, maar dit effect verdween toen de achtergrond van de leerlingen werd meegenomen.

Uit een variantie analyse (ANOVA) blijkt dat scholen van verschillende denominaties significant verschillen in prestatiegerichtheid ($F=3.35$, $p=0.022$), maar niet in ontwikkelingsgerichtheid ($F= .38$, $p=0.77$). Uit de gemiddelde scores op de twee cultuurtypen in tabel 5.3 is op te maken dat rooms-katholieke en protestants-christelijke scholen prestatiegerichter zijn dan de openbare en overige scholen.¹⁵

Maslowski (2001) vond dat rooms-katholieke scholen in het Nederlandse voortgezet onderwijs vooral hoog scoorden op de rationele doel oriëntatie (gerichtheid op resultaten) en de protestants-christelijke scholen op de interne proces oriëntatie (gerichtheid op regels en orde). Uit een nadere inspectie van de gemiddelde scores op de afzonderlijke schalen voor de

¹⁵ Uit 'post hoc' analyses blijkt dat het verschil tussen rooms-katholieke en zowel openbare als 'overige' scholen significant is ($p < .05$). Het verschil tussen protestants-christelijke scholen en andere denominaties is (net) niet significant.

schoolcultuur en de schoolorganisatiepraktijk (hier niet vermeld) blijkt echter dat in dit onderzoek zowel rooms-katholieke als protestants-christelijke scholen relatief hoog scoren op beide oriëntaties. Verder zouden volgens Maslowski openbare scholen meer gericht zijn op verandering dan andere scholen. Uit tabel 5.3 blijkt echter dat openbare scholen geen grotere ontwikkelingsgerichtheid vertonen dan andere scholen. Dit kan een gevolg zijn van verschillen in de operationalisering van het open systeem model (onderwijsvernieuwing wordt nu expliciet meegenomen), maar het kan ook betekenen dat andere scholen meer aandacht zijn gaan besteden aan flexibiliteit en onderwijsvernieuwing.

Tabel 5.3 Gemiddelde scores op de twee schoolcultuurtypen naar denominatie

Denominatie	Aantal scholen	Prestatiegerichtheid	Ontwikkelingsgerichtheid
Openbaar	29	2.67 (.22)	2.47 (.26)
Rooms-Katholiek	39	2.80 (.19)	2.47 (.23)
Protestants-Christelijk	12	2.81 (.14)	2.54 (.19)
Overig	17	2.66 (.27)	2.51 (.24)
Totaal	97	2.74 (.22)	2.49 (.23)

De waarden voor de (samengestelde) variabelen variëren van 1 (laag) tot 4 (hoog).
Achter de gemiddelde scores is tussen haakjes de standaarddeviatie vermeld.

Over het verband tussen schoolgrootte en de schoolcultuur zijn verschillende hypothesen mogelijk. In kleine scholen zou het ‘human relations’ model belangrijker kunnen zijn, omdat het personeel meer met elkaar in contact komt. In grote scholen zou het interne proces model de nadruk kunnen krijgen, omdat er in een grote organisatie meer behoefte is aan regels en procedures. Er bestaat echter geen consensus over een optimale schoolgrootte (zie bijv. Bosker, 1992; Luyten, 1994).

Uit een variantie analyse blijkt dat scholen van verschillende grootte significant verschillen in zowel prestatiegerichtheid ($F = 4.67$, $p = .004$) als ontwikkelingsgerichtheid ($F = 4.99$, $p = .003$). Tabel 5.4 geeft een overzicht van de gemiddelde scores van scholen van verschillende grootte op de twee schoolcultuurtypen.

Tabel 5.4 Gemiddelde scores op de twee schoolcultuurtypen naar schoolgrootte

Schoolgrootte	Aantal scholen	Prestatiegerichtheid	Ontwikkelingsgerichtheid
minder dan 500 leerlingen	5	2.53 (.18)	2.22 (.19)
500 tot 1000 leerlingen	35	2.68 (.24)	2.42 (.21)
1000 tot 1500 leerlingen	40	2.82 (.19)	2.55 (.24)
Meer dan 1500 leerlingen	17	2.74 (.17)	2.54 (.19)
Totaal	97	2.74 (.22)	2.49 (.23)

De waarden voor de (samengestelde) variabelen variëren van 1 (laag) tot 4 (hoog).
Achter de gemiddelde scores is tussen haakjes de standaarddeviatie vermeld.

Naarmate de schoolgrootte toeneemt, neemt de prestatiegerichtheid eerst toe en daarna, bij meer dan 1500 leerlingen, (schijnbaar, het verschil is niet significant) weer af.¹⁶ De ontwikkelingsgerichtheid neemt toe tot een schoolgrootte van 1500 leerlingen en blijft dan stabiel. Deze resultaten wijzen erop dat de relatie tussen schoolgrootte en prestatiegerichtheid, respectievelijk ontwikkelingsgerichtheid, niet lineair is. De gemiddelde scores op de afzonderlijke schalen voor de schoolcultuur en de schoolorganisatiepraktijk (hier niet vermeld) vertonen een vergelijkbaar patroon voor de verschillende schoolgroottes. Mogelijk

¹⁶ Zowel voor prestatiegerichtheid als voor ontwikkelingsgerichtheid geldt dat het verschil tussen de laagste categorie (tot 500 leerlingen) en de hoogste twee significant is. Verder is ook het verschil tussen de tweede (500-1000 leerlingen) en de derde categorie (1000-1500 leerlingen) significant.

is er op grotere scholen meer aandacht voor hoge leerprestaties, een ordelijke werksfeer, samenwerking, ontwikkeling en vernieuwing dan op kleinere, omdat de grootte van de organisatie dit vereist, of omdat er meer faciliteiten zijn. Op scholen met een heel hoog leerlingenaantal zijn er wellicht andere prioriteiten. Maslowski (2001) vond alleen een significant (en niet-lineair) verband tussen schoolgrootte en het open systeem model voor schoolcultuur.

5.5 Het verband tussen kenmerken van scholen en de opbrengsten

In deze paragraaf worden de relaties tussen de schoolkenmerken en de opbrengsten van scholen geëxploreerd. Voor de meeste schoolkenmerken in dit onderzoek is nog geen eenduidig verband met de onderwijsopbrengsten aangetoond. Uit een kwantitatieve meta-analyse van Witziers, Bosker en Krüger (2003) blijkt dat een (direct) effect van schoolleiderschap op de leerprestaties in het voortgezet onderwijs niet bewezen kan worden. In het basisonderwijs (buiten Nederland) kunnen wel relaties tussen schoolleiderschap en de uitkomsten worden aangetoond. Witziers et al. concluderen dat het verband tussen schoolleidersgedrag en de leerprestaties zwak is. Het formuleren en uitdragen van een missie door schoolleiders lijkt de meeste invloed te hebben op de leerprestaties. Er zijn echter ook aanwijzingen dat een evenwichtige invulling van diverse managementrollen nodig is om tot goede resultaten te komen (zie bijv. Quinn, 1998; Smart, 2003; Thompson, 2000; Vermeulen, 1997). Voor de schoolcultuur geldt eveneens dat een relatie met de opbrengsten van het onderwijs niet eenduidig is vastgesteld (Maslowski, 2001). Verschillende cultuurfactoren zoals prestatiegerichtheid, een ordelijke werksfeer, consensus onder het personeel en een gerichtheid op aanpassing en vernieuwing, zouden van invloed kunnen zijn op de leerprestaties. Het werkklimaat in de klas wordt in het algemeen gezien als een belangrijke factor voor schooleffectiviteit (Scheerens & Bosker, 1997, p. 100). De waardering van de school en de waardering van onderwijs zouden een effect kunnen hebben op de uitkomsten, bijvoorbeeld door verminderde schooluitval (Mulford & Silins, 2003).

In 5.5.1 worden eerst de uitkomstvariabelen voor dit onderzoek beschreven. Daarna worden in 5.5.2 de kenmerken van scholen in verband gebracht met de opbrengsten.

5.5.1 De uitkomstvariabelen

De opbrengsten van het onderwijs worden in dit onderzoek bepaald aan de hand van twee criteria: het gemiddelde cijfer voor alle vakken op het *centraal schriftelijk eindexamen* en het doorstroompercentage in de bovenbouw, voor de havo-opleiding. Deze gegevens zijn ontleend aan databestanden die door de Inspectie van het Onderwijs worden aangemaakt voor de opbrengstenkaarten van het voortgezet onderwijs. Het gemiddelde eindexamencijfer voor alle vakken is een gewogen gemiddelde, waarbij rekening wordt gehouden met het aantal leerlingen per examenvak. Het doorstroompercentage in de bovenbouw is het percentage leerlingen dat zonder zittenblijven van het derde leerjaar doorgaat naar het diploma. Voor de berekening van dit percentage worden de bevorderingspercentages per leerjaar gemiddeld en daarna wordt de gemiddelde bevorderingskans voor alle leerjaren vermenigvuldigd. Vanwege de variatie in het gemiddelde eindexamencijfer en het doorstroompercentage van jaar tot jaar (vgl. Luyten, 1994; Maslowski, 2001), is zowel voor het gemiddelde eindexamencijfer, als het doorstroompercentage het gemiddelde berekend over een periode van drie jaar (2003-2005). Tabel 5.5 geeft een overzicht van kengetallen voor de gemiddelde eindexamencijfers en de doorstroompercentages per jaar en gemiddeld over drie jaar.

Tabel 5.5 Beschrijvende gegevens voor het gemiddelde cijfer op het centraal examen voor alle vakken en het doorstroompercentage in de bovenbouw, voor de havo-opleiding

	N	Minimum	Maximum	Gemiddelde	Std. deviatie
<i>Gemiddeld eindexamencijfer</i>					
2003	94	5.82	6.86	6.30	.21
2004	97	5.24	6.81	6.31	.22
2005	97	5.49	6.63	6.22	.21
Gemiddelde 2003-2005	97	5.68	6.72	6.28	.19
<i>Doorstroompercentage bovenbouw</i>					
2003	94	22.19	87.59	62.76	12.71
2004	96	34.52	93.15	64.71	11.66
2005	97	44.37	89.95	65.50	9.67
Gemiddelde 2003-2005	97	38.05	88.29	64.23	9.54

Uit tabel 5.5 is af te lezen dat er grote verschillen zijn tussen scholen in het gemiddelde eindexamencijfer en het doorstroompercentage voor de havo-opleiding. In tabel 5.6 worden de correlaties tussen de gemiddelde eindexamencijfers en tussen de gemiddelde doorstroomcijfers voor de jaren 2003 tot 2005 gegeven.

Tabel 5.6 Correlaties tussen gemiddelde cijfers op het centrale examen en tussen doorstroompercentages voor drie opeenvolgende jaren

	2003	2004	2005
<i>Gemiddeld eindexamencijfer</i>			
2003	1.00		
2004	.68**	1.00	
2005	.62**	.63**	1.00
Gemiddelde 2003-2005	.87**	.89**	.86**
<i>Doorstroompercentage bovenbouw</i>			
2003	1.00		
2004	.60**	1.00	
2005	.51**	.50**	1.00
Gemiddelde 2003-2005	.87**	.85**	.77**

** Correlatie is significant op 0.01 niveau

De gemiddelde eindexamencijfers correleren redelijk voor twee achtereenvolgende jaren en ook over een periode van twee jaar. De doorstroompercentages zijn eveneens redelijk stabiel. De gemiddelde eindexamencijfers en de gemiddelde doorstroompercentages over drie achtereenvolgende jaren kunnen worden beschouwd als betrouwbare maten voor de opbrengsten van scholen. De correlatie tussen het gemiddelde eindexamencijfer en het gemiddelde doorstroompercentage over drie jaar bedraagt .36 ($p < .01$). Dit betekent dat er een verband is tussen de twee uitkomstvariabelen, die verschillende dimensies weergeven van de schoolprestaties. Het eindexamencijfer kan worden beschouwd als een maat voor schooleffectiviteit en het doorstroompercentage als een maat voor schoolefficiëntie (Maslowski, 2001, p.107).

5.5.2 Relaties tussen schoolkenmerken en uitkomsten

Om een indicatie te krijgen van de mate waarin schoolleidersvariabelen en andere schoolkenmerken samenhangen met de uitkomsten, werden Pearson correlatiecoëfficiënten berekend. Feitelijk gaat het hierbij om een 'direct-effect-model' (zie hoofdstuk 1). Voor de uitkomsten werden niet alleen de 'ruwe' scores gebruikt, maar ook de 'bolletjes' scores van

de opbrengstenkaarten van scholen voor voortgezet onderwijs. Op de opbrengstenkaarten wordt voor alle kengetallen de positie van de school ten opzichte van andere scholen uitgedrukt in vijf categorieën (visueel weergegeven als ‘bolletjes’). Bij de berekening van de ‘bolletjes’ score wordt rekening gehouden met de samenstelling van de leerlingenpopulatie. Tot en met het jaar 2005 ging het hierbij om het percentage ‘cumi-leerlingen’, dit is het aantal leerlingen waarvoor scholen extra bekostiging ontvangen volgens de *Faciliteitenregeling Culturele Minderheden*. Vanaf 2006 gaat het om het percentage leerlingen uit een ‘armoedeprobleemcumulatiegebied’. De bolletjes geven een indicatie van de ‘toegevoegde waarde’ (de afwijking ten opzichte van de verwachte score, op basis van een regressievergelijking)¹⁷. De correlatie tussen de gemiddelde ‘bolletjes’ score voor het eindexamencijfer en het gemiddelde ruwe eindexamencijfer voor alle vakken over drie jaar (2003-2005) bedraagt .82. De correlatie tussen de gemiddelde ‘bolletjes’ score voor de doorstroom in de bovenbouw en het gemiddelde ruwe doorstroompercentage over drie jaar is .92

In tabel 5.7 zijn de correlaties tussen de schoolkenmerken en de uitkomsten vermeld. Voor de schoolkenmerken zijn geaggregeerde factorscores gebruikt (zie paragraaf 4.6). De uitkomsten zijn gebaseerd op de gemiddelde cijfers voor het centraal schriftelijk eindexamen en de gemiddelde doorstroompercentages in de bovenbouw (zie 5.5.1). De ruwe scores zijn niet gecorrigeerd voor de samenstelling van de leerlingenpopulatie, de ‘bolletjes’ scores wel. De scores zijn gemiddeld over drie jaar (2003-2005).

Tabel 5.7 *Correlaties tussen schoolkenmerken en uitkomsten, aan de hand van ruwe scores en scores die zijn gecorrigeerd voor de samenstelling van de leerlingenpopulatie ('bolletjes')*

	Doorstroompercentage		Gemiddeld eindexamencijfer	
	ruw	'bolletjes'	ruw	'bolletjes'
<i>Schoolleidersgedrag (schoolleiders)</i>				
Rationeel doel	.03	.04	-.07	-.09
Intern proces	-.01	-.00	.03	.06
Human relations	-.01	-.03	.09	.13
Open systeem	-.08	-.04	-.33**	-.32**
<i>Schoolleidersgedrag (docenten)</i>				
Rationeel doel	.05	.05	-.13	-.20
Intern proces	.03	.04	-.10	.03
Human relations	-.04	-.05	.04	-.01
Open systeem	.02	.05	-.20*	-.20*
<i>Schoolcultuur (docenten)</i>				
Prestatiegerichtheid	.36**	.30**	.21*	.13
Ontwikkelingsgerichtheid	.24*	.25*	-.08	-.15
<i>Schoolbeleving (leerlingen)</i>				
Waardering school	.37**	.25*	.11	-.14
Waardering onderwijs	.05	.08	-.12	-.14
Werkklimaat	.25*	.23*	-.13	.05

** Correlatie is significant op 0.01 niveau

* Correlatie is significant op 0.05 niveau

¹⁷ De ‘bolletjes’ score voor een school was 1 als 90% van de scholen een betere afwijkingsscore had, 2 als 75-90% een betere afwijkingsscore had, 3 voor de middelste 50%, 4 voor vestigingen met een betere afwijkingsscore dan 75% van de scholen, maar minder dan de bovenste 10% en 5 voor de top 10%.

Uit tabel 5.7 blijkt dat er een significant en negatief direct verband is tussen open systeem handelingen van schoolleiders en de gemiddelde eindexamencijfers, zowel voor ruwe als gecorrigeerde scores. Uitgaande van de correlaties (-.20 tot -.33) kan worden berekend dat 4-11% van de variantie in de eindexamencijfers is geassocieerd met verschillen in schoolleidersgedrag. Deze waarde is veel hoger dan het percentage van 1% dat werd gevonden door Witziers, Bosker en Krüger (2003) bij een kwantitatieve meta-analyse van onderzoek naar directe effecten van schoolleiderschap op leerlingprestaties. Witziers en zijn collega's vonden overigens, naast een klein positief 'totaal' effect van schoolleidersgedrag, eveneens een negatieve relatie van specifieke handelingen van schoolleiders die zijn gericht op het verbeteren en ontwikkelen van de school met de uitkomsten. Hierbij veronderstellen zij dat schoolleiders in scholen met lage leerprestaties geneigd zullen zijn om actie te ondernemen om de school te verbeteren (zie ook Teddlie & Stringfield, 1993). De negatieve relatie tussen schoolleidersgedrag en de leerprestaties wordt in het volgend hoofdstuk nader besproken.

De variabelen 'prestatiegerichtheid' en 'ontwikkelingsgerichtheid' hebben een significante positieve relatie met het doorstroompercentage in de bovenbouw. De variabele 'prestatiegerichtheid' heeft ook een positieve relatie met het ruwe gemiddelde eindexamencijfer, maar niet met de 'bolletjes' score. Een verband tussen een prestatiegerichte cultuur en de uitkomsten lijkt voor de hand te liggen. Het ontbreken van een dergelijk verband voor de gecorrigeerde scores wijst er echter op dat niet alleen de leerprestaties, maar ook de prestatiegerichtheid van een school wordt beïnvloed door de samenstelling van de leerlingenpopulatie. Hier wordt in het volgende hoofdstuk nader op ingegaan. De waardering van de school door leerlingen en het werkklimaat in de klas zijn positief gerelateerd aan het doorstroompercentage. Deze uitkomsten zijn verklaarbaar vanuit het idee dat betrokkenheid van de leerlingen de schooluitval vermindert (zie bijv. Mulford & Silins, 2003). De variabele 'waardering onderwijs' heeft geen effect op de uitkomsten. Deze variabele is niet erg betrouwbaar op schoolniveau (zie tabel 4.17).

Om meer inzicht te krijgen in het verband tussen schoolkenmerken en de uitkomsten zijn de verschillen tussen scholen met hoge opbrengsten en scholen met lage opbrengsten onderzocht. Hiertoe zijn de tien scholen met de hoogste gemiddelde eindexamencijfers vergeleken met de tien scholen met de laagste gemiddelde eindexamencijfers. Uit tabel 5.7 werd al duidelijk dat er tussen scholen met hoge en lage eindexamencijfers significante verschillen zijn in de scores voor de schoolleidersgedrag. Bij nadere analyse blijkt dat er niet alleen verschillen zijn voor open systeem handelingen, maar in de docentscores ook voor rationele doel handelingen en 'human relations' handelingen. Tabel 5.8 geeft een overzicht van de docentenscores voor de handelingen van schoolleiders van scholen met hoge en lage eindexamencijfers. In de tabel zijn ook de gemiddelde eindexamencijfers voor de groepen vermeld.

Tabel 5.8 Gemiddelde eindexamencijfers en docentscores voor schoolleidersgedrag van scholen met hoge, gemiddelde en lage eindexamencijfers

Indeling Scholen	N	Eindexamen-cijfer	Rationele doel	Interne proces	Human relations	Open systeem
Laagste score	10	5.92 (.15)	2.61 (.30)	2.54 (.33)	2.23 (.24)	2.74 (.38)
Middengroep	77	6.28 (.12)	2.65 (.27)	2.67 (.33)	2.35 (.32)	2.69 (.26)
Hoogste score	10	6.57 (.07)	2.31 (.38)	2.56 (.45)	2.12 (.37)	2.41 (.29)
Totaal	97	6.28 (.18)	2.61 (.30)	2.65 (.35)	2.32 (.32)	2.66 (.29)

De verschillen tussen de drie groepen in de gemiddelde eindexamencijfers zijn significant. Het verschil tussen de hoogste en de laagste groep in het gemiddelde eindexamencijfer is 0.65.

Voor de rationele doel handelingen zijn er significante verschillen tussen de hoogste groep en de andere twee groepen ($F=6.57$, $p=.00$), voor de interne proces handelingen zijn er geen verschillen tussen de groepen ($F=1.01$, $p=.37$), voor de ‘human relations’ handelingen zijn er significante verschillen tussen de hoogste en de middelste groep ($F=2.87$, $p=.06$) en voor de open systeem handelingen zijn er significante verschillen tussen de hoogste groep en de andere twee groepen ($F = 4.56$, $p=.01$). Uit de gegevens in tabel 5.8 kan worden opgemaakt dat schoolleiders op de hoogst presterende scholen (volgens de docenten) minder nadruk leggen op doelen, ondersteuning van het werk van docenten, externe contacten en het verwerven van middelen.

Een nadere analyse van de scores voor prestatiegerichtheid en ontwikkelingsgerichtheid van scholen met hoge en lage eindexamencijfers (hier niet weergegeven) leverde geen nieuwe inzichten op. Scholen met hoge examencijfers verschillen alleen significant van scholen met lage examencijfers voor wat betreft de prestatiegerichtheid. Hoewel er volgens tabel 5.7 geen significante relaties bestaan tussen de waardering van de school of het werkklimaat in de klassen en de eindexamencijfers, blijken er bij nadere inspectie toch verschillen te zijn voor scholen met hoge en lage eindexamencijfers. De waardering van de school is voor scholen met lage eindexamencijfers significant lager dan voor scholen met gemiddelde of hoge eindexamencijfers. De score voor het werkklimaat in de klassen is voor scholen met lage eindexamencijfers significant lager dan voor scholen met hoge eindexamencijfers.

5.6 Discussie

In dit hoofdstuk zijn gegevens over het gedrag van schoolleiders, de schoolcultuur, de schoolorganisatiepraktijk en de schoolbeleving van leerlingen gepresenteerd (zie paragraaf 5.3). Uit een vergelijking van de gegevens van docenten en schoolleiders bleek dat schoolleiders zelf menen dat zij de nadruk leggen op doelen, terwijl docenten hier geen grotere nadruk ervaren. Docenten denken ook anders dan schoolleiders over de ondersteuning van hun werk door de schoolleider. Zij nemen een lagere ‘human relations’ oriëntatie waar dan schoolleiders zelf aangeven. Vanwege de verschillende positie die zij innemen in de organisatie kunnen docenten anders tegen bepaalde zaken aankijken dan schoolleiders (zie bijv. Houtveen e.a., 1996; Van Muijen, Koopman en De Witte, 1998), maar schoolleiders zouden ook meer sociaal wenselijke antwoorden gegeven kunnen hebben. Uit een vergelijking van de uitkomsten voor de schoolorganisatiepraktijk en de schoolcultuur zou de conclusie kunnen worden getrokken dat de schoolorganisatie in het algemeen minder flexibel is dan wenselijk wordt geacht. De scores op de ‘human relations’ en open systeem schalen zijn voor de schoolcultuur (onderliggende normen en waarden) hoger dan voor de schoolorganisatiepraktijk (realisatie). De verschillen in de uitkomsten kunnen echter ook een gevolg zijn verschillen in de operationalisering van het concurrerende waarden model voor de schoolcultuur en de schoolorganisatiepraktijk.

Omdat er conceptuele overlap bleek te zijn tussen de schalen voor het meten van de schoolcultuur en de schoolorganisatiepraktijk, zijn twee nieuwe variabelen geconstrueerd: ‘prestatiegerichtheid’ en ‘ontwikkelingsgerichtheid’. Wanneer scholen worden ingedeeld aan de hand van deze twee ‘cultuurtypen’ (zie paragraaf 5.4), dan blijken de uitkomsten goed overeen te stemmen met de typering van Houtveen et al. (1996) en Maslowski (2001). Dit wijst erop dat de twee cultuurtypen valide zijn. Gemiddeld scores scholen hoger op

prestatiegerichtheid dan op ontwikkelingsgerichtheid. Voor rooms-katholieke en protestants-christelijke geldt dat de nadruk op prestatiegerichtheid nog wat hoger is dan voor andere denominaties, terwijl de ontwikkelingsgerichtheid van de scholen vergelijkbaar is. Naarmate de schoolgrootte toeneemt lijkt ook de prestatiegerichtheid en de ontwikkelingsgerichtheid van scholen toe te nemen. Wellicht is het in grotere organisaties meer van belang om hier aandacht aan te geven, of zijn grotere scholen hiervoor beter toegerust dan kleine scholen. Voor hele grote scholen (meer dan 1500 leerlingen) neemt de prestatiegerichtheid mogelijk wat af, terwijl de ontwikkelingsgerichtheid gelijk blijft. Bij een groot leerlingenaantal is een hoge prestatiegerichtheid over de hele linie misschien moeilijker te realiseren vanwege de variatie in de opleidingen binnen de school en de diversiteit van de leerlingen.

In de schoolorganisatiepraktijk lijkt de nadruk in het algemeen meer te liggen op ‘beheersing’ dan op ‘flexibiliteit’ (zie figuur 5.3), waardoor scholen ook wat sterker zijn gericht op ‘prestaties’ dan op ‘ontwikkeling’. Hieruit kan worden opgemaakt dat scholen wellicht meer aandacht zouden moeten gaan besteden aan professionalisering en aanpassing aan externe omstandigheden, om het ideaal van de ‘lerende organisatie’ te kunnen benaderen (zie paragraaf 2.4). In de lerende school zouden de vier organisatiemodellen van Quinn en Rohrbaugh (1983) met de bijbehorende waarden evenwichtig gehanteerd moeten worden (Houtveen, Voogt, van der Vegt & van de Grift, 1996).

De havo-leerlingen in dit onderzoek zijn redelijk tevreden over hun school en zij waarderen onderwijs gematigd positief. Over het werkklimaat in de klassen zijn zij wat minder tevreden, zij beoordelen het klimaat niet negatief, maar ook niet positief. Wellicht is er een relatie tussen de matige waardering van het werkklimaat en de relatief beperkte aandacht voor de ‘human relations’ oriëntatie in de schoolorganisatie.

In het tweede deel van dit hoofdstuk worden de relaties tussen schoolkenmerken en opbrengsten geëxploreerd (zie paragraaf 5.5). Hierbij worden twee uitkomstvariabelen voor dit onderzoek geïntroduceerd: het gemiddelde cijfer voor alle vakken op het centraal examen en het doorstroompercentage in de bovenbouw, gemiddeld over drie achtereenvolgende jaren (2003-2005). Het eindexamencijfer is een maat voor de effectiviteit, terwijl het doorstroompercentage iets zou kunnen zeggen over de efficiëntie van de school. Voor schoolleidersgedrag blijkt er een negatief verband te bestaan tussen de open systeem oriëntatie en het gemiddelde eindexamencijfer. Dit leidt tot de veronderstelling dat schoolleiders op scholen met relatief lage leerprestaties bezig zijn om de open systeem oriëntatie te versterken. Zij richten zich meer op contacten met derden en het verwerven van middelen. Dit past bij de laatste trend in de effectiviteitsopvattingen in het onderwijs, waarbij flexibilisering en profilering van de instelling voorop staan (zie paragraaf 5.2). Schoolleiders van scholen met de hoogste eindexamencijfers zijn vooral gericht op het interne proces en minder op het stellen van doelen, ondersteuning bij het werk van docenten en profilering van de school. Deze schoolleiders zullen wellicht de stabiliteit en de continuïteit van de organisatie willen bewaken, om de goede prestaties te behouden. Hieruit kan worden opgemaakt dat schoolleiders in hun handelingen reageren op de resultaten van de school. De Amerikaanse onderzoeker Thompson (2000) concludeert uit zijn onderzoek dat leiders effectiever zijn bij een hogere cognitieve complexiteit (zie paragraaf 5.2), ofwel wanneer zij op alle leiderschapsrollen in het concurrerende waarden model voldoende hoog scoren. Ook Smart (2003) wijst op het belang van een hoge ‘leiderschapscomplexiteit’. Door Lawrence, Quinn en Lenk (2003), wordt gesteld dat goede managers alle leiderschapsrollen moeten vervullen (‘behavioral complexity’, zie paragraaf 4.2). Uit de bovenvermelde resultaten van dit onderzoek kan echter worden afgeleid dat het niet vanzelfsprekend is dat schoolleiders van

scholen met hoge eindexamenresultaten ook effectieve leiders zijn met een hoge 'leiderschapscomplexiteit'.

De prestatiegerichtheid en ontwikkelingsgerichtheid van scholen zijn positief gerelateerd aan de doorstroompercentages, maar niet aan de gemiddelde eindexamencijfers, wanneer wordt gecorrigeerd voor de samenstelling van de leerlingenpopulatie. Er is wel een samenhang tussen prestatiegerichtheid en de ruwe eindexamencijfers. Hieruit kan worden geconcludeerd dat de prestatiegerichtheid van een school, net als de leerprestaties, gerelateerd is aan de samenstelling van de leerlingenpopulatie (vgl. Luyten et al., 2005). Zowel de prestatiegerichtheid als de ontwikkelingsgerichtheid van een school hangen positief samen met het doorstroompercentage, evenals de waardering van de school door leerlingen en het werkklimaat in de klassen.

In dit hoofdstuk werd een globale indicatie verkregen van de relaties tussen schoolkenmerken en de opbrengsten met een 'direct-effect model', zonder intermediaire variabelen. Het weglaten van intermediaire variabelen zou echter geen gevolgen moeten hebben voor de grootte van de (totale) effecten van schoolkenmerken op de uitkomsten (Saris & Stronkhorst, 1984, p.86). Context- of achtergrondvariabelen die van invloed zijn op zowel de schoolkenmerken, als de uitkomsten, zijn wel van betekenis voor de relatie tussen deze variabelen. Dergelijke factoren kunnen een deel van deze relatie verklaren, waardoor het feitelijke effect van de ene op de andere variabelen kleiner blijkt te zijn. Bij de analyse van de relaties tussen schoolkenmerken en uitkomsten in dit hoofdstuk werd gecontroleerd voor de samenstelling van de leerlingenpopulatie. Hierbij bleek dat de positieve samenhang tussen de prestatiegerichtheid van een school en de leerprestaties voor een belangrijk deel kan worden verklaard door de samenstelling van de leerlingenpopulatie. Het negatieve verband tussen open systeem schoolleidersgedrag en de leerprestaties blijft wel bestaan als wordt gecorrigeerd voor de samenstelling van de leerlingenpopulatie. Andere schoolkenmerken hangen niet significant samen met de leerprestaties, maar wel met het gemiddelde doorstroompercentage van de derde klas naar het diploma, onafhankelijk van de samenstelling van de leerlingenpopulatie. .

Uit de resultaten die zijn beschreven in dit hoofdstuk kan worden opgemaakt dat er voldoende significante verbanden zijn gevonden tussen de schoolkenmerken en de uitkomstvariabelen, om een test van het volledige onderzoeksmodel (zie figuur 2.5) te kunnen verantwoorden. In dit model zijn verschillende contextvariabelen opgenomen die een effect zouden kunnen uitoefenen op zowel de schoolkenmerken als de uitkomsten. Verder biedt het structurele model de mogelijkheid om te onderzoeken hoe de invloed van de schoolleider verloopt langs intermediaire variabelen (schoolkenmerken). De toetsing van het onderzoeksmodel model is beschreven in het volgende hoofdstuk.

Hoofdstuk 6

Een model voor de invloed van schoolleiders op de leerprestaties

6.1 Inleiding

Onderzoek naar de invloed van schoolleiders op de leerprestaties heeft wisselende resultaten opgeleverd. Uit een kwantitatieve meta-analyse van het onderzoek tussen 1986 en 1996 concluderen Witziers, Bosker en Krüger (2003) dat er geen bewijs is voor een *direct effect* van schoolleiderschap in het voortgezet onderwijs. Buiten Nederland zou in het basisonderwijs wel een klein (direct) effect kunnen bestaan, binnen Nederland niet. In het algemeen verklaart schoolleiderschap niet meer dan 1% van de variatie in leerlingprestaties tussen scholen, volgens Witziers en zijn collega's. Net als Hallinger en Heck (1998) gaan Witziers et al. (2003) er vanuit dat *indirecte effectmodellen* meer inzicht zullen opleveren in de relatie tussen schoolleiderschap en de leerprestaties. In een indirect effectmodel wordt verondersteld dat de invloed van de schoolleider op de leerprestaties verloopt via een of meer interveniërende of *intermediaire* variabelen. Hier is echter nog relatief weinig onderzoek naar gedaan. Witziers, Krüger en Bosker (2003) vonden bij een literatuuronderzoek slechts vijf studies naar indirecte effecten van schoolleiderschap op de leerprestaties (Bosker, de Vos, Witziers & Scheerens, 2000; Hallinger, Bickman & Davis, 1996; Hallinger & Heck, 1996b, 1998; Hill, Rowe & Holmes-Smith, 1996). Meer recent publiceerden bijvoorbeeld ook Leithwood en zijn collega's (zie bijv. Leithwood, Day, Sammons, Harris & Hopkins, 2006a) en Mulford en Silins (2003) resultaten van indirecte effectmodellen. Zowel Bosker, de Vos en Witziers (2003), als De Maeyer en Rymenans (2004; zie ook De Maeyer, Rymenans, Van Petegem, van den Bergh & Rijlaarsdam, 2007) vergeleken de geschiktheid van directe effectmodellen met die van indirecte effectmodellen van schoolleiderschap. Hieruit bleek dat indirecte effectmodellen een betere modelpassing vertoonden. Het is echter moeilijk om algemene conclusies te trekken over indirecte effecten van schoolleiderschap, omdat uiteenlopende concepten van leiderschap werden onderzocht, zoals onderwijkskundig, transformationeel, geïntegreerd, of gedistribueerd leiderschap, waarbij diverse intermediaire variabelen naar voren kwamen, op school-, docenten-, klassen- of leerlingniveau (zie paragraaf 2.8, tabel 2.1).

Vanwege de beperkingen van specifieke leiderschapsmodellen wordt in dit onderzoek uitgegaan van een 'generiek' leiderschapsmodel, zoals gespecificeerd door het concurrerende waarden kader van Quinn c.s. (zie bijv. Quinn, Faerman, Thompson & MacGrath, 2003). De invloed van schoolleiderschap op de leerprestaties wordt onderzocht met behulp van een indirect effectmodel, dat is gebaseerd op het 'algemene beschrijvingsmodel van onderwijkskundig leiderschap' van Bossert et al. (Bossert, Dwyer, Rowan & Lee, 1982; Dwyer et al., 1985; Dwyer, Barnett & Lee, 1987; zie ook Blom, Krüger en Van Roozendaal, 1990; Krüger, Witziers, Slegers & Imants, 1999). In het model van Bossert et al. worden twee domeinen onderscheiden voor intermediaire factoren: het pedagogisch didactisch klimaat en

onderwijsorganisatie. Verder wordt rekening gehouden met drie categorieën van contextfactoren: de omgeving van de school, visie en ervaring van de schoolleider en wettelijke kaders en regelgeving (zie figuur 1.1). Voor de relaties die worden verondersteld in het model van Bossert c.s. bestaat al enige empirische ondersteuning (Hallinger et al., 1990; Heck, Larssen & Marcoulides, 1990; Heck, Marcoulides & Lang, 1991).

Het onderzoeksmodel is verder uitgewerkt aan de hand van het LOLSO (*Leadership for Organisational Learning and Student Outcomes*) model van Mulford en Silins (2003). In de eerste versie van het onderzoeksmodel zijn de schoolcultuur, de schoolorganisatiepraktijk, het werk van docenten, de betrokkenheid van leerlingen bij de school en het rendement (doorstroompercentage) opgenomen als intermediaire variabelen en de leerprestaties als uitkomstvariabele (zie figuur 2.5). Omdat dit onderzoek niet tot doel heeft om het gedrag van schoolleiders te verklaren, maar alleen om de invloed op de leerprestaties te verhelderen, wordt het gedrag van de schoolleiders als onafhankelijke variabele beschouwd. Dat betekent dat in het model alleen contextfactoren zijn opgenomen die van invloed zouden kunnen zijn op de leerprestaties. Het gaat hierbij om covariabelen zoals de samenstelling van de leerlingenpopulatie, schoolgrootte, stedelijkheid van de omgeving (zie bijv. Scheerens, 1990).

In het vorige hoofdstuk werden alleen ‘directe’ relaties tussen de onderzoeksvariabelen en de uitkomsten geëxploreerd. In dit hoofdstuk staat de volgende onderzoeksvraag centraal:

Hoe groot is het directe en het indirecte effect van schoolleidersgedrag op de uitkomstvariabelen en hoe verloopt deze invloed langs de intermediaire variabelen, gegeven de contextfactoren?

Om deze vraag te kunnen beantwoorden, is het onderzoeksmodel ingevuld met de onderzoeksvariabelen die in de voorgaande hoofdstukken zijn beschreven en relevante contextvariabelen. In paragraaf 6.2 wordt ingegaan op de invloed van contextfactoren op schooleffectiviteit. Paragraaf 6.3 bevat een beschrijving van het onderzoeksmodel en alle variabelen. Het onderzoeksmodel is getoetst aan de data met behulp van *structural equation modeling* in LISREL (zie bijlage 2B). In paragraaf 6.4 wordt eerst uiteengezet welke strategie werd gevolgd om te komen tot een acceptabel model. Daarna worden de uitkomsten gepresenteerd, waarbij wordt ingegaan op directe en indirecte effecten van schoolleiderschap en op de invloed van de intermediaire factoren op de uitkomsten. In paragraaf 6.5 wordt de invloed van contextvariabelen in het model voor de effecten van schoolleiderschap nader onderzocht. De conclusies worden besproken in paragraaf 6.5.

6.2 De invloed van contextfactoren op de leerprestaties

Contextfactoren die in verband kunnen worden gebracht met de leerprestaties zijn bijvoorbeeld: samenstelling van de leerlingenpopulatie, schoolgrootte, schooltype, schoolomgeving, onderwijssector en bestuursstructuur (zie bijv. Scheerens, 1989, 1990; Teddlie, Stringfield & Reynolds, 2000). Deze factoren worden hieronder toegelicht.

De samenstelling van de schoolpopulatie kan worden beschreven aan de hand van geaggregeerde leerlingkenmerken, zoals de gemiddelde sociaaleconomische status (SES) en het percentage leerlingen dat behoort tot culturele minderheden. De samenstelling van de leerlingenpopulatie heeft een substantieel effect op de uitkomsten (zie bijv. Luyten, Scheerens, Visscher, Maslowski, Witziers & Steen, 2005). Uit de resultaten van schooleffectiviteitsonderzoek kan worden opgemaakt dat door effectieve scholen

verschillende strategieën worden gehanteerd, afhankelijk van de SES samenstelling van de school (Teddlie, Stringfield & Reynolds, 2000). De Maeyer en Rymenans (2004, p.22) geven aan dat over het algemeen wordt vastgesteld dat leerlingen beter presteren op een school met een hogere gemiddelde sociaal economische status (Brandsma & Knuver, 1989; Martin, Mullis, Gregory, Hoyle & Shen, 2000; Willms & Raudenbusch, 1989). Overigens wordt wel verondersteld dat de effecten van het klimaat van de school sterk verweven zijn met de gemiddelde SES van de schoolpopulatie. Scholen met een beter klimaat hebben vaak ook een hoger SES niveau (Luyten et al., 2005). In Nederland wordt geregeld een negatief verband geconstateerd tussen het percentage leerlingen uit culturele minderheden en de gemiddelde leerprestaties op een school (Driessen, 2002; Veenstra, 1999). Hierbij speelt een combinatie van diverse achtergrondfactoren van leerlingen (SES, anderstaligheid) en schoolfactoren een rol (voor een overzicht zie Driessen, 2002). Op scholen met veel leerlingen uit achterstandsgroepen wordt bijvoorbeeld meer aandacht besteed aan basisvaardigheden en minder aan hogere orde vaardigheden, zoals informatiezoeken en studievoordigheden (Driessen, Doesborgh, Ledoux, van der Veen & Vergeer, 2003). Een gebrek aan hogere orde vaardigheden zou de leerprestaties negatief kunnen beïnvloeden.

Onderzoek naar het verband tussen de grootte van de school en de leerprestaties heeft tot nu toe geen duidelijke resultaten opgeleverd. Er is geen consensus over een optimale schoolgrootte (Bosker, 1992). De hypothese dat kleine scholen de leerprestaties zouden bevorderen, bijvoorbeeld door meer samenwerking en betrokkenheid van docenten, kan niet voldoende empirisch worden onderbouwd en de hypothese dat grote scholen een positieve invloed zouden hebben op de leerprestaties, bijvoorbeeld doordat meer faciliteiten beschikbaar zijn, ook niet (Luyten, 1994). Bij een analyse van zowel Nederlandse, als Zweedse en Amerikaanse data, ontleend aan een tweetal internationale onderzoeken, konden geen significante verbanden aangetoond worden tussen schoolgrootte en leerprestaties, wanneer rekening wordt gehouden met de achtergrond van de leerlingen (Luyten, 1994).

Uit de resultaten van PISA 2000 blijkt dat de schoolomgeving en het schooltype slechts in enkele landen een significant effect hebben op de leerprestaties (Luyten et al., 2005). Teddlie, Stringfield en Reynolds (2000) geven aan dat in het algemeen scholen in een sterk stedelijke omgeving karakteristieken vertonen van scholen met een laag gemiddeld SES niveau en scholen in een matig stedelijke omgeving karakteristieken van scholen met een middelmatig SES niveau. Scholen in landelijke gebieden geven een gevarieerd beeld. In Nederland zijn er bijvoorbeeld in niet-stedelijke gebieden meer autochtone achterstandsleerlingen dan in de grote steden (G. Ledoux, I. van der Veen, M. Vergeer, G. Driessen & J. Doesborgh, 2003). De Maeyer en Rymenans (2004, p.22) wijzen op twee Nederlandse onderzoeken waarin een verband werd gevonden tussen de urbanisatiegraad en respectievelijk de leerprestaties en het welbevinden van de leerlingen. In grote steden is het slaagpercentage kleiner en het welbevinden lager (Lugthart, Roeders, Bosker & Bos, 1989). Kleintjes en Kremers vonden in 1992 in Nederland geen samenhang tussen de grootte van de school en de leerprestaties, maar wel een verband tussen het schooltype en de leerprestaties voor bepaalde vakken. Voor Engels, biologie en vooral voor wiskunde werden op categorale scholen betere prestaties gehaald dan in scholengemeenschappen. Opendakker en Van Damme (2006) vonden in een onderzoek binnen het Vlaamse voortgezet onderwijs dat er een significant effect is van het schooltype (gebaseerd op het aangeboden opleidingen) op de leerresultaten voor het vak wiskunde, ook wanneer rekening werd gehouden met leerlingkenmerken.

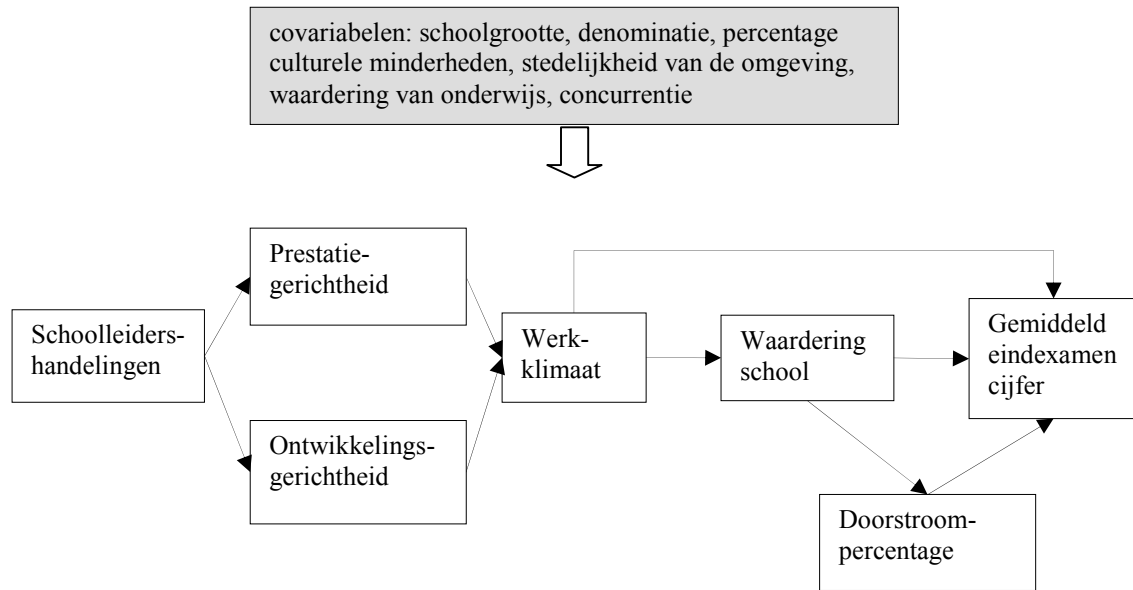
Er is relatief weinig onderzoek gedaan naar verschillen in schooleffectiviteit voor verschillende onderwijssectoren. Teddlie, Stringfield en Reynolds (2000) geven aan dat in het

algemeen het schooleffect groter lijkt te zijn in het basisonderwijs dan in het voortgezet onderwijs. Een verklaring hiervoor is, dat er in het voortgezet onderwijs sprake is van een groter aantal leidinggevend en een complexer curriculum dan in het basisonderwijs, naast andere contextverschillen. Uit het schoolleidersonderzoek blijkt ook dat de invloed van schoolleiders op de leerprestaties groter is in het basisonderwijs dan in het voortgezet onderwijs (Witziers, Bosker & Krüger, 2003).

Verschillen tussen openbare en bijzondere scholen en tussen scholen met een verschillende religieuze grondslag zijn volgens Teddlie, Stringfield en Reynolds (2000) vooral een gevolg van de mate waarin de scholen zeggenschap hebben over de gang van zaken binnen de school. Scholen met meer zeggenschap halen betere resultaten. Dronkers (2004) geeft aan dat er in Europese landen nauwelijks verschillen zijn te vinden in niet-cognitieve uitkomsten tussen openbare en bijzondere scholen. In Nederland was er volgens Dronkers (2004) van eind jaren zestig tot in de negentiger jaren sprake van een grotere effectiviteit van bijzonder onderwijs (katholiek of protestant) ten opzichte van het openbaar onderwijs, maar deze verschillen zouden inmiddels verdwenen zijn. Driessen en van der Slik (2001) vonden in het basisonderwijs in Nederland geen effect van denominatie op cognitieve en niet-cognitieve uitkomsten. Opendakker en Van Damme (2006) vonden in het Vlaamse voortgezet onderwijs een klein effect van denominatie op de leerresultaten dat verdween toen de achtergrond van leerlingen mede in beschouwing werd genomen. Zij geven aan dat er relaties bestaan tussen de samenstelling van de leerlingenpopulatie en de gang van zaken op school (Opendakker en Van Damme, 2001, 2006). Ook uit een analyse van de PISA 2000 data bleek dat de samenstelling van de leerlingenpopulatie, de schoolcontext, het schoolklimaat, het schoolbeleid en de beschikbare middelen niet helemaal onafhankelijk zijn van elkaar (Luyten et al., 2005).

6.3 Het onderzoeksmodel

In figuur 6.1 is het (tweede) concept model afgebeeld met alle onderzoeksvariabelen. Schoolleidersgedrag wordt gemeten door middel van de variabelen 'rationele doel handelingen', 'interne proceshandelingen', 'human relations' handelingen' en 'open systeem handelingen'. Rationele doel handelingen zijn gericht op het stellen van doelen, interne proces handelingen op het handhaven van orde en regels en het vastleggen van procedures, 'human relations' handelingen op ondersteuning van het werk van docenten en open systeem handelingen op externe contacten en het verwerven van middelen (zie paragraaf 4.3). De schoolcultuur en de schoolorganisatiepraktijk worden gemeten aan de hand van de variabelen 'prestatiegerichtheid' en 'ontwikkelingsgerichtheid' (zie paragraaf 4.7). Zoals blijkt uit figuur 6.1 wordt in het concept model verondersteld dat handelingen van de schoolleider direct invloed hebben op de prestatiegerichtheid en ontwikkelingsgerichtheid van de school. De variabele prestatiegerichtheid geeft aan in hoeverre waarde wordt gehecht aan hoge leerprestaties en een ordelijke werksfeer. De variabele ontwikkelingsgerichtheid meet in welke mate samenwerking, ontwikkeling en vernieuwing worden nagestreefd in de schoolorganisatie. In het model wordt ervan uitgegaan dat er verbanden kunnen bestaan tussen alle vier typen handelingen van schoolleiders en de twee cultuurvariabelen. Het ligt echter in de verwachting dat rationele doel en interne proces handelingen vooral samenhangen met prestatiegerichtheid, terwijl 'human relations' handelingen en open systeem handelingen met name verband houden met ontwikkelingsgerichtheid.



Figuur 6.1 Tweede versie van het onderzoeksmodel, met alle onderzoeksvariabelen.

Zowel de prestatiegerichtheid als de ontwikkelingsgerichtheid van de schoolorganisatie zouden het ‘werk-klimaat’ kunnen beïnvloeden. Het werk-klimaat wordt bepaald door de omgang van leerlingen met leraren, de ondersteuning door leraren bij het leerproces, de werksfeer in de klas en de organisatie van de vakken en de lessen (zie paragraaf 4.5). In het LOLSO-model van Mulford en Silins (2003, zie ook figuur 2.2) verloopt het effect van schoolleiderschap op het werk van docenten via *organisational learning*. Hierbij gaat het onder meer om: onderling vertrouwen, samenwerking, het nemen van initiatieven, een gedeelde missie en voortdurende professionele ontwikkeling (Mulford & Silins, 2003; Silins & Mulford, 2002). Deze aspecten zijn vooral terug te vinden in de ‘ontwikkelingsgerichtheid’ van de school en minder in ‘prestatiegerichtheid’. Leithwood en zijn collega’s vonden significante effecten van *total leadership* (van schoolleiders en anderen) op zowel de capaciteiten, als de motivatie en betrokkenheid van docenten en de perceptie van de werkomstandigheden, terwijl deze factoren significante effecten hebben op de leerprestaties (Leithwood et al., 2006b; Leithwood & Mascall, 2008). Deze laatste resultaten zouden erop kunnen wijzen, dat de invloed van de schoolleider vooral verloopt via de ontwikkelingsgerichtheid van de schoolorganisatie.

Het werk van docenten zou volgens het LOLSO-model invloed kunnen uitoefenen op de leerprestaties via identificatie van leerlingen met de school en participatie van leerlingen in schoolactiviteiten binnen en buiten de klas. Zowel door Mulford en Silins (2003), als door Leithwood en Jantzi (2000b) werd echter geconstateerd dat schoolleiderschap een groter effect heeft op identificatie dan op participatie. In het onderzoeksmodel in figuur 6.1 is identificatie van leerlingen met de school vertaald in de variabele ‘waardering van de school’. Deze variabele meet hoe leerlingen de sfeer op school en de omgang met de medeleerlingen ervaren (zie paragraaf 4.5). Omdat verondersteld mag worden dat invloed van het werk-klimaat voor de leerlingen niet alleen zal verlopen via waardering van de school, is in het onderzoeksmodel in figuur 6.1 zowel een direct als een indirect verband gelegd tussen werk-klimaat en leerprestaties.

De leerprestaties worden bepaald aan de hand van het gemiddelde cijfer voor alle vakken op het centrale schriftelijke eindexamen, gemiddeld over de jaren 2003-2005. Het gemiddelde

doorstroompercentage in de bovenbouw over drie jaar wordt als tweede uitkomstvariabele meegenomen in het onderzoek. De uitkomstvariabelen zijn nader beschreven in paragraaf 5.5.1. De veronderstelling dat de waardering van de school door leerlingen invloed heeft op het doorstroompercentage is gebaseerd op het LOLSO-model van Mulford en Silins (2003). Uit dit model is ook af te leiden dat het doorstroompercentage een effect zou moeten hebben op de leerprestaties.

In het concept onderzoeksmodel in figuur 6.1 zijn de volgende contextvariabelen opgenomen: schoolgrootte, denominatie, percentage culturele minderheden, stedelijkheid van de omgeving, waardering van onderwijs door leerlingen en de concurrentie die ervaren wordt van andere scholen. De contextfactoren schoolgrootte, denominatie, percentage culturele minderheden en stedelijkheid van de omgeving worden besproken in paragraaf 6.2. De contextfactoren schooltype en onderwijssector, die ook worden genoemd in paragraaf 6.2, zijn voor dit onderzoek niet relevant, omdat de deelnemende scholen allemaal vo-scholen zijn met een havo-bovenbouw.

Hoewel het niet zeker is dat er een verband bestaat tussen leerprestaties en de grootte of denominatie van de school (zie paragraaf 6.2), leek toch relevant om deze contextvariabelen mee te nemen. In dit onderzoek bleek namelijk dat rooms-katholieke en protestants-christelijke scholen prestatiegerichter zijn dan openbare en overige scholen (zie paragraaf 5.4). Verder bleken zowel prestatiegerichtheid als ontwikkelingsgerichtheid van scholen toe te nemen met het leerlingenaantal, tot een bepaalde schoolgrootte (tussen 1000 en 1500 leerlingen). Aangezien er een samenhang bestaat tussen de prestatiegerichtheid van een school en de uitkomsten (zie tabel 5.7) zou het weglaten van de contextvariabelen schoolgrootte en denominatie de uitkomsten kunnen beïnvloeden. In het LOLSO-model is er een negatief effect van schoolgrootte op schoolleiderschap en de participatie van leerlingen (Mulford & Silins, 2003). Krüger, Witziers en Slegers (2007) vonden negatieve effecten van de schoolgrootte op schoolleiderschap (*instructional leadership*, gericht op directe ondersteuning van docenten bij hun werk) en de uitkomsten (betrokkenheid van leerlingen). De schoolgrootte wordt uitgedrukt in het aantal leerlingen van de school, zoals opgegeven door de schoolleider. Gegevens over denominatie zijn afkomstig uit een databestand voor de opbrengstenkaart 2006 (beschikbaar gesteld door DANS¹⁸).

Het percentage 'culturele minderheden' is een belangrijke indicator voor de samenstelling van de leerlingenpopulatie. Hierbij gaat het om het aantal leerlingen waarvoor de school extra bekostiging ontvangt volgens de *Faciliteitenregeling Culturele Minderheden*. Tot en met 2006 werd dit percentage gebruikt als controlevariabele bij de berekening van vergelijkingsscores voor de opbrengstenkaarten (zie ook paragraaf 5.6). Vanaf 2007 wordt hiervoor het percentage leerlingen uit 'armoedeprobleemcumulatiegebieden' gebruikt. Deze indicator geeft beter weer dat het vooral gaat om de sociaaleconomische status van leerlingen. Aangezien in dit onderzoek gebruik werd gemaakt van gegevens voor de opbrengstenkaarten van 2004 tot en met 2006 (over de schooljaren 2002-2003 tot en met 2004-2005), is uitgegaan van het percentage cumi-leerlingen als contextfactor. Vanwege een trendbreuk in de percentages, is het gemiddelde percentage cumi-leerlingen in de havo-bovenbouw alleen berekend over de laatste twee jaar. Vanaf schooljaar 2003-2004 tellen leerlingen die langer dan acht jaar in Nederland zijn niet meer mee voor de 'cumi regeling'.

¹⁸ *Data Archiving and Networked Services*, een instituut van de KNAW en NWO

Het effect van de stedelijkheid van de omgeving zou gekoppeld kunnen zijn aan het gemiddelde SES-niveau van de leerlingen (Teddlie, Stringfield & Reynolds, 2000). Krüger, Witziers en Slegers (2007) vonden een negatief effect van stedelijkheid op zowel leiderschap als de uitkomsten (betrokkenheid van leerlingen). Gegevens over de stedelijkheid van de omgeving van de scholen die deelnamen aan het onderzoek zijn afgeleid uit demografische kerncijfers per gemeente, die zijn gepubliceerd door het CBS (CBS, 2006). Voor de indeling naar stedelijkheid zijn de numerieke waarden van de omgevingsadressendichtheid voor gemeenten gecategoriseerd in vijf klassen: van zeer sterk stedelijk naar niet stedelijk. In dit onderzoek is de volgorde van de indeling omgedraaid: van niet stedelijk (1) naar zeer sterk stedelijk (5)¹⁹.

De contextvariabele ‘waardering van onderwijs’ is toegevoegd als extra kenmerk van de leerlingenpopulatie. Deze variabele koppelt de motivatie van de leerling (het belang dat wordt gehecht aan onderwijs) aan de *family (home) educational culture* (onderwijsondersteuning van thuis, zie paragraaf 4.5). Volgens verschillende auteurs heeft de *family educational culture* een grotere voorspellende waarde voor schoolsucces dan de sociaaleconomische status (Leithwood, Aitken & Jantzi, 2000; Marzano, 2000, p. 77; Walberg, 1984). Mulford en Silins vonden in het LOLSO-project een negatief verband tussen de sociaaleconomische status van leerlingen en de perceptie van het werk van docenten, terwijl er tussen *home educational culture* en het werk van docenten een sterke positieve relatie bestond. Overigens had SES daarbij een directe relatie met de leerprestaties en *home educational culture* niet.

De laatste contextvariabele die in het model is opgenomen, is de concurrentie die wordt ervaren van andere scholen, volgens de opgave van de schoolleiders. De mate waarin concurrentie wordt ervaren zou iets kunnen zeggen over het imago van de school, omdat verwacht mag worden dat scholen die goed bekend staan minder concurrentie te duchten hebben. Daaraan gekoppeld zou de concurrentie met andere scholen gevolgen kunnen hebben voor de samenstelling van de leerlingenpopulatie, omdat minder selectief kan worden opgetreden. Tenslotte wordt verondersteld dat scholen die veel concurrentie ervaren, meer beleid maken dan andere scholen (Slegers, 1991, p. 144). De ervaren competitie werd in dit onderzoek gemeten aan de hand van vier categorieën: 8% van de schoolleiders gaf aan ‘geen competitie’ van andere scholen te ervaren, 24% ervaart ‘zwakke competitie’, 49% ‘sterke competitie’ en 19% ‘zeer sterke’ competitie (zie ook Schmidt, 2009).

6.4 Analyse van het onderzoeksmodel

Het onderzoeksmodel is geschat en getoetst met behulp van LISREL versie 8.8. Hierbij werd gebruik gemaakt van een procedure die is beschreven door Mels (2004) en aanvullende informatie van de website van *Scientific Software International*, de producent van LISREL (SSIcentral, z.j.). Bij deze procedure werden ruwe data gebruikt als *input*. LISREL produceerde op basis van de ruwe data een covariantiematrix. Er werd alleen een structureel model getoetst, omdat de grootte van de steekproef (97 scholen) te beperkt was om een compleet model met zowel een structureel gedeelte als een meetgedeelte (inclusief alle geobserveerde variabelen) te schatten (zie bijlage 2B). In het structurele model werden de leiderschapsvariabelen en de variabelen prestatiegerichtheid, ontwikkelingsgerichtheid, werkklimaat voor leerlingen, waardering van onderwijs, waardering van de school gerepresenteerd door geaggregeerde ‘latente variabele scores’ (zie hoofdstuk 4). De

¹⁹ 1 = niet stedelijk, minder dan 500 adressen per km², 2 = weinig stedelijk, 500 - 1000 adressen per km², 3 = matig stedelijk, 1000 tot 1500 adressen per km², 4 = sterk stedelijk, 1500-2500 adressen per km², 5 = zeer sterk stedelijk, 2500 adressen of meer per km².

toegepaste tweefase procedure, waarbij eerst de afzonderlijke meetmodellen worden geschat en daarna een structureel model, heeft als voordeel dat de modellen niet met elkaar kunnen interfereren. Door wederzijdse beïnvloeding zou een onjuist beeld kunnen ontstaan van zowel het meetgedeelte als het structurele gedeelte (Verschuren, 1991, p. 583).

Voor *structural equation modeling* moet de steekproef niet te klein zijn. Garson (z.j., b) geeft aan dat in het algemeen een steekproef van minder dan 100 cases ongeschikt wordt geacht. Verschillende auteurs (in Garson, z.j., b) adviseren een steekproefgrootte van tenminste 100 tot 200 cases. Met meer dan tien variabelen zou de steekproefgrootte meer dan 200 moeten bedragen. Een vuistregel is, dat voor elke predictor in het model tien tot twintig cases nodig zijn. Een andere vuistregel geeft aan dat er tenminste 15 cases per gemeten variabele zouden moeten zijn. In het onderzoeksmodel zitten 16 variabelen, inclusief de uitkomstvariabele(n), waaronder zeven gemeten variabelen. De steekproef van 97 scholen is dus eigenlijk te klein, gezien het aantal variabelen in het model.

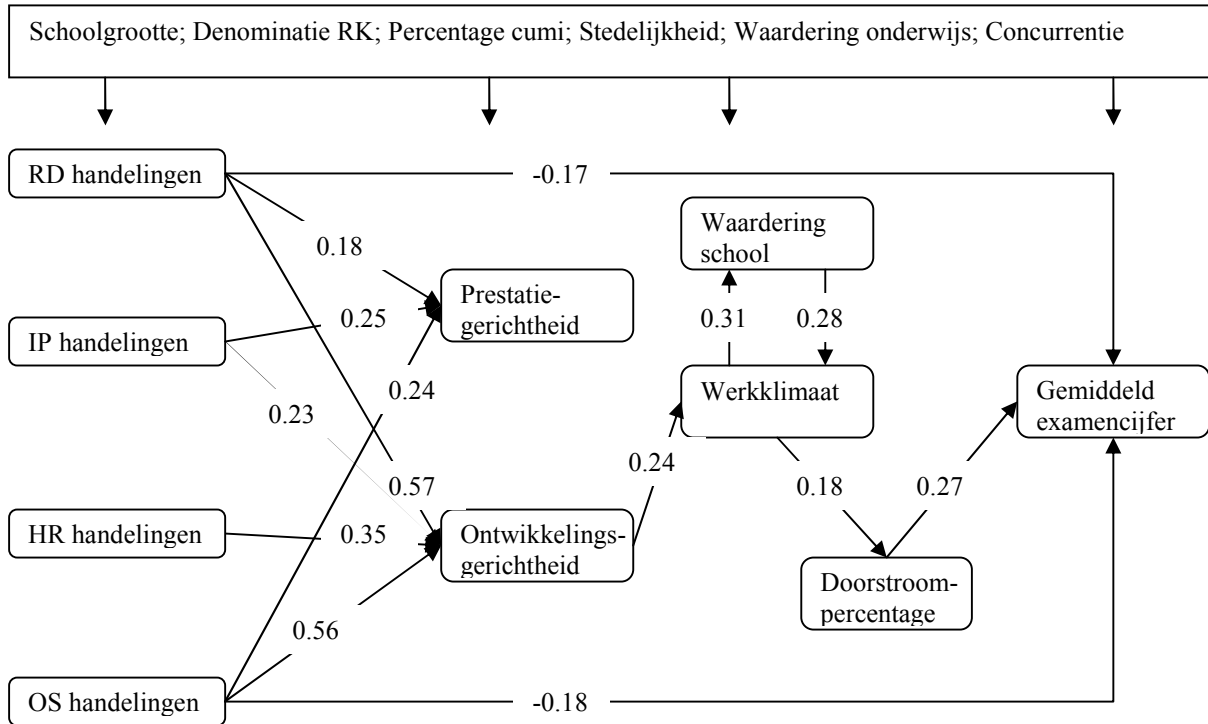
Vanwege de beperkte grootte van de steekproef is voor de toetsing van het onderzoeksmodel een fasegewijze strategie gehanteerd:

1. Eerst werd een model geconstrueerd voor de invloed van de contextvariabelen en de leiderschapsvariabelen op de 'prestatiegerichtheid' en de 'ontwikkelingsgerichtheid' van scholen.
2. Daarna werd een optimaal model vastgesteld voor de invloed van de contextvariabelen en de schoolbelevingsvariabelen 'werkklimaat' en 'waardering van de school' op het doorstroompercentage en het gemiddelde eindexamencijfer.
3. Tenslotte werden beide modellen gekoppeld.

Hierbij werd ervan uitgegaan dat de contextvariabelen van invloed zouden kunnen zijn op alle andere variabelen in het model. Voor de optimalisering van elk (deel-)model werden eerst niet significante relaties verwijderd ($\alpha = 0.05$). Daarna werden eventueel nog wijzigingen aangebracht op basis van de modificatie indices die LISREL genereerde, maar alleen wanneer deze aanpassingen plausibel leken op basis van theoretische overwegingen.

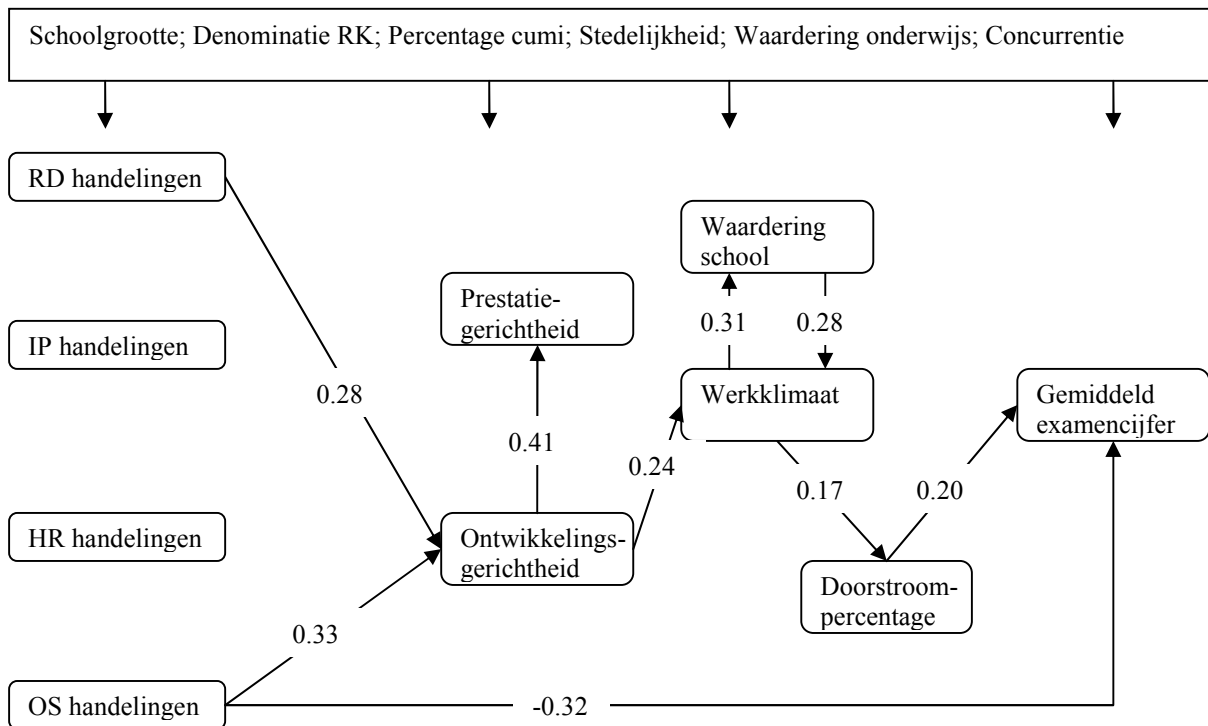
De hierboven beschreven aanpak voor de toetsing van het onderzoeksmodel werd twee maal uitgevoerd: eerst met gegevens van docenten over het schoolleidersgedrag dat zij waarnamen, daarna met gegevens van de schoolleiders zelf over de handelingen die zij uitvoerden. Een model met variabelen voor zowel door schoolleiders zelf opgegeven handelingen, als door docenten waargenomen handelingen was te complex om te testen, in verband met het beperkte aantal cases. Bovendien zou dit weinig nieuwe informatie opleveren, omdat de samenhang tussen de handelingen die schoolleiders zelf rapporteerden en door docenten waargenomen handelingen al eerder is onderzocht (zie paragraaf 4.7).

De causale modellen die met behulp van de bovenvermelde strategie werden gevonden voor de waargenomen handelingen (model 1) en voor de handelingen die schoolleiders zelf rapporteerden (model 2), zijn afgebeeld in respectievelijk de figuren 6.2 en 6.3. Bij elk model zijn de passingswaarden vermeld. Voor beide modellen voldoen deze waarden aan de criteria voor een goede modelpassing: RMSEA is kleiner dan .06, SRMR is kleiner dan .08 en CFI is groter dan .95 (zie bijlage 2F). Ook de verhouding tussen chi-kwadraat en df en de p-waarde duiden op een acceptabel model.



Chi-kwadraat = 62.44, df = 59, p-value = 0.35, RMSEA = 0.025, SRMR = 0.0662, CFI = 0.984

Figuur 6.2 Model 1: Effecten van door docenten waargenomen schoolleidersgedrag; alleen significante relaties zijn weergegeven.



Chi-kwadraat = 66.22, df = 73, p-value = 0.70, RMSEA = 0.000, SRMR = 0.0704, CFI = 1.00

Figuur 6.3 Model 2: Effecten van door schoolleiders gerapporteerd schoolleidersgedrag; alleen significante relaties zijn weergegeven

De effecten van de contextvariabelen op de andere onderzoeksvariabelen in bovenstaande modellen zijn niet afgebeeld, omdat de figuren dan te onoverzichtelijk zouden worden. Deze effecten komen later aan de orde.²⁰

Uit model 1 in figuur 6.2 is op te maken dat er een significante en positieve samenhang bestaat tussen waargenomen rationele doel (RD), interne proces (IP) en open systeem (OS) handelingen en de prestatiegerichtheid van scholen. Er is geen verband tussen waargenomen 'human relations' (HR) handelingen en de prestatiegerichtheid. Verder zijn er volgens model 1 significante positieve relaties tussen alle vier typen van schoolleidersgedrag en de ontwikkelingsgerichtheid van scholen. De samenhang van rationele doel en open systeem handelingen met de ontwikkelingsgerichtheid is groter dan de samenhang met prestatiegerichtheid. Door schoolleiders zelf gerapporteerde handelingen in model 2 (figuur 6.3) vertonen geen verband met de prestatiegerichtheid, maar wel met de ontwikkelingsgerichtheid van scholen. In dit model hangen alleen de rationele doel en open systeem handelingen samen met de ontwikkelingsgerichtheid. Daarnaast is er een effect van ontwikkelingsgerichtheid op prestatiegerichtheid.

In het rechterdeel van de modellen 1 en 2 wordt de invloed van de schoolbelevingsvariabelen op de uitkomsten weergegeven. Deze invloed verloopt anders dan in het concept model (figuur 6.1) werd verondersteld. Er is geen direct effect van de waardering van de school op de uitkomstvariabelen, maar wel direct effect van het werkklimaat in de klassen op het doorstroompercentage en een wederkerig effect tussen het werkklimaat en de waardering van de school. Het doorstroompercentage heeft net als in het concept model een direct effect op het gemiddelde eindexamencijfer.

Opvallend is dat er zowel in model 1 als in model 2 een significant en negatief *direct* verband is tussen open systeem handelingen van de schoolleider en de leerprestaties. In model 1 is er bovendien een negatief direct effect van rationele doel handelingen op de leerprestaties. Deze relaties werden al gesignaleerd bij de exploratie van correlaties tussen de onderzoeksvariabelen en de uitkomstvariabelen in paragraaf 5.5.2 (zie tabel 5.7), evenals de relaties tussen het doorstroompercentage en de ontwikkelingsgerichtheid van de school, de waardering van de school door leerlingen en het werkklimaat zoals dat wordt ervaren door leerlingen. Verder is er in de modellen 1 en 2 geen (significant) verband tussen prestatiegerichtheid en de uitkomstvariabelen, zoals ook bij de analyse van directe relaties werd vastgesteld, wanneer werd gecontroleerd voor de samenstelling van de leerlingenpopulatie (zie tabel 5.7). Een gerichtheid op leerprestaties, gecombineerd met het hebben van hoge verwachtingen van leerlingen, wordt in het algemeen gezien als een effectiviteitsbevorderende factor (Scheerens & Bosker, 1997). Er wordt echter ook wel verondersteld dat de prestatiegerichtheid van een school (voor een deel) kan worden verklaard door de samenstelling van de leerlingenpopulatie (zie bijv. Luyten et al., 2005). In de hier gepresenteerde modellen zou het veronderstelde effect van prestatiegerichtheid op de leerprestaties 'verdwenen' kunnen zijn doordat relevante covariabelen zijn opgenomen.

In hoeverre kunnen nu met de twee modellen die zijn gevonden, *indirecte* effecten van schoolleidersgedrag, via intermediaire variabelen, op de uitkomstvariabelen worden aangetoond? In tabel 6.1 staat een overzicht van de indirecte effecten van schoolleidersgedrag

²⁰ Bij de toetsing van de structurele modellen zijn de latente variabelen feitelijk opgevat als geobserveerde (gemeten) variabelen. In model 1 is uitgegaan van (error) covariantie tussen de scores voor alle vier typen handelingen van schoolleiders. Bij model 2 is uitgegaan van (error) covariantie tussen de scores voor rationele doel en open systeem handelingen en tussen de scores voor interne proces en 'human relations' handelingen.

op het doorstroompercentage en het gemiddelde eindexamencijfer, voor beide modellen. Geen van de effecten is significant wanneer wordt uitgegaan van een kritieke waarde van 1.96 ($\alpha = 0.05$), maar bij een lager significantieniveau ($\alpha = 0.1$, kritieke waarde 1.65) zijn in model 1 de indirecte effecten van rationele doel (RD), 'human relations' (HR) en open systeem (OS) handelingen op het doorstroompercentage wel significant. Het lagere significantieniveau kan worden toegepast vanwege de relatief kleine steekproef. De gevonden effecten zijn klein.

Tabel 6.1 *Gestandaardiseerde indirecte effecten van schoolleidersgedrag en schoolcultuur op werkklimaat in de klassen, doorstroompercentage en gemiddeld eindexamencijfer.*

Model ¹	RD handelingen		IP handelingen		HR handelingen		OS handelingen	
	1	2	1	2	1	2	1	2
Doorstroompercentage	.026*	.013	.010		.016*		.026*	.015
Gemiddeld examencijfer	.007	.002	.003		.004		.007	.003

¹ model 1 bevat gegevens van docenten over schoolleidersgedrag en model 2 van schoolleiders.

* significant bij $\alpha = 0.1$

Tabel 6.2 geeft een overzicht van de totale (directe en indirecte) effecten die worden uitgeoefend door de leiderschapsvariabelen op de intermediaire en uitkomstvariabelen. Uit deze tabel blijkt (zoals ook zichtbaar is in de figuren 6.2 en 6.3) dat er een grotere samenhang bestaat tussen de handelingen die worden waargenomen door de docenten en de intermediaire variabelen prestatiegerichtheid en ontwikkelingsgerichtheid (model 1), dan tussen handelingen die schoolleiders zelf rapporteren en de cultuurvariabelen (model 2). Dit ligt voor de hand, omdat de gegevens over de schoolcultuur afkomstig zijn van docenten. De perceptie van de handelingen van schoolleiders door docenten staat waarschijnlijk niet helemaal los van de mening van docenten over de schoolcultuur. Overigens vonden Kim en Yukl (1995) bijvoorbeeld dat de effectiviteit van leiders beter gerelateerd kon worden aan de perceptie van ondergeschikten dan aan door leiders zelf gerapporteerd gedrag. Model 1 is dus wellicht toch het meest informatief.

Tabel 6.2 *Gestandaardiseerde totale effecten van schoolleidersgedrag op intermediaire en uitkomstvariabelen*

Model ¹	RD handelingen		IP handelingen		HR handelingen		OS handelingen	
	1	2	1	2	1	2	1	2
Prestatiegerichtheid	.18**	.11**	.25**		-		.24**	.14**
Ontwikkelingsgerichtheid	.57**	.28**	.23**		.35**		.56**	.33**
Werkklimaat	.15**	.07**	.06**		.09**		.15**	.09**
Waardering school	.05*	.02	.02		.03*		.05*	.03
Doorstroompercentage	.03*	.01	.01		.02*		.03*	.02
Gemiddeld examencijfer	-.16**	.00	.00		.00		-.18**	-.31**

¹ model 1 bevat gegevens van docenten over schoolleidersgedrag en model 2 van schoolleiders.

* significant bij $\alpha = 0.1$

** significant bij $\alpha = 0.05$

Uit tabel 6.2 is verder op te maken dat er significante en positieve indirecte relaties bestaan tussen de vier typen handelingen van schoolleiders en het werkklimaat zoals dat wordt ervaren door de leerlingen. Rationele doel (RD) en open systeem (OS) handelingen vertonen een grotere samenhang met het werkklimaat dan interne proces (IP) en 'human relations' (HR) handelingen, via de intermediaire variabele ontwikkelingsgerichtheid. De kleine significante ($\alpha = 0.1$) en positieve effecten van rationele doel, 'human relations' en open systeem handelingen op het doorstroompercentage en de waardering van de school verlopen

via de intermediaire variabelen ontwikkelingsgerichtheid en werkklimaat (zie figuur 6.2 en 6.3). Het totale effect van zowel rationele doel als open systeem handelingen op het gemiddelde eindexamencijfer is opgebouwd uit een substantieel negatief direct effect en een klein positief indirect effect (zie tabel 6.1).

Tabel 6.3 geeft een overzicht van de totale effecten van de intermediaire variabelen in het causale model. Vanwege de overeenkomsten tussen de twee modellen zijn alleen de uitkomsten van model 1 weergegeven. De directe effecten van ontwikkelingsgerichtheid op het werkklimaat voor leerlingen en van het werkklimaat op zowel het doorstroompercentage als de waardering van de school, zijn hierboven al genoemd.

Tabel 6.3 Gestandaardiseerde totale effecten van intermediaire variabelen.

	Ontwikkelings- Gerichtheid	Werkklimaat	Waardering school
Werkklimaat	.26**	.	.31**
Waardering school	.08*	.34**	
Doorstroompercentage	.05*	.19**	.05
Gemiddeld examencijfer	.01	.05*	.02

* significant bij $\alpha = 0.1$

** significant bij $\alpha = 0.05$

Uit tabel 6.3 is af te leiden dat er ook nog een klein indirect positief effect is van ontwikkelingsgerichtheid, via het werkklimaat voor leerlingen, op zowel de waardering van de school als het doorstroompercentage ($\alpha = 0.1$). Verder blijkt er een klein maar significant indirect effect te zijn van het werkklimaat op het gemiddelde eindexamencijfer, via het doorstroompercentage. Een beter werkklimaat voor leerlingen heeft dus in eerste instantie een positief effect op het doorstroompercentage en in tweede instantie op de hoogte van de examencijfers.

De contextvariabelen hebben aanzienlijke effecten op de andere onderzoeksvariabelen, zoals blijkt uit de totale effecten (de som van alle directe en indirecte effecten) in tabel 6.4. In deze tabel zijn alleen de uitkomsten voor model 1 weergegeven.

Tabel 6.4 Gestandaardiseerde totale effecten van contextvariabelen (in model 1).

	School- grootte	Denominatie RK	Percentage cumi	Stedelijkheid	Waardering Onderwijs	Concurrentie
RD handelingen		.16**		.25**	.13**	-.13**
HR handelingen	-.12**	.19**				
OS handelingen	.25**			.28**		
Prestatiegerichtheid	.06**	.03*		.11**	.02**	-.35**
Ontwikkelingsgerichtheid	.26**	-.03		.30**	.08**	-.33**
Werkklimaat	.07**	-.01	-.10*	.08**	.50**	-.18**
Waardering school	.02	-.00	-.35**	.02	.39**	-.36**
Doorstroompercentage	.01	.20**	-.34**	.01	.09**	-.34**
Gemiddeld examencijfer	-.04*	.03	-.37**	-.09**	.00	-.07*

* significant bij $\alpha = 0.1$

** significant bij $\alpha = 0.05$

De effecten van de contextvariabelen in model 2 zijn vergelijkbaar, alleen is er in model 2 geen invloed van denominatie en concurrentie op de handelingen van schoolleiders. Dit zou er weer op kunnen wijzen dat de perceptie van docenten van schoolleidersgedrag sterker is gekoppeld aan de cultuur van de school, dan de perceptie van schoolleiders zelf (zie boven).

Interne proces handelingen ontbreken in tabel 6.4. Deze handelingen worden niet beïnvloed door de contextvariabelen die in het model zijn opgenomen.

Uit tabel 6.4 is op te maken dat vooral de concurrentie die wordt ervaren van andere scholen en het percentage culturele minderheden negatief samenhangen met schoolkenmerken en uitkomsten. De schoolgrootte heeft een lichte negatieve invloed op ‘human relations’ handelingen van de schoolleider en het gemiddelde examencijfer, een matige positieve invloed op open systeem handelingen en de ontwikkelingsgerichtheid van scholen en een kleine positieve invloed op de prestatiegerichtheid en het werkklimaat voor leerlingen. De contextvariabele ‘denominatie RK’ geeft aan dat alleen de rooms-katholieke denominatie een substantiële invloed heeft in het model. Het model is voor elke denominatie (openbaar, rooms-katholiek, protestants-christelijk en overig) apart getest. De rooms-katholieke denominatie vertoont een matige positieve samenhang met rationele doel en ‘human relations’ handelingen van de schoolleider en met het doorstroompercentage en verder een kleine samenhang met de prestatiegerichtheid. De mate van stedelijkheid hangt vooral positief samen met rationele doel en open systeem handelingen en de ontwikkelingsgerichtheid van de school en in mindere mate met de prestatiegerichtheid en het werkklimaat in de klassen. Er is een lichte negatieve samenhang tussen stedelijkheid en het gemiddelde examencijfer. De mate waarin leerlingen gemotiveerd zijn en vanuit thuis worden ondersteund (‘waardering onderwijs’) heeft vooral een positieve invloed op de beleving van het werkklimaat in de klas en de waardering van de school en een licht positief effect op de prestatiegerichtheid, de ontwikkelingsgerichtheid en het doorstroompercentage.

Wanneer de tabellen 6.2, 6.3 en 6.4 worden vergeleken blijkt dat de variabelen prestatiegerichtheid en ontwikkelingsgerichtheid vooral samenhangen met schoolleidersgedrag terwijl de schoolbeleving en de uitkomsten voornamelijk samenhangen met de context. Dit is ook te zien in tabel 6.5. Deze tabel geeft een overzicht van de mate waarin de variantie in de verschillende onderzoeksvariabelen wordt verklaard door respectievelijk alleen (directe) effecten van contextvariabelen of effecten van alle variabelen, voor de twee modellen die hierboven zijn besproken.

Tabel 6.5 *Percentage verklaarde variantie, door directe effecten van contextvariabelen of het volledige model, in model 1 (handelingen gerapporteerd door docenten) en model 2 (handelingen gerapporteerd door schoolleiders).*

	Contextvariabelen		Volledig model	
	model 1	model 2	model 1	model 2
RD handelingen	8	12	8	12
HR handelingen	5	8	5	8
OS handelingen	14	17	14	17
Prestatiegerichtheid	9	6	26	24
Ontwikkelingsgerichtheid	9	5	95	25
Werkklimaat	27	27	49 ¹	49 ¹
Waardering school	44	44	59	59
Doorstroompercentage	37	37	40	40
Gemiddeld examencijfer	15	15	28 ²	30 ²

¹ Alleen context + ontwikkelingsgerichtheid: Model 1: $R^2 = 0.33$; Model 2: $R^2 = 0.32$

² Alleen context + handelingen: Model 1: $R^2 = 0.22$; Model 2: $R^2 = 0.28$

Uit tabel 6.5 kan worden afgeleid dat open systeem handelingen meer worden beïnvloed door contextvariabelen dan de overige handelingen. Dat is plausibel omdat het bij open systeem handelingen juist gaat om het reageren op externe omstandigheden. Interne proces

handelingen worden daarentegen helemaal niet beïnvloed door contextvariabelen. Dat is ook verklaarbaar, omdat interne proces handelingen gericht zijn op regulering van de interne organisatie, hierbij is de context minder van belang. In beide modellen wordt iets minder dan 20% van de variantie in prestatiegerichtheid verklaard door schoolleidersgedrag (het verschil tussen het percentage verklaarde variantie in het hele model en alleen door de contextvariabelen). Voor ontwikkelingsgerichtheid lopen de percentages verklaarde variantie sterk uiteen: in model 1 wordt 86% van de variantie in de variabele 'ontwikkelingsgerichtheid' verklaard door de handelingen van schoolleiders en in model 2 niet meer dan 20%. De mate waarin de handelingen van schoolleiders worden waargenomen door docenten hangt kennelijk sterk samen met de (perceptie van de) ontwikkelingsgerichtheid van de school. Het werkklimaat in de klassen wordt vooral beïnvloed door contextvariabelen en de waardering van de school. Slechts zo'n 6% van de variantie in werkklimaat wordt verklaard door de ontwikkelingsgerichtheid van de school (zie ook tabel 6.3). De waardering van de school wordt vooral bepaald door contextvariabelen en het werkklimaat in de klas. Ongeveer 3% van de variantie in het doorstroompercentage wordt verklaard door het werkklimaat. Het doorstroompercentage verklaart 2-6% van de variantie in het gemiddelde examencijfer, terwijl 7-13% wordt verklaard door schoolleidersgedrag.

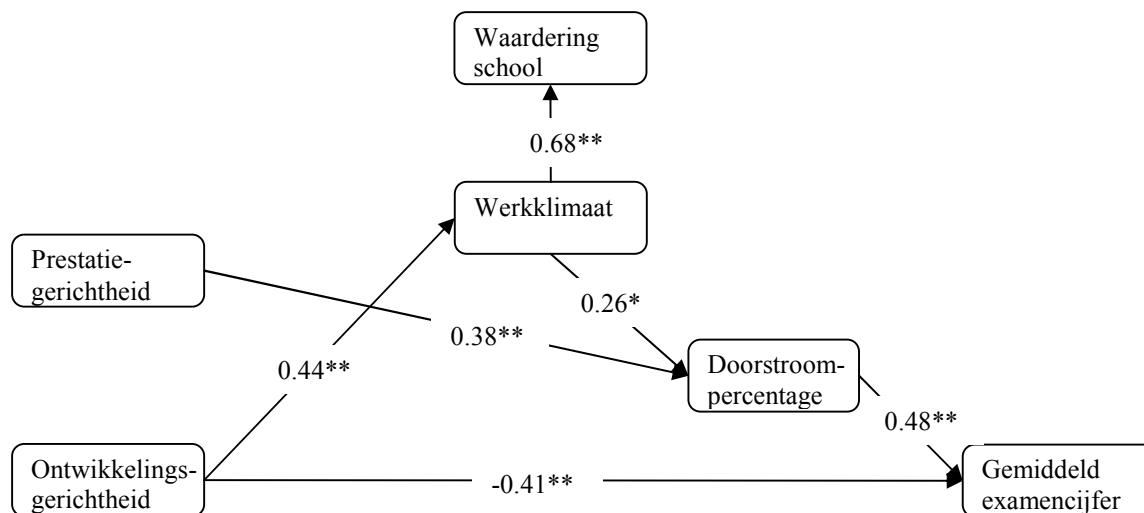
Het relatief hoge percentage van de variantie in het gemiddelde examencijfer dat wordt verklaard door schoolleidersgedrag, is vooral een gevolg van de substantiële negatieve samenhang tussen schoolleidersgedrag en het eindexamencijfer. Daarnaast bestaat echter ook een zeer klein maar significant positief effect van de handelingen van schoolleiders, via de ontwikkelingsgerichtheid van de school en het werkklimaat voor leerlingen, op het doorstroompercentage in de bovenbouw (zie tabel 6.1).

Hoewel een effect van schoolleidersgedrag op het gemiddelde examencijfer, zoals in de hierboven beschreven modellen (figuur 6.2 en 6.3), een betere modelpassing bleek op te leveren dan een omgekeerd effect van de uitkomsten op de schoolleidersgedrag (model niet afgebeeld), is het aannemelijk dat het toch gaat om een reactie van schoolleiders op de leerresultaten. In eerder onderzoek werden namelijk ook negatieve verbanden gevonden tussen de leerprestaties (of de betrokkenheid van leerlingen) en activiteiten van de schoolleider die gericht zijn op het ontwikkelen en verbeteren van de school (zie bijv. Krüger, Witziers & Slegers, 2007; Teddlie & Stringfield, 1993; Witziers, Bosker & Krüger, 2003). De richting van het negatieve effect in de hier gepresenteerde modellen kan worden verklaard, wanneer wordt aangenomen dat hoge rationele doel en open systeem activiteiten van schoolleiders vaker samengaan met lage leerprestaties dan omgekeerd. Waarschijnlijk reageren niet alle schoolleiders op lage leerprestaties door meer rationele doel en open systeem handelingen uit te voeren. Aan de andere kant kan de gevonden richting van het effect natuurlijk ook betekenen dat schoolleiders van scholen met goede leerprestaties niet altijd een beperkte rationele doel en open systeem activiteit vertonen (vgl. tabel 5.8).

6.5 De invloed van contextvariabelen nader onderzocht

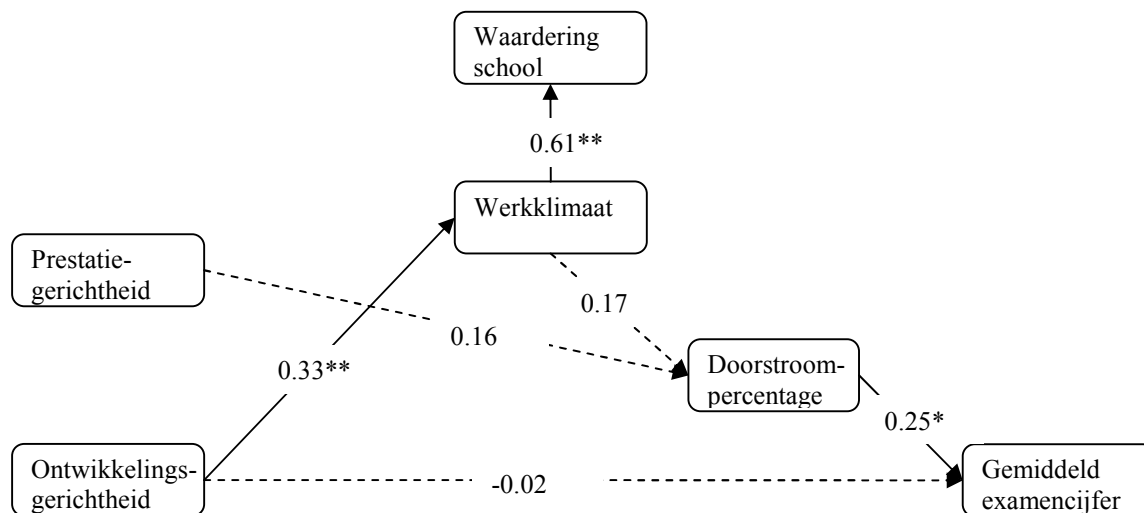
Uit een onderzoek van Krüger, Witziers en Slegers (2007) naar het effect van schoolleiderschap op schoolkenmerken en opbrengsten, bleek dat de wijze waarop de effecten van contextvariabelen in het model werden verdisconteerd, van grote invloed was op het resultaat. Krüger en haar collega's vonden dat effecten van de schoolcultuur en de schoolorganisatie op de betrokkenheid van leerlingen verdwenen, zodra rekening werd gehouden met de invloed van een aantal contextvariabelen, zoals de grootte van de school, mate van stedelijkheid van de omgeving en het aantal leerlingen uit een milieu met een lage

sociaal economische status. Om na te gaan wat de impact is van de contextvariabelen in dit onderzoek, is een deel van het model opnieuw getest, zonder contextvariabelen, voor respectievelijk rooms-katholieke scholen (n = 39) en 'niet rooms-katholieke' scholen (n = 58). De uitkomsten staan in de figuren 6.4 en 6.5.



Chi-kwadraat = 6.21, df = 8, p-value = 0.62, RMSEA = 0.000, SRMR = 0.10, CFI = 1.00
 * significant bij $\alpha = 0.1$ ** significant bij $\alpha = 0.05$

Figuur 6.4 Effecten van prestatiegerichtheid en ontwikkelingsgerichtheid voor rooms-katholieke scholen in de onderzoeksgroep (n=39)



Chi-kwadraat = 15.64, df = 8, p-value = 0.05, RMSEA = 0.129, SRMR = 0.11, CFI = 0.84
 * significant bij $\alpha = 0.1$ ** significant bij $\alpha = 0.05$

Figuur 6.5 Effecten van prestatiegerichtheid en ontwikkelingsgerichtheid voor niet rooms-katholieke scholen in de onderzoeksgroep (n = 58)

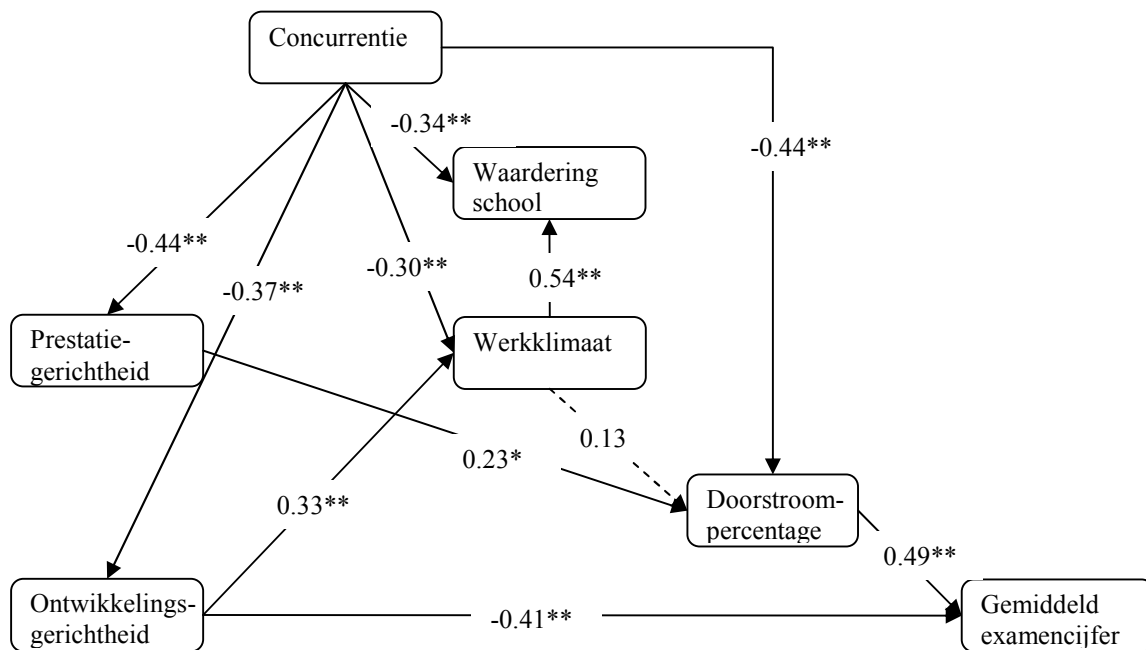
Er is gekozen voor een verdeling in katholieke en niet-katholieke scholen, omdat gebleken is dat de rooms-katholieke denominatie een grotere invloed heeft op de onderzoeksvariabelen dan de andere denominaties (zie paragraaf 6.4) en omdat katholieke scholen het grootste deel uitmaken van de onderzoeksgroep (40%, zie paragraaf 3.4). Het onderzoeksmodel werd verder vereenvoudigd door de leiderschapsvariabelen weg te laten, omdat de samenhang tussen leiderschap en de cultuurvariabelen voldoende duidelijk was. Vereenvoudiging van het model was bovendien nodig vanwege de opsplitsing van de onderzoeksgroep, waardoor het aantal scholen per groep te beperkt werd om een complex structureel model te kunnen analyseren. Uitgaande van het model zoals afgebeeld in figuur 6.2, is eerst een optimaal model vastgesteld voor de katholieke scholen, door het verwijderen van niet significante relaties en het aanbrengen van nieuwe relaties op basis van de modificatie indices van LISREL (zie ook bijlage 1D). Het model voor katholieke scholen is vervolgens getoetst aan data van niet-katholieke scholen.

Het model voor katholieke scholen in figuur 6.4 heeft een acceptabele modelpassing (zie bijlage 2D): RMSEA is kleiner dan .06 en CFI is groter dan .95, alleen SRMR is aan de hoge kant (SRMR moet kleiner zijn dan .08). Uit figuur 6.5 blijkt dat hetzelfde model geen goede modelpassing heeft voor niet katholieke scholen. Wanneer de niet-significante relaties verwijderd worden uit dit model, verbetert de modelpassing: Chi-kwadraat = 6.39, $df = 6$, $p\text{-value} = 0.38$, RMSEA = 0.034, SRMR = 0.07, CFI = 0.99 (model niet afgebeeld).

Uit de figuren 6.4 en 6.5 kan worden afgeleid dat, wanneer de invloed van contextvariabelen buiten beschouwing wordt gelaten, er voor katholieke scholen een significante en positieve relatie is tussen prestatiegerichtheid en het doorstroompercentage, terwijl deze relatie voor de overige scholen niet significant is. De relatie tussen prestatiegerichtheid en het doorstroompercentage is ook in het volledige model in figuur 6.2 niet zichtbaar. Dit maakt het aannemelijk dat het verband wordt ‘weggefilterd’ door de contextvariabelen die zijn opgenomen in het volledige model.

De figuren 6.4 en 6.5 laten verder zien dat de verbanden tussen ontwikkelingsgerichtheid en het werkklimaat voor leerlingen, tussen het werkklimaat en het doorstroompercentage en tussen het doorstroompercentage en het eindexamencijfer groter zijn voor katholieke scholen dan voor niet katholieke scholen, als de contextfactoren buiten beschouwing worden gelaten. Verder is er voor katholieke scholen een negatief verband tussen ontwikkelingsgerichtheid en het gemiddelde examencijfer en voor de overige scholen niet. Deze negatieve relatie is in het volledige model in figuur 6.2 waarschijnlijk terug te vinden in het negatieve verband tussen rationele doel en open systeem handelingen en het examencijfer, omdat deze handelingen sterk samenhangen met de ontwikkelingsgerichtheid van een school (zie paragraaf 6.4). Kennelijk geldt de negatieve samenhang tussen de handelingen van schoolleiders en het eindexamencijfer vooral voor katholieke scholen.

Om na te gaan wat de invloed is van de context op de relatie tussen prestatiegerichtheid en het doorstroompercentage, is de variabele ‘concurrentie’ toegevoegd aan het model in figuur 6.4. Uit tabel 6.4 is af te leiden dat de concurrentie die wordt ervaren van andere scholen een substantiële negatieve samenhang vertoont met zowel de prestatiegerichtheid als de ontwikkelingsgerichtheid van scholen. Het aangepaste model is afgebeeld in figuur 6.6.



Chi-kwadraat = 9.65, df = 10, p-value = 0.47, RMSEA = 0.000, SRMR = 0.08, CFI = 0.99

* significant bij $\alpha = 0.1$ ** significant bij $\alpha = 0.05$

Figuur 6.6 Effecten van prestatiegerichtheid en ontwikkelingsgerichtheid voor rooms-katholieke scholen in de onderzoeksgroep, onder invloed van de concurrentie van andere scholen (n=39)

Uit figuur 6.6 blijkt dat voor katholieke scholen het verband tussen prestatiegerichtheid en het doorstroompercentage, evenals het verband tussen ontwikkelingsgerichtheid en het werkklimaat voor leerlingen, kleiner wordt, wanneer rekening wordt gehouden met de concurrentie van andere scholen. Het negatieve verband tussen ontwikkelingsgerichtheid en het eindexamencijfer blijft gelijk. Het is dus aannemelijk dat de positieve relatie tussen de prestatiegerichtheid van scholen en het doorstroompercentage feitelijk wordt veroorzaakt door contextfactoren. De vraag is nu welke verschillen er zijn tussen katholieke en niet katholieke scholen en waar deze verschillen toe leiden. Tabel 6.6 geeft een overzicht van de gemiddelde waarden van contextvariabelen en uitkomstvariabelen voor katholieke scholen en overige scholen, met een indicatie van de mate waarin deze waarden verschillen (F-toets).

Tabel 6.6 Verschillen in schoolkenmerken en uitkomsten voor tussen rooms-katholieke (n = 39) en overige scholen (n = 58)

	RK scholen	Overige scholen	F	p
Schoolgrootte	1181	1123	.358	.551
Percentage cumi school ¹	4.6	12.2	5.064	.027
Percentage cumi havo bovenbouw	.97	1.90	2.460	.120
Stedelijkheid	3.18	3.50	2.316	.131
Concurrentie	2.54	2.91	4.710	.032
Doorstroompercentage	67.8	61.8	10.154	.002
Gemiddeld eindexamencijfer	6.34	6.23	8.108	.005

¹ naar inschatting van de schoolleider (voor 91 scholen)

Uit de tabel blijkt dat het gemiddelde doorstroompercentage voor katholieke scholen significant hoger is dan voor andere scholen en ook het gemiddelde eindexamencijfer is hoger. De gemiddelde schoolgrootte blijkt niet significant te verschillen tussen katholieke en overige scholen, evenmin als het percentage culturele minderheden in de havo bovenbouw, of de ‘gemiddelde stedelijkheid’ van de omgeving. De gemiddelde stedelijkheid geeft echter een weinig genuanceerd beeld: van de katholieke scholen ligt een kleiner percentage scholen (44%) in sterk of zeer sterk stedelijk gebied dan van de overige scholen (60%). In het algemeen geldt dat in grote steden de leerprestaties en het welbevinden van leerlingen lager zijn (zie bijv. Lugthart Roeders, Bosker & Bos, 1989). Het percentage culturele minderheden binnen een school heeft daarbij een negatief effect op de uitkomstvariabelen (zie tabel 6.4). Voor katholieke scholen is het gemiddelde percentage culturele minderheden in de gehele school significant lager dan voor overige scholen en de concurrentie die wordt ervaren van andere scholen is minder. Van de katholieke scholen ondervindt 51% sterke of zeer sterke concurrentie van andere scholen, tegen 76% van de overige scholen. De veronderstelling dat de verschillen tussen katholieke en overige scholen vooral worden veroorzaakt door de samenstelling van de leerlingenpopulatie, lijkt hiermee te worden ondersteund. Zowel de locatie van de school, als het imago zouden kunnen leiden tot een geselecteerde leerlingenpopulatie. Katholieke scholen die veel concurrentie ondervinden van andere scholen hebben minder gunstige schoolkenmerken dan katholieke scholen die weinig concurrentie ervaren (zie figuur 6.6).

Uit de bovenstaande exploratie kan worden opgemaakt dat de contextvariabelen in de modellen die zijn gepresenteerd in de vorige paragraaf, inderdaad relevant zijn om het effect van schoolleidersgedrag op de opbrengsten van scholen goed in kaart te kunnen brengen.

6.6 Discussie

Dit hoofdstuk geeft aan de hand van een aantal modellen inzicht in de relaties tussen schoolleidersgedrag, de prestatiegerichtheid en ontwikkelingsgerichtheid van scholen, de schoolbeleving van leerlingen en de uitkomsten, onder invloed van relevante contextvariabelen. Uit de resultaten blijkt dat de keuzes die gemaakt worden bij de opzet van een model voor schoolleiderschap van groot belang zijn. Dit werd eerder ook door andere onderzoekers aangetoond. In het algemeen wordt gesteld dat directe effectmodellen minder geschikt zijn om effecten van schoolleiderschap aan te tonen, dan indirecte effectmodellen (zie bijv. Bosker, de Vos & Witziers, 2000; De Maeyer & Rymenans, 2007; Hallinger & Heck, 1998; Krüger, Witziers & Bosker, 2007; Witziers, Bosker & Krüger, 2003). Daarbij wordt ook het belang van het opnemen van contextvariabelen in het model onderstreept (zie paragraaf 6.2).

In dit onderzoek is een indirect effectmodel gehanteerd om de effecten van de handelingen van schoolleiders te onderzoeken (zie paragraaf 6.3). Bij de analyse van het structurele model is een *model generating* benadering toegepast, dat betekent dat het model werd bijgesteld aan de hand van de uitkomsten (zie bijlage 2C). Uiteindelijk zijn twee modellen geconstrueerd, respectievelijk voor handelingen die werden waargenomen door docenten en handelingen die werden gerapporteerd door schoolleiders (zie paragraaf 6.4). In deze modellen zijn de volgende contextvariabelen opgenomen: schoolgrootte (leerlingenaantal), denominatie (rooms-katholiek), percentage culturele minderheden, stedelijkheid van de omgeving, waardering van het onderwijs door leerlingen en de concurrentie die wordt ervaren van andere scholen. De volgende vier typen handelingen zijn onderscheiden, op basis van het concurrerende waarden model (Quinn & Rohrbaugh, 1983): rationele doel handelingen,

interne proceshandelingen, 'human relations' handelingen en open systeem handelingen. Intermediaire variabelen voor de invloed van schoolleidersgedrag zijn de prestatiegerichtheid en ontwikkelingsgerichtheid van de school, de waardering van de school door leerlingen en het werkklimaat in de klas. De uitkomsten werden gemeten aan de hand van het percentage leerlingen dat doorstroomt van de derde klas naar het diploma zonder zittenblijven en het gemiddelde eindexamencijfer voor alle vakken.

Uit het concurrerende waarden model is af te leiden dat vooral rationele doel en interne proces handelingen van schoolleiders een effect zouden moeten hebben op de prestatiegerichtheid van de school. Dit klopt ook, alleen blijken open systeem handelingen in minstens zo grote mate bij te dragen aan de prestatiegerichtheid. Overigens geldt dit alleen voor handelingen die door docenten werden waargenomen (zie figuur 6.2). Handelingen die door de schoolleiders zelf werden opgegeven, lijken niet samen te hangen met prestatiegerichtheid (zie figuur 6.3). Rationele doel en open systeem handelingen hebben een sterke relatie met de ontwikkelingsgerichtheid van een school. Ook interne proces en 'human relations' handelingen die worden waargenomen door docenten lijken hiermee samen te hangen. Uitgaande van het concurrerende waarden model was te verwachten dat vooral 'human relations' en open systeem handelingen zouden bijdragen aan de ontwikkelingsgerichtheid, maar kennelijk zijn hierbij alle typen handelingen van belang.

Wanneer rekening wordt gehouden met contextvariabelen, lijkt er een heel klein positief effect te zijn van rationele doel, 'human relations' en open systeem handelingen, via de ontwikkelingsgerichtheid van de school en het werkklimaat in de klassen, op het doorstroompercentage (zie figuur 6.2 en tabel 6.1). Er is geen effect van de prestatiegerichtheid op de uitkomsten. Als alleen de katholieke scholen in beschouwing worden genomen, die een groot deel (40%) vormen van de onderzoeksgroep, is er in een beperkt model zonder context- en leiderschapsvariabelen wel een (directe) relatie tussen prestatiegerichtheid en het doorstroompercentage aantoonbaar (zie paragraaf 6.5). Daarnaast is er een negatieve relatie tussen ontwikkelingsgerichtheid en de eindexamencijfers. De relatie tussen prestatiegerichtheid en het doorstroompercentage voor katholieke scholen neemt af, wanneer contextvariabelen in het model worden opgenomen. Bij niet-katholieke scholen is er geen significante relatie tussen prestatiegerichtheid en het doorstroompercentage. Het ontbreken van een relatie tussen prestatiegerichtheid en de uitkomsten is opmerkelijk, omdat in het algemeen wordt aangenomen dat een prestatiegerichte atmosfeer de leerprestaties bevordert (Scheerens & Bosker, 1997). Er wordt echter ook wel verondersteld dat de samenstelling van de leerlingpopulatie hierbij een rol speelt. Met leerlingen die een hogere sociaal economische status hebben, zouden scholen een beter leerklimaat kunnen scheppen, waardoor het leerresultaat wordt bevorderd (zie bijv. Luyten et al., 2005).

Maslowski (2001) vond een negatief effect van een flexibele schoolcultuur ('human relations' en open systeem model) op het gemiddelde eindexamencijfer, ten opzichte van een intern gerichte cultuur (interne proces en 'human relations' model) en een positief effect van de rationele doel en interne proces oriëntatie op het gemiddelde doorstroompercentage. De uitkomsten van Maslowski komen overeen met de resultaten die hierboven zijn beschreven voor katholieke scholen, aan de hand van het model zonder contextvariabelen. In het onderzoek van Maslowski (2001) werd alleen het percentage culturele minderheden als controlevariabele gebruikt en katholieke scholen vormden ongeveer eenderde van de onderzoeksgroep. Het is dus mogelijk dat speciale kenmerken van katholieke scholen de resultaten hebben beïnvloed. Als er binnen katholieke scholen een bijzondere prestatiegerichte schoolcultuur wordt gecreëerd, waardoor deze scholen betere resultaten halen, zijn de

conclusies van Maslowski terecht. Aan de andere kant wijzen de resultaten van dit onderzoek erop dat de prestatiegerichtheid van katholieke scholen vooral verband houdt met de context en dus waarschijnlijk met de samenstelling van de leerlingenpopulatie. Diverse andere onderzoekers geven ook aan dat verschillen in leerresultaten tussen katholieke en niet-katholieke scholen voor een belangrijk deel zijn toe te schrijven aan verschillen in leerlingenkenmerken (zie bijv. Bryk & Raudenbusch, 1992; Opdenakker & Van Damme, 2006).

In de Verenigde Staten gebruikte Griffith (2003) het concurrerende waarden model als raamwerk voor het analyseren van schooleffectiviteit in het basisonderwijs. Hierbij werd voor elk organisatie-model afzonderlijk, door middel van structural equation modeling, getoetst in hoeverre het geschikt was om een vooruitgang in de leerprestaties te verklaren. De leerprestaties werden gemeten door test scores van dezelfde leerlingen in twee opeenvolgende jaren te vergelijken. Het effect van het rationele doel model op de leerprestaties werd direct gemeten, het effect van de andere modellen indirect. Voor het interne proces model werd een beoordeling van de veiligheid, ordelijkheid en discipline door leerlingen gebruikt als intermediaire variabele, voor het 'human relations' model de tevredenheid van de staf en voor het open systeem model de relatie tussen de staf en ouders. Het 'human relations' model bleek een goede modelpassing te vertonen, gevolgd door het open systeem model. Het rationele doel model en het interne proces model pasten minder goed. Het open systeem model verklaarde de meeste variantie (19%) in de leerprestaties, gevolgd door het interne proces model en het 'human relations' model (elk 11%). Het rationele doel model verklaarde slechts 3% van de variantie in de leerprestaties. Uit het onderzoek van Griffith (2003) kan worden opgemaakt dat een flexibele schoolorganisatie meer bijdraagt aan de leerprestaties dan een op beheersing gerichte organisatie. Deze resultaten zijn in lijn met in de uitkomsten van dit onderzoek, in die zin dat in dit onderzoek een flexibele, 'ontwikkelingsgerichte' organisatie een (klein) positief verband vertoont met de uitkomsten en een rationele, 'prestatiegerichte' organisatie niet.

De samenhang tussen door docenten waargenomen schoolleidersgedrag en de ontwikkelingsgerichtheid van scholen is heel groot. De waargenomen handelingen verklaren meer dan 80% van de variantie in ontwikkelingsgerichtheid. De ontwikkelingsgerichtheid van scholen zoals die wordt gemeten in dit onderzoek vertoont overeenkomsten met *organisational learning* zoals dat wordt beschreven door Mulford (2006, p. 51): "establishing a trusting and collaborative climate; followed by having a shared and monitored vision; and then taking risks and initiatives (that is, being innovative); within a context of ongoing, relevant professional development." Volgens het model van Mulford en Silins (2003) is *organisational learning* een belangrijke intermediaire variabele voor de invloed van schoolleiders op de leerprestaties. In het model voor schoolleiderschap in dit onderzoek is ontwikkelingsgerichtheid de intermediaire variabele voor een klein indirect effect van schoolleiderschap op het doorstroompercentage via het werkklimaat voor de leerlingen (zie tabel 6.1).

Het model voor de handelingen die door schoolleiders zelf werden gerapporteerd (figuur 6.3) bevat een direct effect van ontwikkelingsgerichtheid op prestatiegerichtheid. In dit model wordt de samenhang tussen prestatiegerichtheid en ontwikkelingsgerichtheid kennelijk niet verklaard door het verband van beide variabelen met schoolleidersgedrag, zoals in het model voor door docenten waargenomen handelingen (figuur 6.2). Heck, Larsen en Marcoulides (1990) vonden in hun model voor schoolleiderschap een enigszins vergelijkbaar direct effect van de schoolorganisatie (ondersteunend voor het werk van docenten) op het (prestatiegerichte) schoolklimaat (zie paragraaf 2.6).

Het meest opvallende in de modellen die zijn gevonden in dit onderzoek is het substantiële directe negatieve effect van rationele doel en (met name) open systeem handelingen op het gemiddelde examencijfer. Een mogelijke verklaring is dat op veel scholen waar schoolleiders intensief bezig zijn met doelen en externe betrekkingen, de eindexamencijfers relatief laag zijn. De lage eindexamencijfers zijn daarbij waarschijnlijk de oorzaak van de acties en niet het gevolg. Het veronderstelde effect van schoolleidersgedrag op het gemiddelde examencijfer komt voort vanuit het theoretische model dat handelingen van de schoolleiders voorafgaan aan de resultaten. Vanwege het cross-sectionele karakter van het onderzoek is het echter niet mogelijk om echte causale verbanden aan te tonen, tussen gebeurtenissen die elkaar opvolgen in de tijd. In dit onderzoek bleek echter dat een effect van schoolleidersgedrag op het examencijfer een beter passend model opleverde dan een omgekeerd effect van het examencijfer op schoolleidersgedrag. Het zou zo kunnen zijn dat hoge rationele doel en open systeem activiteit vaak gepaard gaan met lage eindexamencijfers, terwijl lage eindexamencijfers niet altijd betekenen dat de schoolleidersactiviteit verhoogd is, of dat lage rationele doel en open systeem activiteit vaak samen gaan met hoge examencijfers, terwijl hoge examencijfers niet altijd tot gevolg hebben dat de schoolleidersactiviteit laag is. Krüger en haar collega's vonden een negatief verband tussen *strategic leadership* en de betrokkenheid van studenten, waarbij 'strategic leadership' wordt gekenmerkt door een gerichtheid op doelen en plannen (Krüger, Witziers & Slegers, 2007). Deze negatieve samenhang is wellicht vergelijkbaar met het negatieve verband tussen het gemiddelde eindexamencijfer en de rationele doel en open systeem handelingen. Schoolleiders van scholen met lage leerprestaties zouden zich verplicht kunnen voelen om actie te ondernemen (zie ook Teddlie & Stringfield, 1993; Witziers, Bosker & Krüger, 2003). Het grote negatieve 'directe' verband dat in dit onderzoek werd gevonden zou kunnen wijzen op een trend in het onderwijs. Op (een deel van) de scholen die in moeilijke omstandigheden verkeren, is de schoolleiding bezig met het formuleren van doelen en het uitdragen van een visie en ook met het intensiveren van het contact met ouders en andere externe partijen, om de resultaten en het imago van de scholen te verbeteren. Omdat het hier gaat om een cross-sectioneel onderzoek geeft het causaal model een momentopname. Misschien verandert het effect van leiderschapsvariabelen wel in de tijd, afhankelijk van ontwikkelingen in de schoolorganisatie. Dergelijke effecten kunnen alleen worden blootgelegd door middel van longitudinaal onderzoek. In een stabiele situatie is er wellicht geen effect van leiderschapshandelingen te zien, alleen onder bijzondere omstandigheden. Ervan uitgaande dat leiderschapshandelingen vooraf gaan aan de resultaten, is er wellicht op een later moment, wanneer de resultaten zijn verbeterd, geen verhoogde leiderschapactiviteit meer zichtbaar. Enige ondersteuning voor deze veronderstelling kan worden gevonden in tabel 5.8 in hoofdstuk 5. Schoolleiders van de hoogst presterende scholen vertonen minder rationele doel, 'human relations' en open systeem handelingen dan schoolleiders van andere scholen.

In modellen voor de invloed van schoolleidersgedrag zijn vaker zowel positieve als negatieve verbanden tussen schoolleiderschap en uitkomsten te zien (zie bijv. Krüger, Witziers & Slegers, 2007; Leithwood et al., 2006a). Door de combinatie van positieve en negatieve verbanden zou het totale effect beperkt kunnen zijn en moeilijk aantoonbaar. Dit kan worden geïllustreerd aan de hand van het werk van De Maeyer en zijn collega's (De Maeyer et al., 2007). De Maeyer et al. tonen aan dat een 'direct en indirect' effectmodel met contextvariabelen de beste keuze is om het effect van schoolleiderschap op de leerprestaties aan te tonen. In hun model is er een significant positief indirect effect van leiderschap op de leerprestaties. Het negatieve directe effect van leiderschap is weliswaar niet significant, maar

wel ongeveer even groot als het indirecte effect.²¹ Het totale effect van schoolleiderschap zal in het model van De Maeyer et al. dus toch nog erg klein zijn.

In dit onderzoek werd, uitgaande van een indirect effect model, een klein positief indirect effect gevonden van schoolleidersgedrag op het doorstroompercentage en een groter negatief direct effect op de leerprestaties. De negatieve relatie zou een actueel (tijdelijk) effect kunnen zijn, dat toevallig in dit cross-sectionele onderzoek is blootgelegd. De laatste jaren is er veel aandacht voor *accountability* (het afleggen van rekenschap) in het onderwijs. Als de resultaten van de school te wensen overlaten, voelen schoolleiders zich wellicht meer dan vroeger geroepen om bepaalde activiteiten te ondernemen. Dit zou kunnen verklaren waarom er vroeger geen effect kon worden aangetoond van schoolleiderschap in het voortgezet onderwijs in Nederland (zie bijv. Witziers, Bosker en Krüger, 2003) en nu wel. In tegenstelling tot wat vaak wordt verondersteld, maakt het gebruik van indirecte of directe effectmodellen waarschijnlijk weinig verschil voor het aantonen van een effect van schoolleiderschap. Een indirect effectmodel zou een betere modelpassing kunnen vertonen, maar de grootte van een effect in een causaal model verandert niet door het toevoegen of weglaten van intermediaire variabelen (Saris & Stronkhorst, 1984, p. 35; Verschuren, 1991, p. 74). Intermediaire variabelen zijn uiteraard wel van belang om in kaart te kunnen brengen hoe de invloed van de schoolleider precies verloopt. Het indirecte effect van schoolleiderschap op de leerprestaties wordt echter kleiner naarmate er meer intermediaire variabelen zijn: het totale indirecte effect is het product van alle tusseneffecten. Contextvariabelen die van invloed zijn op de handelingen van schoolleiders, maar niet op andere variabelen in het model (zoals de visie van schoolleiders, zie Krüger, Witziers & Slegers, 2007) kunnen worden weggelaten, zonder dat dit het effect van schoolleidersgedrag op de uitkomsten beïnvloedt. Contextvariabelen die van invloed zijn op twee of meer variabelen in het model kunnen echter niet worden weggelaten (Saris & Stronkhorst, 1984, p. 35; Verschuren, 1991, p. 74).

Ten aanzien van de derde onderzoeksvraag, die in dit hoofdstuk centraal staat, kan worden geconcludeerd dat het erg moeilijk is om effecten van schoolleiderschap op de leerprestaties in kaart te brengen. Het totale effect is mogelijk opgebouwd uit diverse positieve en negatieve, directe en indirecte verbanden. De effecten van schoolleiderschap zijn afhankelijk van de context waarin de school zich bevindt. Niettemin kan uit de modellen die hier zijn gepresenteerd, worden afgeleid dat er relaties zijn tussen schoolleidersgedrag en de uitkomsten. Naast een negatieve samenhang tussen specifieke handelingen van schoolleiders en de leerprestaties is er ook een klein positief indirect verband met het doorstroompercentage gevonden. Het is niet uit te sluiten dat er hierbij sprake is van reciproque relaties. De schoolleider kan de schoolorganisatie beïnvloeden, maar de situatie op school zal ook weer een effect hebben op het gedrag van de schoolleider. Reciproque modellen zijn echter alleen goed te onderzoeken aan de hand van longitudinale gegevens.

De effecten van de handelingen van schoolleiders en andere schoolkenmerken op het doorstroompercentage zijn groter dan de effecten op de leerprestaties. Het percentage leerlingen dat onvertraagd doorstroomt van de derde klas naar het diploma, is dus beter te beïnvloeden door de school(leiding) dan de eindexamencijfers, die afhankelijk zijn van de landelijke toetsing. In het doorstroompercentage is zichtbaar wat er ‘onderweg’ naar het eindexamen gebeurt: het geeft aan in welke mate leerlingen doubleren of uitvallen. Een hoger doorstroompercentage betekent dat meer leerlingen zonder zittenblijven het examen halen,

²¹ Het indirecte effect van leiderschap op de leesprestaties is $.59 \cdot .43 = .25$ en het directe effect $-.27$ (zie De Maeyer et al., 2007, p. 138). Met dank aan prof. dr. C.A.W. Glas.

maar niet noodzakelijk hogere eindexamencijfers. Scholen die meer leerlingen laten doubleren (of afstromen) voor het eindexamenjaar zouden uiteindelijk hogere gemiddelde eindexamencijfers kunnen halen. Voor een goede beoordeling van de prestaties van een school zijn beide uitkomstmaten dus van belang.

Hoofdstuk 7

Samenvatting en conclusies

7.1 Inleiding

Een sterke schoolleider bevordert de effectiviteit van de school. Deze veronderstelling, die eigenlijk meer is gebaseerd op consensus dan op empirische resultaten, vormde het uitgangspunt voor dit proefschrift. Sinds de jaren zeventig wordt in schooleffectiviteitsonderzoek gezocht naar veranderbare schoolkenmerken die de leerprestaties positief beïnvloeden (zie bijv. Scheerens, 1989; Scheerens & Bosker, 1997). Hierbij kwam al in een vroeg stadium naar voren, dat een sterke schoolleiding een van de belangrijkste kenmerken zou zijn van een effectieve school (Edmonds, 1979). Het bleek echter moeilijk om daadwerkelijk een effect van schoolleiderschap op de leerprestaties aan te tonen. Uit een analyse van onderzoek tussen 1980 en 1995, concludeerden Hallinger en Heck (1998), dat schoolleiders een klein maar significant *indirect effect* uitoefenen op de leerresultaten, via verschillende ‘intermediaire’ variabelen, zoals de schoolcultuur en de motivatie van docenten. Onderzoek waarin gebruik werd gemaakt van *directe effectmodellen* (zonder intermediaire variabelen) leverde volgens Hallinger en Heck weinig consistente resultaten op.

Aan de hand van een kwantitatieve meta-analyse van onderzoek naar directe effecten van schoolleiderschap tussen 1986 en 1996, stelden Witziers, Bosker en Krüger (2003) vast dat er geen bewijs was voor een direct effect van schoolleiderschap in het voortgezet onderwijs, terwijl er voor het primaire onderwijs wel een klein direct effect aangetoond kon worden. Witziers en zijn collega's (2003) concludeerden dat schoolleiderschap niet meer dan één procent van de variantie in de leerprestaties zou verklaren. Leithwood en Levin (2005) merken echter op dat Witziers et al. bij hun analyse uitgingen van een ‘beperkt’ model van schoolleiderschap (namelijk de *Principal Instructional Management Rating Scale* van Hallinger, 1989, 1994). Zij geven aan dat uit grootschalige kwantitatieve onderzoeken zou blijken dat drie tot vijf procent van de *verschillen tussen scholen* verklaard kan worden door effecten van schoolleiderschap. In een recenter document gaan Leithwood et al. (2006b) ervan uit dat schoolleiderschap vijf tot zeven procent zou verklaren van de verschillen in leerprestaties “across schools (not to be confused with the typically very large differences among pupils within schools)” (p. 4). Schoolleiderschap zou daarmee volgens Leithwood et al. een van de belangrijkste effectiviteitsbevorderende factoren op schoolniveau zijn. Ervan uitgaande dat ongeveer 15% van de variantie in leerprestaties is gesitueerd op schoolniveau (Luyten, 1994, p. 55), betekent vijf tot zeven procent van de *variantie op schoolniveau* echter dat niet meer dan één procent ($0.07 \cdot 0.15$) van de *totale variantie* in de leerprestaties wordt verklaard door schoolleiderschap. De uitkomsten van Leithwood et al. (2006b) en Witziers et al. (2003) lopen dus eigenlijk minder uiteen dan op het eerste gezicht het geval lijkt te zijn. Dat neemt echter niet weg dat inderdaad een aanzienlijk deel (een kwart) van de *main school effect* (goed voor 25 procent van de variantie op schoolniveau, volgens Luyten, 1994, p.55) zou kunnen worden toegeschreven aan schoolleiderschap.

In de afgelopen jaren is het takenpakket van schoolleiders in het voortgezet onderwijs in Nederland verzaamd, vanwege de schaal- en autonomievergroting van scholen en de toegenomen verantwoordingsplicht. Het functioneren van de schoolleiding is daarom een belangrijk thema geworden in het onderzoek naar schooleffectiviteit en schoolverbetering (Slegers, 1999). Om iets te kunnen zeggen over de mate waarin schoolleiders ‘het verschil kunnen maken’, bijvoorbeeld voor de invoering van onderwijsvernieuwingen (al dan niet geïnitieerd vanuit de overheid), is het van belang om te onderzoeken hoe groot de invloed van schoolleiders op de leerprestaties precies is en langs welke wegen deze verloopt. De bovenstaande ontwikkelingen hebben geleid tot de volgende hoofdvraag voor dit proefschrift:

Wat is de directe en indirecte invloed van schoolleiders op de leerprestaties in het voortgezet onderwijs in Nederland?

Zoals uit het voorgaande kan worden opgemaakt, is het onderzoek naar de effecten van schoolleiderschap niet eenvoudig. Het wordt gehinderd door een aantal problemen: er bestaat bijvoorbeeld geen overeenstemming over wat schoolleiderschap precies inhoudt en het is ook nog onduidelijk in hoeverre andere variabelen een rol spelen bij de invloed van schoolleiders op de leerprestaties. Bovendien is het effect van schoolleiderschap relatief klein, zodat geavanceerde analysetechnieken toegepast moeten worden. Leithwood en Levin (2005) stelden op basis van literatuuronderzoek een uitgebreid overzicht op van de conceptuele en methodologische ‘uitdagingen’. Zij bevelen aan om een bredere set van leiderschapsactiviteiten te onderzoeken dan de specifieke modellen van schoolleiderschap die eerder werden onderzocht. Voorbeelden van deze specifieke modellen zijn: ‘onderwijskundig leiderschap’ dat direct is gericht op de onderwijsactiviteiten van docenten, ‘transformationeel leiderschap’ dat vooral tot doel heeft om de betrokkenheid en capaciteiten van docenten te verhogen, of ‘strategisch leiderschap’, waarbij de nadruk ligt op coördineren, plannen, monitoren en het verdelen van middelen. Bij het meten van leiderschap zou volgens Leithwood en Levin het accent veel meer moeten liggen op die leiderschapsactiviteiten die daadwerkelijk een effect zouden kunnen hebben op de leerlingenresultaten. De auteurs beschrijven een algemeen model om de effecten van leiderschap te onderzoeken, waarin contextvariabelen zijn opgenomen als ‘modererende’ variabelen en schoolcondities, klassencondities en docentfactoren als ‘mediërende’ (intermediaire) variabelen. Zij geven aan dat inmiddels een groot aantal potentiële intermediaire variabelen bekend is, maar dat niet voldoende duidelijk is hoe schoolleiders deze variabelen beïnvloeden. Leithwood en Levin (2005) benadrukken dat er meer aandacht geschonken moet worden aan de modererende contextvariabelen, die het effect van leiderschap zouden kunnen verkleinen, neutraliseren of zelfs vergroten. Het gebrek aan aandacht voor modererende variabelen zou een belangrijke oorzaak kunnen zijn van de tegenstrijdige uitkomsten in schoolleiderschapsonderzoek.

7.2 Onderzoeksmodel en variabelen

Vanwege de knelpunten in het onderzoek naar schoolleiderschap (zie ook van Vilsteren, 1999), werd bij de start van dit promotie-onderzoek (in 2003) een zodanige onderzoekszet gekozen, dat aan een aantal van de aanbevelingen van Leithwood en Levin (2005) tegemoet gekomen kon worden. Voor het meten van schoolleiderschap is uitgegaan van een ‘generiek’ leiderschapsmodel, gebaseerd op het concurrerende waarden model van Quinn en Rohrbaugh (1983). Verder werd een (causaal) onderzoeksmodel geconstrueerd aan de hand van het algemene beschrijvingsmodel van schoolleiderschap van Bossert et al. (Bossert, Dwyer, Rowan & Lee, 1982; Dwyer et al., 1985; Dwyer, Barnett & Lee, 1987). In dit algemene beschrijvingsmodel wordt zowel een direct effect verondersteld van schoolleidersgedrag op

de leerprestaties, als indirecte effecten, via de intermediaire variabelen ‘pedagogische onderwijskundig klimaat’ en ‘onderwijsorganisatie’. Daarnaast wordt ervan uitgegaan dat contextfactoren, zoals de omgeving van de school, visie en ervaring van de schoolleider en wettelijke kaders en regelgeving een belangrijke rol spelen.

Om het onderzoeksmodel nader te kunnen specificeren werd een literatuuronderzoek uitgevoerd aan de hand van de eerste onderzoeksvraag:

Welke contextvariabelen, schoolleidersvariabelen, intermediaire variabelen en uitkomstvariabelen zijn relevant voor het in kaart brengen van de effectiviteit van de schoolleider en welke relaties zijn er tussen deze variabelen te onderscheiden?

Het onderzoek naar de effecten van schoolleiderschap wordt gekenmerkt door uiteenlopende conceptuele en methodologische benaderingen (zie bijv. Leithwood & Levin, 2005; Heck & Hallinger, 2000; Krüger & Witziers, 2003; Scheerens & Witziers, 2005). Dit heeft als voordeel dat het complexe begrip van vele kanten wordt belicht, maar als nadeel dat er nog weinig eenduidige conclusies zijn te trekken over de effectiviteit van schoolleiderschap. In de loop der jaren is het begrip schoolleiderschap verbreed van onderwijskundig leiderschap, via (onder meer) transformationeel en gedistribueerd leiderschap, naar geïntegreerd leiderschap. Marks en Printy (2003) beargumenteren dat transformationeel leiderschap nodig is voor het bevorderen van veranderingen en gedeeld onderwijskundig leiderschap om de resultaten van leerlingen te verbeteren. Leithwood en zijn collega's veronderstellen dat er een repertoire bestaat van leiderschapsactiviteiten waar alle succesvolle (school-)leiders gebruik van maken (Leithwood et al., 2006a; NCLS, 2007). Dit repertoire van leiderschapsactiviteiten vertoont grote overeenkomsten met het generieke managementmodel van Quinn c.s. (1988, 1998; Quinn et al., 2003).

Het concurrerende waarden model van Quinn & Rohrbaugh (1983) is een raamwerk waarin verschillende theorieën van organisatie-effectiviteit zijn geïntegreerd. Het raamwerk omvat vier organisatiemodellen die uitgaan van verschillende effectiviteitscriteria (Quinn & Rohrbaugh, 1983; Quinn, 1998; zie ook Maslowski, 2001; Vermeulen, 1997; van Muijen, 1994; van Wieringen, Ax, Karstanje & Voogt, 2000):

- het rationele doel model: door middel van planning en het stellen van doelen wordt gestreefd naar efficiëntie en productiviteit;
- het interne proces model: hierbij zijn informatie-management en communicatie de middelen om beheersing en stabiliteit van de organisatie te bereiken;
- het ‘human relations’ model: door het bevorderen van inzet, samenhang en moreel wordt de ontwikkeling van menselijk potentieel nagestreefd;
- het open systeem model: alertheid en aanpassingsvermogen moeten leiden tot het verwerven van middelen en groei.

De vier organisatiemodellen vertonen zowel tegenstellingen als overeenkomsten. De relaties tussen de modellen kunnen in twee dimensies worden weergegeven: van nadruk op de interne organisatie tot nadruk op de externe organisatie en van flexibiliteit tot beheersing. Het rationele doel model en het open systeem model zijn extern gericht, terwijl bij het interne proces model en het ‘human relations’ model de focus ligt op de gang van zaken binnen de organisatie. Het ‘human relations’ en het open systeem model benadrukken de flexibiliteit en het aanpassingsvermogen van de organisatie, terwijl bij rationele doel en het interne proces model zijn gericht op beheersing en stabiliteit. Aan de vier organisatiemodellen zijn verschillende leiderschapsrollen en -competenties gekoppeld (Quinn, 1998; Quinn, Faerman.

Thompson & McGrath, 2003). Het concurrerende waarden model biedt dus enerzijds een integratie van verschillende managementmodellen en anderzijds de mogelijkheid om relaties te leggen tussen leiderschapsgedrag en organisatiekenmerken.

Door de verschillende benaderingen in het onderzoek naar de effecten van schoolleiderschap is een grote verscheidenheid aan potentiële intermediaire variabelen aan het licht gekomen (Leithwood & Levin, 2005; Scheerens & Witziers, 2005; Witziers, Bosker & Krüger, 2003). Dit zijn variabelen op het niveau van de school, de klas, de docenten als collectief, of de individuele docent (Leithwood & Levin, 2005). Hallinger en Heck (1998) onderscheiden vier 'domeinen' waardoor de schoolleider invloed zou kunnen uitoefenen op de leerprestaties: 'visie en doelstellingen', 'organisatiestructuur', 'menselijk kapitaal' en 'organisatiecultuur'. De intermediaire variabelen zijn in het schooleffectiviteitsmodel van Scheerens (1990) terug te vinden als 'proces' kenmerken, die een effect uitoefenen op de leerprestaties. Door Scheerens en Bosker (1997) is een overzicht samengesteld van proceskenmerken die de schooleffectiviteit bevorderen.

Een belangrijke leidraad voor de uitwerking van het onderzoeksmodel was het LOLSO (*Leadership for Organisational Learning and Student Outcomes*)-model. Dit model is tot stand gekomen na een grootschalig onderzoek in het voortgezet onderwijs in Australië (zie bijv. Mulford en Silins, 2003; Mulford, Silins & Leithwood, 2004). Het LOLSO-model geeft aan dat transformationeel en 'gedistribueerd' (over verschillende functionarissen verdeeld) schoolleiderschap een effect heeft op de betrokkenheid van leerlingen, via een 'lerende' schoolorganisatie en het werk van docenten. De grotere betrokkenheid van leerlingen bij de school zou, volgens het LOLSO-model, leiden tot verminderde schooluitval ('*retention*') en betere leerprestaties. In het LOLSO-model staat het concept van de 'lerende schoolorganisatie' centraal. Hierbij gaat om een klimaat van onderling vertrouwen en samenwerking, een gedeelde missie die wordt 'gemonitord' en het nemen van initiatieven en risico's, binnen de context van ondersteunende, continue en relevante professionele ontwikkeling (Mulford, Silins & Leithwood, 2004, p.5).

In dit onderzoek is ervoor gekozen om het concurrerende waarden model (zie boven) te gebruiken voor de operationalisering van leiderschapsvariabelen en intermediaire variabelen. Een belangrijk deel van de kenmerken van de lerende organisatie is echter wel terug te vinden in het bredere kader van Quinn en Rohrbaugh (1983), met name in het rationele doel, 'human relations' en open systeem model. Voor de invulling van de leiderschapsvariabelen werd uitgegaan van een instrument voor het meten van schoolleidersgedrag van Krüger (1994), dat was gebaseerd op de *Principal Instructional Management Rating Scale* van Hallinger (1989, 1994). Gebruik makend van het werk van Houtveen en haar collega's (Houtveen, Voogt, van der Vegt en van de Grift, 1995) en van Maslowski (2001) werden de intermediaire variabelen 'pedagogisch onderwijskundig klimaat' en 'onderwijsorganisatie' uit het algemene beschrijvingsmodel van schoolleiderschap van Bossert et al. (zie boven) vertaald in de variabelen 'schoolcultuur' en 'schoolorganisatiepraktijk'. Hierbij werd de schoolcultuur ingevuld als de 'gemeenschappelijke vooronderstellingen, waarden en normen op school' en de schoolorganisatiepraktijk 'als processen, praktijken en organisatorische voorwaarden die van invloed zijn op de manier waarop de organisatie functioneert' (zie Houtveen e.a, 1995, p.33). Bij de schoolcultuur gaat het dan vooral om ideeën en verwachtingen en bij de schoolorganisatiepraktijk om concrete praktijken en gedragingen. De verwachting was dat schoolleiders meer 'vat' zouden hebben op de schoolorganisatiepraktijk dan op de schoolcultuur en wellicht via de organisatie de cultuur zouden kunnen beïnvloeden (zie bijv. Heck, Larsen & Marcoulides, 1990; Krüger, Witziers & Slegers, 2007). Bij het

operationaliseren van de intermediaire variabelen schoolcultuur en schoolorganisatiepraktijk werd rekening gehouden met de schooleffectiviteitsbevorderende factoren die werden geïnventariseerd door Scheerens en Bosker (1997).

Aan de hand van de resultaten van het literatuuronderzoek werd een concept-onderzoeksmodel opgesteld, waarmee de eerste onderzoeksvraag is beantwoord. In dit model werd verondersteld dat de schoolleider een indirect effect zou hebben op de leerprestaties door invloed uit te oefenen op de schoolcultuur en de schoolorganisatiepraktijk. Uitgaande van het LOLSO-model werd aangenomen dat dit effect verder zou verlopen via twee andere intermediaire variabelen, namelijk ‘werk van docenten’ en ‘betrokkenheid van leerlingen’. Deze schoolbelevingsvariabelen werden geoperationaliseerd aan de hand van een instrument van Leithwood c.s. en enkele andere leerlingenvragenlijsten (Engels, Aelterman, Scheppens & Van Petegem, 2003; Houtveen, Vermeulen & van de Grift, 1993; Krüger, 1994; Leithwood, Aitken & Jantzi, 2001; OECD, 2003). Naar aanleiding van zowel het LOLSO-model van Mulford en zijn collega’s (zie boven), als het werk van Maslowski (2001) werd het gemiddelde doorstroompercentage in de bovenbouw als extra (intermediaire) uitkomstvariabele toegevoegd aan het concept-onderzoeksmodel. Maslowski vond geen significant verband tussen dimensies van de schoolcultuur en de leerprestaties, maar hij constateerde wel een relatie tussen zowel een rationele doel, als een interne proces oriëntatie en het rendement, uitgedrukt in het gemiddelde doorstroompercentage in de bovenbouw.

In het concept-onderzoeksmodel zijn alleen contextvariabelen opgenomen die van invloed zijn op de leerprestaties en mogelijk op andere variabelen in het model. Het gaat hierbij om ‘covariabelen’ zoals de samenstelling van de leerlingpopulatie, schoolgrootte, schoolsoort en stedelijkheid van de omgeving. De mate waarin contextvariabelen het gedrag van de schoolleider kunnen verklaren is in een complementair onderzoek bestudeerd door Schmidt (2009).

7.3 Dataverzameling en validering van de onderzoeksinstrumenten

Om gegevens te kunnen verzamelen voor de toetsing van het onderzoeksmodel, werden alle scholen voor voortgezet onderwijs met een havo-bovenbouw benaderd. Dit leverde een respons op van 21% van de havo-vestigingen. Een dergelijk responspercentage is normaal voor dit soort onderzoek. Schoolleiders geven aan dat zij vele verzoeken ontvangen om mee te werken aan onderzoek en dat zij dus selectief moeten zijn. Verder laat de situatie op school het niet altijd toe om deel te nemen, bijvoorbeeld als de school bezig is met ingrijpende onderwijsvernieuwingen of andere speciale projecten. In totaal werkten 103 schoolleiders, 998 docenten van de havo bovenbouw en 4336 leerlingen van havo-5 klassen mee aan het onderzoek. De groep van deelnemende scholen is representatief voor de totale populatie wat betreft de verdeling over de provincies en de schoolgrootte. Het percentage rooms-katholieke scholen onder de deelnemende scholen is groter dan in de totale populatie en het percentage overige bijzonder scholen kleiner. Het onderzoek is gericht op de havo, omdat dit een homogene opleiding is, met relatief geringe verschillen in aanleg tussen leerlingen. De schoolleider die centraal staat in dit onderzoek, is de eindverantwoordelijke schoolleider van de havo-locatie. Het streven was om het onderzoek uit te voeren onder schoolleiders met een vergelijkbaar takenpakket. De taken en verantwoordelijkheden van locatieleiders blijken echter toch nog behoorlijk te verschillen. Bovendien hebben niet alle locaties een locatiedirecteur, in sommige scholen is de schoolleiding bijvoorbeeld verdeeld over sectordirecteuren of afdelingsdirecteuren.

Zowel schoolleiders als docenten van de deelnemende scholen vulden een vragenlijst in over schoolleidersgedrag (42 items). Hiermee werd nagegaan in hoeverre schoolleiders handelingen uitvoeren die kenmerkend zijn voor een respectievelijk een rationele doel, interne proces, ‘human relations’, of open systeem benadering, volgens het concurrerende waarden model. Met behulp van vragenlijsten over de schoolcultuur en de schoolorganisatiepraktijk (elk 40 items) werd bij docenten geïnventariseerd in welke mate de schoolcultuur kan worden gekarakteriseerd door een rationele doel, interne proces, ‘human relations’ of open systeem oriëntatie en in hoeverre deze vier oriëntaties zijn terug te vinden in de schoolorganisatiepraktijk. Tenslotte vulden leerlingen een vragenlijst in over hun schoolbeleving (42 items). Van 97 van de 103 scholen die meewerkten aan het onderzoek werd een volledige dataset verkregen, met gegevens van de locatieleider, gemiddeld 10 docenten en gemiddeld 44 leerlingen per school.

Aan de hand van de verzamelde data werd eerst de kwaliteit van de onderzoeksinstrumenten vastgesteld. Hierbij werd aandacht besteed aan de constructvaliditeit en de betrouwbaarheid (of de interne consistentie), van de instrumenten. De constructvaliditeit is onder te verdelen in convergerende en discriminante validiteit. Voor een goede convergerende validiteit moeten correlaties tussen meetwaarden voor dezelfde concepten, gemeten met verschillende methodes, voldoende hoog zijn. De discriminante validiteit is in orde, als correlaties tussen meetwaarden voor verschillende concepten, gemeten met dezelfde methode, laag zijn. De constructvaliditeit en de betrouwbaarheid werden onderzocht met behulp van confirmatieve factoranalyse. Daartoe werd voor elk onderzoeksinstrument een meetmodel geconstrueerd, waarin de indicatoren (items) werden gekoppeld aan latente variabelen (schalen). Voor de toetsing van meetmodellen aan docenten- en leerlingendata werd meerniveau confirmatieve factoranalyse het meest geschikt geacht, vanwege de hiërarchische structuur van de data, met docenten en leerlingen ‘geclusterd’ in scholen (Hox, 2002; Snijders & Bosker, 1999). Met meerniveau confirmatieve factoranalyse kunnen (verschillende) meetmodellen tegelijkertijd worden getoetst op verschillende analyseniveaus (bijvoorbeeld schoolniveau en docentenniveau). Hierbij werd ervan uitgegaan, dat factorscores (latente variabele scores) op schoolniveau een zo zuiver mogelijke schatting van de schoolkenmerken zouden opleveren (zie De Maeyer & Rymenans, 2004, p. 260; De Maeyer, Rymenans, Van Petegem, van den Bergh & Rijlaarsdam, 2007).

Bij de uitvoering van meerniveau confirmatie factoranalyses met behulp van LISREL werd gebruik gemaakt van twee schattingsmethoden: *Muthén’s approximate Maximum Likelihood* (MUML, zie Muthén, 1990) en de meer geavanceerde *Full Information Maximum Likelihood* (FIML, zie bijv. Hox & Maas, 2001) methode. De eerste methode bleek robuuster voor de omvangrijke meetmodellen die werden getest, de tweede methode leverde minder accurate resultaten op. De beperkte steekproefgrootte op *between groups* niveau (het aantal scholen) en de relatief lage intraklassecorrelatie speelden hierbij wellicht een rol (zie 4.3.6). Omdat de LISREL software (versie 8.8) geen factorscores bleek te kunnen berekenen voor meerniveau modellen, werden uiteindelijk alleen voor docentendata (schoolleidersgedrag, schoolcultuur en schoolorganisatiepraktijk) meerniveau confirmatieve factoranalyses uitgevoerd, terwijl voor leerlingendata éénniveau analyse werd toegepast. Een voordeel van de meerniveau analyse was dat de modelpassing op schoolniveau kon worden bepaald. Dit was voor de leerlingvariabelen minder relevant.

Met behulp van de uitkomsten van de éénniveau en meerniveau confirmatieve factoranalyses zijn de meetmodellen geoptimaliseerd. Slecht ladende indicatoren werden verwijderd of verplaatst. Daarbij werd ook gebruik gemaakt van exploratieve factoranalyse. Voor

schoolleidersgedrag, de schoolcultuur en de schoolorganisatiepraktijk zijn de vier factoren, die de vier organisatiemodellen in het concurrerende waarden model vertegenwoordigen, in gewijzigde vorm gehandhaafd. Voor de schoolbeleving van leerlingen werden drie factoren gevonden: 'waardering van de school', 'waardering van onderwijs' en 'werkklimaat'. De aangepaste meetmodellen vertonen een betere modelpassing dan de oorspronkelijke modellen. Door het verwijderen van items met lage factorladingen, is de interne consistentie (betrouwbaarheid) verhoogd. De interne consistentie kan worden gezien als een vorm van convergerende validiteit. Verder verminderde de correlatie tussen de schalen, zodat de discriminante validiteit werd bevorderd. Voor alle schalen is de betrouwbaarheid zoals bepaald door middel van Cronbach's alpha voldoende (0.7 of hoger).

Op basis van de verbeterende meetmodellen werden factorscores gegenereerd en geaggregeerd naar schoolniveau. De betrouwbaarheid van de geaggregeerde scores (zie tabel 4.16) is voor de meeste schalen acceptabel (0.7-0.8), maar voor enkele schalen (interne proces en human relations schoolcultuur, waardering van onderwijs door leerlingen) aan de lage kant (0.4-0.6). De lage betrouwbaarheid van de naar schoolniveau geaggregeerde scores voor 'waardering van onderwijs' ligt voor de hand, omdat het hier meer gaat om een persoonlijke dan een schoolgebonden factor.

Nadat kwaliteit van de bijgestelde onderzoeksinstrumenten in principe was vastgesteld, werd de validiteit nader onderzocht aan de hand van de tweede onderzoeksvraag:

Is het concurrerende waarden model een geschikt kader om zowel het schoolleidersgedrag als de schoolcultuur en de schoolorganisatiepraktijk te karakteriseren en relaties tussen deze variabelen bloot te leggen?

Door sommige onderzoekers wordt verondersteld dat factoren die tegengestelde organisatiemodellen vertegenwoordigen in het concurrerende waarden model (rationele doel vs. 'human relations' en interne proces vs. open systeem), minder met elkaar zouden moeten correleren, dan factoren die gerelateerde organisatiemodellen representeren (Kalliath, Bluedorn & Gillespie, 1999; Lawrence, Quinn & Lenk, 2003). In dit onderzoek kon deze hypothese niet worden bevestigd. De correlaties tussen de schalen binnen de meetmodellen voor schoolleidersgedrag, schoolcultuur en schoolorganisatiepraktijk, vertonen niet het verwachte patroon. Andere onderzoekers menen echter dat het belang dat gehecht wordt aan de waarden die de vier organisatiemodellen vertegenwoordigen, kan afhangen van de specifieke kijk op de organisatie van de respondentengroep (Lamond 2003; van Muijen, Koopman & De Witte, 1998, p.109), de ontwikkelingsfase van de organisatie (Quinn, 1998), of actuele maatschappelijke ontwikkelingen (Kalliath, Bluedorn & Gillespie, 1999). Een afwijkend patroon van correlaties tussen de factoren dan op basis van overeenkomsten en verschillen tussen de vier kwadranten in het concurrerende waarden model zou worden verwacht, wordt dus wel vaker gevonden. De correlaties tussen de factoren geven daarom geen uitsluitsel over de validiteit van de onderzoeksinstrumenten.

De verbanden tussen schoolleidersgedrag, de schoolcultuur en de schoolorganisatiepraktijk werden nader onderzocht in een *multitrait-multimethod* matrix (Campbell & Fiske, 1959; zie ook Trochim, 2006). Tussen de scores van schoolleiders en docenten voor de vier typen handelingen blijken significante, maar matige correlaties te bestaan. De convergerende validiteit is dus acceptabel, maar er kan ook worden geconcludeerd dat schoolleiders en docenten (deels) verschillende percepties hebben van de uitgevoerde handelingen. Handelingen zoals die door docenten werden waargenomen, vertonen veel sterkere relaties

met respectievelijk schoolcultuur- en schoolorganisatievariabelen dan handelingen zoals die door schoolleiders zelf werden opgegeven. Tussen de schoolcultuur- en de schoolorganisatievariabelen onderling blijken ook aanzienlijke verbanden te bestaan. Dit zou kunnen wijzen op conceptuele overlap, hoewel volgens de literatuur pas bij een correlatie van .85 of hoger moet worden uitgegaan van onvoldoende discriminante validiteit (zie Garson, z.j., d, Kenny, 1998).

Met betrekking tot de tweede onderzoeksvraag kan worden geconcludeerd dat op basis van het concurrerende waarden model valide en betrouwbare onderzoeksinstrumenten zijn ontwikkeld voor het meten van schoolleidersgedrag, de schoolcultuur en de schoolorganisatiepraktijk. Tussen de schoolleiderschaps- schoolcultuur en schoolorganisatievariabelen bestaan de nodige relaties, het is daarbij echter de vraag of schoolcultuur en schoolorganisatiepraktijk voldoende van elkaar kunnen worden onderscheiden. In dit onderzoek wordt de schoolorganisatiepraktijk opgevat als een 'manifestatie van de cultuur van de organisatie' (zie ook Houtveen et al., 1995, p33). Overlap tussen schoolcultuur- en schoolorganisatievariabelen ligt dus voor de hand. Een principale componenten analyse van de acht variabelen voor schoolcultuur en schoolorganisatiepraktijk leverde op, dat twee componenten kunnen worden onderscheiden, die samen 82% van de variantie verklaren.

Hoewel de discriminante validiteit in principe voldoende was om alle variabelen te kunnen handhaven, werd om om inhoudelijke en praktische redenen gekozen voor het combineren van cultuur- en organisatievariabelen. Voor de toetsing van een structureel model was het voordelig om het aantal variabelen te verminderen, vanwege het relatief beperkte aantal cases. Een kleiner aantal variabelen zou bovendien de interpretatie van de uitkomsten vereenvoudigen. Met behulp van exploratieve factoranalyse zijn twee nieuwe 'cultuur' variabelen geconstrueerd: 'prestatiegerichtheid' en 'ontwikkelingsgerichtheid'. De variabele prestatiegerichtheid wordt bepaald door de som van de scores voor rationele doel cultuur, interne proces cultuur en rationele doel praktijk. De variabele 'ontwikkelingsgerichtheid' wordt bepaald door de som van de scores voor human relations cultuur, open systeem cultuur, interne proces praktijk en human relations praktijk.

7.4 Effecten van schoolkenmerken op de opbrengsten

Schoolleidersvariabelen, cultuurvariabelen en schoolbelevingsvariabelen in dit onderzoek worden beschouwd als 'schoolkenmerken' die van invloed kunnen zijn op de opbrengsten van het onderwijs. Voorafgaand aan de toetsing van een causaal model, werden deze schoolkenmerken nader geëxploreerd.

Eerder werd al duidelijk dat schoolleiders en docenten een verschillende kijk hebben op schoolleidersgedrag (zie paragraaf 7.3). Uit een vergelijking van de gemiddelde scores voor alle scholen blijkt dat de scores van schoolleiders en docenten voor interne proces en open systeem handelingen elkaar niet veel ontlopen, maar voor rationele doel en human relations handelingen wel. Deze verschillen worden veroorzaakt door een relatief hoge gemiddelde score van schoolleiders op rationele doel handelingen en een relatief lage score van docenten op human relations handelingen. hetgeen dus wil zeggen dat schoolleiders het eigen gedrag als meer rationele doel bepaald zien dan docenten, terwijl docenten schoolleiders als minder gericht op human relations zien, dan schoolleiders van zichzelf vinden. Schoolleiders ervaren over het algemeen in hun werk kennelijk een nadruk op het ontwikkelen en realiseren van doelstellingen, terwijl docenten deze extra aandacht niet lijken waar te nemen. Docenten ervaren relatief weinig aandacht van schoolleiders voor de ondersteuning van hun werk,

terwijl schoolleiders dit niet zo zien. Over het handhaven van orde, regels en procedures door de schoolleiders en over het onderhouden van externe contacten en het verwerven van middelen bestaat minder verschil van inzicht. Hoewel de gegevens uit dit onderzoek beperkt zijn, zou vooral uit de relatief hoge score van schoolleiders voor rationele doel handelingen en de relatief lage score van docenten voor human relations handelingen de conclusie kunnen worden getrokken, dat het gedrag van locatieleiders meer kenmerken vertoont van strategisch leiderschap dan van onderwijskundig leiderschap. Strategisch leiderschap omvat taken zoals plannen, coördineren, monitoren, het verdelen van middelen en ook het bevorderen van innovatie, terwijl onderwijskundig leiderschap is gericht op de activiteiten van docenten die het leren direct beïnvloeden (zie bijv. Leithwood & Levin, 2005). De strategische insteek van de locatieleider ligt voor de hand vanwege de grootte van schoollocaties in het voortgezet onderwijs: gemiddeld ruim 1100 leerlingen en 90 docenten. De locatieleider is eindverantwoordelijk voor de locatie en zal dus richting moeten geven, faciliteren en controleren. De mate waarin de schoolleider onderwijskundig leiderschap uitoefent, zal onder meer afhangen van de taakverdeling in de schoolleiding. Door de omvang van de scholen ligt het misschien voor de hand om het onderwijskundig leiderschap te verdelen over verschillende personen.

De twee variabelen ‘prestatiegerichtheid’ en ‘ontwikkelingsgerichtheid’, vertonen overeenkomsten met cultuurtypen die werden gevonden door Houtveen en haar collega’s (1995) en Maslowski (2001). De variabele ‘prestatiegerichtheid’ geeft aan in hoeverre belang wordt gehecht aan hoge leerprestaties en een ordelijke werksfeer. De variabele ‘ontwikkelingsgerichtheid’ meet de gerichtheid op samenwerking, ontwikkeling en vernieuwing in de schoolorganisatie. Van de onderzochte scholen heeft 14% een prestatiegerichte cultuur, 20% een ontwikkelingsgerichte cultuur en 35% een gecombineerde prestatiegerichte en ontwikkelingsgerichte cultuur. De rest van de scholen is relatief weinig prestatie- en ontwikkelingsgericht. De gemiddelde score van scholen voor prestatiegerichtheid ligt wat hoger dan die voor ontwikkelingsgerichtheid. Rooms-katholieke en protestants-christelijke scholen zijn prestatiegerichter dan openbare en overige scholen. De prestatiegerichtheid en ontwikkelingsgerichtheid van scholen nemen toe met het leerlingenaantal tot een schoolgrootte van 1000 tot 1500 leerlingen, daarna neemt de prestatiegerichtheid wat af, terwijl de ontwikkelingsgerichtheid vrijwel gelijk blijft. Dit zou erop kunnen wijzen dat grotere scholen beter toegerust zijn dan kleinere om een ordelijke en prestatiegerichte werksfeer te realiseren, met voldoende aandacht voor professionele ontwikkeling en innovatie. Het verband tussen zowel denominatie als schoolgrootte met prestatiegerichtheid en ontwikkelingsgerichtheid kan echter ook het gevolg zijn van een combinatie van contextfactoren, waarbij de samenstelling van de leerlingenpopulatie een belangrijke rol speelt (zie bijv. Luyten et al., 2005; Opdenakker en Van Damme, 2001, 2006).

De schoolbelevingsvariabele ‘waardering van de school’ meet hoe leerlingen de sfeer op school en de omgang met medeleerlingen ervaren. De gemiddelde score is vrij hoog. Havo-5 leerlingen blijken zich dus behoorlijk goed thuis te voelen op hun school. Hierbij moet echter rekening worden gehouden met de mogelijkheid dat leerlingen die zich minder identificeren met de school, in het voorgaande traject al zijn uitgevallen of vertrokken naar een andere school. De variabele ‘werkklimaat’ geeft de mening weer van leerlingen over het werkklimaat in de klassen. Hierbij gaat het om de omgang met docenten, de sfeer tijdens de lessen en de organisatie en invulling van de lessen. De gemiddelde score van leerlingen voor het werkklimaat is neutraal: niet negatief, maar ook niet positief. Hieruit zou op het eerste gezicht kunnen worden geconcludeerd dat het werkklimaat in de klassen voor verbetering vatbaar is. Welke variabelen hierbij van invloed zouden kunnen zijn, komt later aan de orde, bij de

bespreking van het structurele model. De variabele ‘waardering van onderwijs’ is een maat voor de motivatie van leerlingen en de ondersteuning van thuis (‘family educational culture’). De gemiddelde score voor alle scholen is licht positief. Ook hier geldt echter dat minder gemotiveerde leerlingen wellicht al eerder in het traject zijn uitvallen.

Nadat de schoolkenmerken in kaart waren gebracht, werden de relaties tussen schoolkenmerken en uitkomsten verkend, aan de hand van de derde onderzoeksvraag:

Zijn er (directe) relaties tussen de kenmerken van scholen voor voortgezet onderwijs in Nederland en de opbrengsten?

De opbrengsten van scholen werden bepaald met behulp van twee uitkomstvariabelen: het gemiddelde eindexamencijfer voor alle vakken (‘leerprestaties’) en het gemiddelde doorstroompercentage van het derde leerjaar naar het diploma (‘rendement’). Hiervoor werden gegevens ontleend aan de opbrengstenkaarten voor het voortgezet onderwijs (gepubliceerd door de Inspectie van het Onderwijs). Vanwege de variatie van jaar tot jaar in de uitkomsten (vgl. Luyten, 1994; Maslowski, 2001) werden gemiddelde eindexamencijfers en doorstroompercentages berekend over een periode van drie jaar (2003-2005).

Om een indicatie te krijgen van de verbanden tussen schoolkenmerken en opbrengsten, werden (Pearson) correlaties berekend. Feitelijk is hier sprake van een ‘direct-effect-model’, zonder intermediaire of contextvariabelen. Als er tussen twee variabelen een causale relatie bestaat, dan moet er ook een correlatie zijn tussen de twee variabelen (Verschuren, 1991, p.46). Uit de correlatiecoëfficiënten (zie tabel 5.7) blijkt dat er geen significante verbanden zijn tussen schoolleidersgedrag en het doorstroompercentage. Er is ook geen samenhang tussen rationele doel, interne proces of human relations handelingen en het gemiddelde examencijfer, maar er is wel een significante negatieve correlatie tussen open systeem handelingen en het gemiddelde examencijfer. Negatieve relaties tussen leerprestaties en handelingen die zijn gericht op het ontwikkelen en verbeteren van de school worden wel vaker gevonden. Er wordt verondersteld dat schoolleiders bij lage leerprestaties meer geneigd zijn om actie te ondernemen om de school te verbeteren (Teddlie & Stringfield, 1993; Witziers, Bosker & Krüger, 2003).

Een substantiële directe relatie tussen de handelingen van schoolleiders en de leerprestaties werd in dit onderzoek overigens niet verwacht, omdat over het algemeen wordt aangenomen dat dergelijke relaties moeilijk aantoonbaar zijn, met name in het voortgezet onderwijs in Nederland (Hallinger & Heck, 1998; Witziers, Bosker & Krüger, 2003).

De relaties tussen schoolleidersgedrag en uitkomsten werden nader onderzocht door scholen met hoge, gemiddelde en lage eindexamencijfers te vergelijken. Op tien scholen met de hoogste gemiddelde eindexamencijfers bleken schoolleiders significant minder rationele doel en open systeem handelingen te vertonen dan op de andere scholen. Schoolleiders van de tien scholen met de hoogste cijfers leggen ook minder nadruk op human relations handelingen dan schoolleiders op ‘gemiddelde’ scholen, maar het verschil met de tien laagst scorende scholen is niet significant. Schoolleiders van scholen met hoge eindexamencijfers zouden dus (gemiddeld) minder nadruk leggen op doelstellingen, ondersteuning van het werk van docenten, externe contacten en het verwerven van middelen, dan schoolleiders van gemiddeld of laag scorende scholen. Interne proces handelingen, die zijn gericht op stabiliteit en continuïteit, zijn voor alle schoolleiders echter even belangrijk. Wellicht is het dus niet zo dat schoolleiders van scholen met lage leerprestaties actiever zijn dan andere schoolleiders, maar

vooral dat schoolleiders van scholen met hoge leerprestaties wat minder gericht zijn op handelingen die volgens het concurrerende waarde model de effectiviteit van de school zouden moeten bevorderen.

De prestatiegerichtheid van scholen is positief gecorreleerd met het doorstroompercentage en het gemiddelde eindexamencijfer, terwijl de ontwikkelingsgerichtheid alleen samenhangt met het doorstroompercentage. De waardering van de school en het werkklimaat in de klassen hangen samen met het gemiddelde doorstroompercentage, maar niet met het gemiddelde examencijfer. Er is geen significante correlatie tussen de gemiddelde waardering van het onderwijs en de uitkomsten. Dit laatste kan echter het gevolg zijn van de beperkte betrouwbaarheid de geaggregeerde scores voor de waardering van het onderwijs (zie tabel 4.16).

Bij het bepalen van de samenhang tussen twee variabelen aan de hand van de correlatiecoëfficiënt, wordt geen rekening gehouden met intermediaire variabelen en contextvariabelen. Het weglaten van intermediaire variabelen zou echter geen gevolgen moeten hebben voor de grootte van het (totale) effect van de ene variabele op de andere (Saris & Stronkhorst, 1984, p.86; Verschuren, 1991, p.74). Het weglaten van contextvariabelen die van invloed zijn op beide variabelen, kan echter wel tot foutieve conclusies leiden (Verschuren, 1991, p.74). De correlaties tussen de schoolkenmerken zijn daarom opnieuw berekend aan de hand van ‘gecorrigeerde’ uitkomstgegevens. Hiervoor werden vergelijkende scores gebruikt van de opbrengstenkaarten voor het voortgezet onderwijs. Deze scores geven een indicatie van de relatieve positie van een school ten opzichte van andere scholen, waarbij rekening wordt gehouden met de achtergrond van leerlingen. Uit de correlaties met de vergelijkende scores (zie tabel 5.7) blijkt dat het positieve verband tussen prestatiegerichtheid en het gemiddelde eindexamencijfer niet meer significant is, wanneer rekening wordt gehouden met de samenstelling van de leerlingenpopulatie. Hieruit kan worden opgemaakt dat de mate van prestatiegerichtheid van een school (evenals het gemiddelde eindexamencijfer) is gekoppeld aan de samenstelling van de leerlingpopulatie. De negatieve samenhang van open systeem handelingen met het gemiddelde eindexamencijfer blijft bestaan, wanneer wordt gecontroleerd voor de samenstelling van de leerlingenpopulatie, evenals de positieve samenhang van prestatiegerichtheid, ontwikkelingsgerichtheid, waardering van de school en werkklimaat in de klassen met het doorstroompercentage.

Uit deze globale verkenning van de verbanden tussen schoolkenmerken en de uitkomsten kan worden geconcludeerd dat de schoolcultuurtypen en de schoolbeleving van leerlingen vooral samenhangen met het gemiddelde doorstroompercentage en niet met het gemiddelde eindexamencijfer. Maslowski (2001) vond in zijn onderzoek ook geen relaties tussen de schoolcultuur en het gemiddelde eindexamencijfer (gecorrigeerd voor de samenstelling van de leerlingenpopulatie), maar wel een samenhang van rationele doel cultuur en interne proces cultuur met het doorstroompercentage. Hij concludeert dat “scholen die sterk gericht zijn op stabiliteit en continuïteit, alsmede scholen die sterk prestatiegericht zijn, betere doorstroomcijfers behalen dan scholen die aan de genoemde waarden minder belang toekennen” (Maslowski, 2001, p.163). Verder merkt hij op, dat verwacht zou worden dat een prestatiegerichte (rationele doel) cultuur leidt tot hogere leerprestaties, terwijl hogere leerprestaties weer tot betere doorstroompercentages zouden leiden (Maslowski, 2001, p.116). Aangezien er geen verband werd gevonden tussen schoolcultuur en leerprestaties, kan het doorstroompercentage dus hoogstens ten dele zijn gebaseerd op leerprestaties. Maslowski (2001, p. 113) geeft aan dat de mate waarin leerlingen doorstromen ook afhankelijk is van de houding van docenten tegenover ‘zittenblijven’. Het doorstroompercentage zou dus tevens

gerelateerd zijn aan waarden die de school en docenten aanhangen. Een ordelijke en op leren gerichte atmosfeer zou kunnen bijdragen aan het beperken van de uitval van leerlingen, doordat leerlingen snel worden gecorrigeerd wanneer zij zich niet aan de regels houden. Uit de correlaties tussen schoolkenmerken en uitkomsten in dit onderzoek (tabel 5.7) blijkt echter dat niet alleen prestatiegerichte, maar ook sterk ontwikkelingsgerichte scholen, die veel waarde toekennen aan samenwerking, professionele ontwikkeling en vernieuwing, betere doorstroomcijfers behalen, terwijl er geen effect is op de gemiddelde eindexamencijfers. Kennelijk slagen deze scholen er eveneens in om zittenblijven en uitval te beperken, zodat leerlingen een hogere kans hebben om zonder vertraging het diploma te behalen.

De eerste verkenning van de relaties tussen schoolkenmerken en de leerprestaties leverde voldoende aanknopingspunten op voor de toetsing van een causaal model. Tussen schoolkenmerken en uitkomstvariabelen werden correlaties gevonden die het bestaan van causale verbanden aannemelijk maken. Hiermee is de derde onderzoeksvraag beantwoord.

7.5 Een causaal model voor de invloed van schoolleiders op de leerprestaties

De grootte van een causaal verband tussen twee variabelen verandert niet door intermediaire variabelen. Een direct-effect-model voor de invloed van schoolleiders op de leerprestaties zou dus in principe tot een goede schatting van het effect kunnen leiden. Een correlatie tussen twee variabelen is echter op zichzelf nog geen bewijs voor een causale samenhang. De correlatie kan het gevolg zijn van de invloed van een derde (context-) variabele op beide variabelen. Bij het berekenen van correlaties tussen de schoolkenmerken die werden gemeten in dit onderzoek en vergelijkende uitkomstcores (zie tabel 5.7) werd gecontroleerd voor één contextvariabele, namelijk het percentage culturele minderheden. In Nederland is dit een van de belangrijkste covariabelen voor het verband tussen schoolkenmerken en de uitkomsten (zie Maslowski, 2001; Rekers-Mombarg, Lodewick & Bosker, 2000). In de relaties tussen schoolkenmerken en de uitkomsten en tussen de schoolkenmerken onderling kunnen echter ook nog andere contextvariabelen een rol spelen. Verder hoeft de afwezigheid van een correlatie tussen twee variabelen niet per se te betekenen dat er geen causale samenhang is. Een latente correlatie tussen twee variabelen kan bijvoorbeeld worden ‘onderdrukt’ door een derde variabele die negatief samenhangt met de ene en positief met de andere, waardoor het totale (directe + indirecte) effect wordt verlaagd (Verschuren, 1991, p.278). Een analyse van de correlaties tussen schoolkenmerken en uitkomsten geeft dus nog geen uitsluitend over de causale verbanden. Een direct-effect-model geeft bovendien ook geen inzicht in de wijze waarop de invloed van schoolleiders op de leerprestaties precies verloopt. De vierde en laatste onderzoeksvraag was daarom:

Wat zijn de directe en indirecte effecten van schoolleidersgedrag op de leerprestaties en hoe verlopen deze effecten langs intermediaire variabelen, gegeven de contextfactoren?

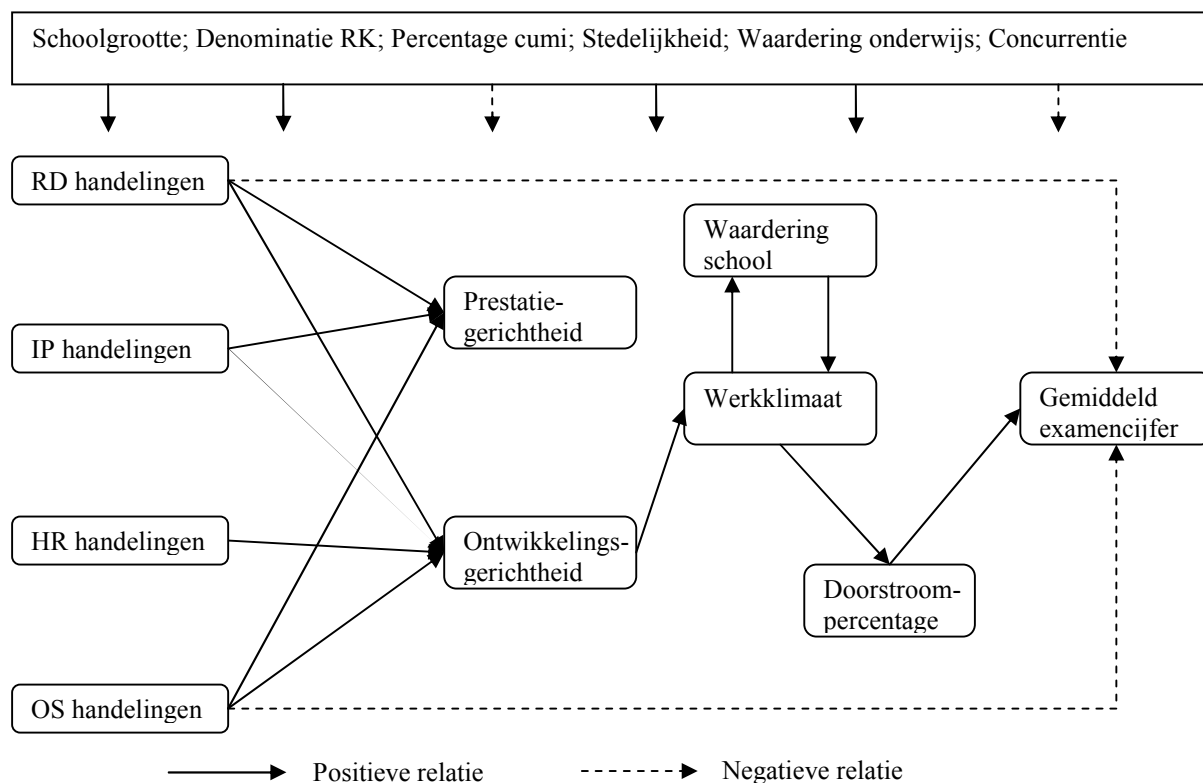
Om directe en indirecte causale verbanden in kaart te kunnen brengen, met de effecten van contextvariabelen die daarbij een rol spelen, werd een structureel (causaal) model opgesteld en getoetst aan de data. In het structurele model zijn de volgende contextvariabelen opgenomen: schoolgrootte, denominatie, percentage culturele minderheden in de havo-bovenbouw, stedelijkheid van de omgeving, waardering van onderwijs door leerlingen en de mate van concurrentie die wordt ervaren van andere scholen. Van het percentage culturele minderheden is bekend dat dit een belangrijke covariabele is (zie boven). Onderzoek naar de samenhang van schoolgrootte en denominatie met de leerprestaties leverde tot nu toe geen

eenduidige resultaten op (Bosker, 1992; Dronkers, 2004; Luyten, 1994; Opdenakker & Van Damme, 2006). Schoolgrootte en denominatie blijken samen te hangen met de prestatiegerichtheid en ontwikkelingsgerichtheid van scholen (zie paragraaf 7.4). Dit zou echter ook een effect kunnen zijn van een combinatie van context en leerlingkenmerken, waarbij de samenstelling van de leerlingenpopulatie een belangrijke rol speelt (vgl. Luyten et al., 2005). Een effect van de stedelijkheid van de omgeving zou eveneens gekoppeld kunnen zijn aan de achtergrond van de leerlingen (de gemiddelde sociaaleconomische status, zie bijv. Teddlie, Stringfield & Reynolds, 2000). De variabele ‘waardering van onderwijs’ verbindt de motivatie van de leerling aan de ondersteuning die de leerling ervaart van thuis. Volgens verschillende onderzoekers heeft de ondersteuning van thuis, of *family educational culture* een grotere voorspellende waarde dan de sociaaleconomische status (Lethwood, Aitken & Jantzi, 2000; Marzano, 2000; Walberg, 1984). De mate van concurrentie die wordt ervaren van andere scholen is opgenomen als contextvariabele, omdat scholen die veel concurrentie ervaren meer beleid zouden voeren dan andere scholen (Sleegers, 1991, p. 144). De concurrentiepositie zou bovendien iets kunnen zeggen over het imago van de school en zou ook gevolgen kunnen hebben voor de samenstelling van de leerlingenpopulatie.

Aangezien vooraf niet precies duidelijk was, welke effecten de contextvariabelen zouden kunnen hebben op verschillende variabelen in het model, werden in het onderzoeksmodel relaties verondersteld tussen de contextvariabelen en alle andere variabelen. Daarnaast werd aangenomen dat er verbanden zouden kunnen bestaan tussen alle vier type handelingen van schoolleiders (rationele doel, interne proces, human relations en open systeem) en zowel de prestatiegerichtheid als de ontwikkelingsgerichtheid van de school. Verwacht werd dat schoolleiders via de ontwikkelingsgerichtheid en prestatiegerichtheid van de school invloed zouden uitoefenen op het werkklimaat voor de leerlingen. Analoog aan het LOLSO-model (Mulford, Silins & Leithwood, 2004, zie figuur 2.2) werd verder verondersteld dat het werkklimaat voor leerlingen via de betrokkenheid van leerlingen bij de school (de variabele ‘waardering van de school’) een effect zou hebben op de uitkomsten.

Omdat het onderzoeksmodel veel te groot was om in één keer te analyseren, werd het in twee delen gesplitst. Eerst werd een optimaal model vastgesteld voor de invloed van contextvariabelen en leiderschapsvariabelen op de variabelen ‘prestatiegerichtheid’ en ‘ontwikkelingsgerichtheid’. Daarbij werden alle niet-significante relaties verwijderd. Vervolgens werd een optimaal model vastgesteld voor de invloed van de contextvariabelen en de schoolbelevingsvariabelen ‘werkklimaat’ en ‘waardering van de school’ op de uitkomsten. Hierbij bleek dat de waardering van de school wel verband houdt met het werkklimaat in de klassen, maar niet direct van invloed is op de uitkomsten. Tenslotte werden de twee deelmodellen gekoppeld, waarbij werd nagegaan in hoeverre er nog directe effecten van de handelingen van schoolleiders of de cultuurvariabelen op de uitkomsten toegevoegd zouden moeten worden. Het onderzoeksmodel werd twee maal getoetst: eerst met gegevens van docenten over de handelingen die zij waarnamen, daarna met gegevens van de schoolleiders zelf over de uitgevoerde handelingen. Beide keren werd een goed passend model verkregen. Toetsing aan de gegevens van de docenten leverde meer inzicht op in de relaties tussen schoolleidersgedrag en de cultuurvariabelen dan toetsing aan de gegevens van schoolleiders.

Een overzicht van het causale model dat werd geconstrueerd, is afgebeeld in figuur 7.1 (voor de padcoëfficiënten zie figuur 6.2).



Figuur 7.1 Model voor de effecten van de handelingen van schoolleiders op de leerprestaties

Uit een analyse van de effecten blijkt dat er kleine, maar significante, positieve effecten zijn van rationele doel, human relations en open systeem handelingen, via de ontwikkelingsgerichtheid van scholen en het werkklimaat in de klassen, op het doorstroompercentage. Er zijn geen significante indirecte effecten van de handelingen van schoolleiders op het gemiddelde eindexamencijfer, alleen negatieve directe effecten van zowel rationele doel als open systeem handelingen. Hiermee is de vierde onderzoeksvraag beantwoord. De kleine positieve indirecte relaties tussen schoolleidersgedrag en het doorstroompercentage zijn in een direct-effect-model, zonder intermediaire variabelen, niet zichtbaar (zie paragraaf 7.5 en tabel 5.7). Overigens is de structurele analysetechniek (*structural equation modeling*) ook gevoeliger, waardoor kleine effecten, zoals het indirecte effect van het werkklimaat in de klassen op het gemiddelde eindexamencijfer, toch significant blijken te zijn (vergelijk tabellen 5.7 en 6.3).

Uit figuur 7.1 is op te maken dat er wel een samenhang bestaat tussen schoolleidersgedrag en de prestatiegerichtheid van scholen, terwijl prestatiegerichtheid geen effect heeft op de uitkomsten. Ook hier blijkt dus dat de correlatie die werd gevonden tussen prestatiegerichtheid en het ruwe doorstroompercentage (zie paragraaf 7.4) kan worden toegeschreven aan de invloed van contextkenmerken op beide variabelen. Verder blijkt uit een analyse van de relaties in het structurele model, dat er een verband is tussen de waardering van de school en het werkklimaat in de klassen, maar geen significant (indirect) effect van de waardering van de school op het doorstroompercentage of het gemiddelde eindexamencijfer. De correlatie die eerder werd gevonden tussen de waardering van de school en het doorstroompercentage (zie paragraaf 7.4) is waarschijnlijk het gevolg van de samenhang van zowel de waardering van de school, als het doorstroompercentage, met het werkklimaat voor leerlingen. Een effect van de betrokkenheid van leerlingen op het rendement, zoals dat in het

LOLSO-model (Mulford & Silins, 2003) wordt voorgesteld, kan hier dus niet worden onderschreven.

De effecten van de contextvariabelen op de andere variabelen in het model zijn complex en daarom niet weergegeven in figuur 7.1. De concurrentie die wordt ervaren van andere scholen en het percentage culturele minderheden blijken negatief samen te hangen met verschillende schoolkenmerken en de uitkomsten. De overige contextvariabelen (schoolgrootte, denominatie, mate van stedelijkheid en waardering van het onderwijs) vertonen voornamelijk een positieve samenhang met diverse schoolkenmerken en een wisselende invloed op de uitkomsten. Het model werd voor verschillende denominaties apart getest, maar alleen de rooms-katholieke denominatie blijkt een substantiële invloed te hebben. Het doorstroompercentage wordt positief beïnvloed door de rooms-katholieke denominatie en de waardering van het onderwijs, terwijl het percentage culturele minderheden en de ervaren concurrentie een behoorlijke negatieve impact hebben. Het percentage culturele minderheden heeft ook een aanzienlijke negatieve invloed op het gemiddelde eindexamencijfer, terwijl schoolgrootte, mate van verstedelijking en concurrentie een licht negatieve invloed hebben. De verwachting dat de mate van concurrentie die wordt ervaren van andere scholen een indicatie zou zijn voor de samenstelling van de leerlingenpopulatie (zie boven) lijkt uit te komen. Scholen die veel concurrentie ervaren van andere scholen (evenals scholen in sterk verstedelijkte gebieden en zeer grote scholen) zouden wellicht minder kansrijke leerlingen kunnen aantrekken en daardoor minder hoge eindexamencijfers behalen.

Uit de verschillen in de uitkomsten van de analyse van correlaties (paragraaf 7.4) en van het causale model (hierboven) blijkt dat de wijze waarop contextvariabelen worden verdisconteerd, van grote invloed is op de resultaten. Bij de analyse van de correlaties tussen schoolkenmerken en de uitkomsten werd alleen gecontroleerd voor het percentage culturele minderheden, terwijl in het causale model veel meer contextvariabelen in beschouwing werden genomen. Om de impact van de contextvariabelen nader te onderzoeken, werd een structureel model geanalyseerd zonder contextvariabelen. Uit de analyse van dit model blijkt dat de effecten verschillen voor katholieke en niet-katholieke scholen. Voor katholieke scholen is er in een model zonder contextvariabelen een significant positief effect van prestatiegerichtheid op het doorstroompercentage. Dit effect neemt af wanneer contextvariabelen in het model worden opgenomen. Hierbij speelt vooral de concurrentie van andere scholen een belangrijke rol. Voor niet-katholieke scholen is er geen significant verband tussen prestatiegerichtheid en het doorstroompercentage, met of zonder contextvariabelen. De correlatie die werd gevonden tussen prestatiegerichtheid en het doorstroompercentage (zie paragraaf 7.4) is dus waarschijnlijk vooral toe te schrijven aan katholieke scholen. Vermoedelijk is de samenstelling van de leerlingenpopulatie op katholieke scholen de oorzaak van zowel een (gemiddeld) hogere prestatiegerichtheid als een hoger doorstroompercentage. Bryk & Raudenbusch (1992) onderzochten de verschillen tussen katholieke en openbare scholen. Zij constateerden dat de verschillen in leerresultaten tussen beide denominaties vooral werden veroorzaakt door het 'contextuele' effect op schoolniveau van de sociaaleconomische status van leerlingen (zie Scheerens & Bosker, 1997, p.63). Verondersteld wordt dat er meer leerlingen met een hoge sociaaleconomische status naar katholieke scholen gaan (Deinum, 2000, p.28). In dit onderzoek zijn er echter geen significante verschillen tussen de denominaties en het percentage leerlingen uit een milieu met een hoge, gemiddelde of lage sociaaleconomische status geconstateerd. Wel is het percentage culturele minderheden op rooms-katholieke en protestants-christelijke scholen significant lager dan op andere scholen. Voor katholieke scholen is er geen correlatie tussen het percentage culturele minderheden en het percentage leerlingen uit een milieu met een

lage sociaaleconomische status (0.01), voor niet-katholieke scholen is die correlatie hoog (0.60). Hoewel de verschillen tussen katholieke en niet-katholieke scholen dus waarschijnlijk in belangrijke mate zijn toe te schrijven aan verschillen in de samenstelling van de leerlingenpopulatie, is het natuurlijk niet uit te sluiten dat er binnen katholieke scholen een gunstige schoolcultuur wordt gecreëerd. In Nederland was er volgens Dronkers (2004) tussen de jaren zestig en de jaren negentig sprake van een grotere effectiviteit van het bijzonder onderwijs ten opzichte van het openbaar onderwijs, maar deze verschillen zouden inmiddels verdwenen zijn, door fusies van scholen en de daaruit voortkomende schaalvergroting. Verschillen tussen het openbare en het bijzondere onderwijs zouden een gevolg kunnen zijn van de mate waarin scholen zeggenschap hebben over de gang van zaken binnen de school (Teddlie, Stringfield & Reynolds, 2000). Rowan, Raudenbush en Kang (1991) merken op dat katholieke scholen vaker ondersteunend leiderschap, gedeelde besluitvorming en een hoog niveau van samenwerking tussen de staf te vinden is. Uit een analyse van het structurele model blijkt dat de katholieke denominatie positief samenhangt met zowel rationele doel als human relations handelingen van schoolleiders. Hierbij gaat het om (gedeelde) doelstellingen en ondersteuning van het werk van docenten. Deze uitkomsten sluiten dus goed aan bij de resultaten van Rowan et al. (1991).

7.6 Conclusies en reflectie

Bij de aanvang van dit onderzoek werd verondersteld dat een indirect-effect-model beter inzicht zou geven in de impact van schoolleiders op de leerprestaties dan een direct-effect-model. Dat blijkt ook zo te zijn, vooral omdat in het indirecte (causale) model de relaties tussen schoolkenmerken en opbrengsten 'uiteengerafeld' worden. Daarbij blijken er niet alleen indirecte, maar ook directe en zowel positieve, als negatieve effecten van schoolleidersgedrag op de opbrengsten te zijn. Naast kleine, maar significante, positieve indirecte effecten van schoolleidersgedrag op het gemiddelde doorstroompercentage, werden grotere, negatieve, directe effecten gevonden van zowel rationele doel als open systeem handelingen, op de gemiddelde eindexamencijfers. Directe effecten van schoolleiderschap op de uitkomsten werden eigenlijk niet verwacht, omdat deze in het Nederlandse voortgezet onderwijs tot nu toe niet onomstotelijk aangetoond zijn (Witziers, Bosker & Kruger, 2003). Een direct effect in een causaal model hoeft overigens niet per se een 'echt' direct effect te zijn, het kan ook een 'onverklaard' effect zijn, omdat de intermediaire variabelen voor dit effect niet zijn opgenomen in het model. Negatieve effecten worden wel vaker gevonden in (internationaal) onderzoek naar de effecten van schoolleiderschap op de leerprestaties (Witziers, Bosker & Krüger, 2003). Daarbij wordt aangenomen dat schoolleiders van scholen met lage leerprestaties meer geneigd zijn om actie te ondernemen om de school te verbeteren en dat het hier dus eigenlijk gaat om een reactie van de schoolleiders op de leerprestaties. Uit een vergelijking van het gedrag van schoolleiders op scholen met hoge, gemiddelde of lage leerprestaties in dit onderzoek, zou echter ook opgemaakt kunnen worden dat het negatieve effect voornamelijk een gevolg is van het feit dat schoolleiders van scholen met goede leerprestaties minder geneigd zijn om dit soort acties te ondernemen. Deze schoolleiders zijn vooral gericht op het bewaren van de stabiliteit en continuïteit in de schoolorganisatie.

Schoolleiders blijken een substantiële invloed te hebben op de ontwikkelingsgerichtheid van scholen. Door docenten waargenomen handelingen van schoolleiders verklaren meer dan 80% van de variantie in ontwikkelingsgerichtheid. Rationele doel en open systeem handelingen hebben het grootste effect, gevolgd door human relations handelingen en tenslotte interne proces handelingen. De prestatiegerichtheid van scholen hangt veel minder samen met schoolleidersgedrag. Er zijn wel bescheiden relaties met rationele doel, interne proces en open

systeem handelingen, maar human relations handelingen spelen hierbij geen rol. Wanneer rekening wordt gehouden met een aantal relevante contextvariabelen, blijkt de prestatiegerichtheid van scholen niet samen te hangen met de uitkomsten, terwijl er wel een indirect effect is van ontwikkelingsgerichtheid op het doorstroompercentage. Het ontbreken van een verband tussen prestatiegerichtheid en de uitkomsten is opmerkelijk, omdat er in het algemeen van wordt uitgegaan dat een prestatiegerichte atmosfeer de schooleffectiviteit bevordert (Scheerens & Bosker, 1997). Verschillende onderzoekers wezen er echter al op dat de prestatiegerichtheid van een school gerelateerd zou kunnen zijn aan de samenstelling van de leerlingenpopulatie (vgl. Luyten et al., 2005). Op scholen met veel kansrijke leerlingen zijn hoge verwachtingen van leerlingen en goede leerprestaties bijna vanzelfsprekend.

Met de variabelen ontwikkelingsgerichtheid en prestatiegerichtheid nu wel of niet de schoolcultuur? Maslowski (2001, p.130) wijst erop, dat het vaak moeilijk is om cultuur te onderscheiden van 'klimaat' of van procedures en structuren, bijvoorbeeld wanneer aspecten van gedrag worden gemeten. Om de cultuur te bepalen zou de meetmethode alleen gericht moeten zijn op onderliggende waarden. In dit onderzoek wordt een combinatie gemeten van onderliggende waarden en de neerslag in de praktijk. Feitelijk gaat het hier dus niet om 'zuivere' cultuurvariabelen, hoewel prestatiegerichtheid en ontwikkelingsgerichtheid zijn terug te vinden in de cultuurtypen die werden onderscheiden door Maslowski (2001) en Houtveen et al. (1995). De vraag is of 'zuivere' cultuurvariabelen veel meer informatie zouden opleveren over het effect van schoolleiders op de leerprestaties. Volgens eerder onderzoek heeft de schoolorganisatie een grote impact op de schoolcultuur (zie bijv. Heck, Larsen & Marcoulides, 1990; Krüger, Witziers & Slegers, 2007). Effecten van de schoolcultuur en de schoolorganisatie zijn waarschijnlijk moeilijk van elkaar te onderscheiden.

De variabele ontwikkelingsgerichtheid meet in welke mate belang wordt gehecht aan samenwerking, ontwikkeling en vernieuwing. Een ontwikkelingsgerichte school vertoont overeenkomsten met de 'lerende organisatie' in het LOLSO-model van de Australische onderzoekers Mulford en Silins (2003). Kenmerken van de lerende organisatie in dit model zijn: onderling vertrouwen en samenwerking, het nemen van initiatieven en risico's, een gedeelde missie en voortdurende professionele ontwikkeling. Volgens het LOLSO-model is 'organisatieleren' een belangrijke intermediaire variabele voor de invloed van schoolleiders op het werk van docenten (Mulford, Silins & Leithwood, 2004). Organisatieleren zou cruciaal zijn om schoolverbetering te kunnen realiseren (zie ook Hallinger, 1999). Volgens Mulford, Silins & Leithwood (2004) wordt de lerende organisatie bevorderd door transformationeel en gedistribueerd leiderschap. Een transformationele schoolleider is gericht op individuele ondersteuning, een zorgzame en veilige schoolcultuur, delegatie en gedeelde besluitvorming, consensus over visie en doelen en hoge verwachtingen van staf en leerlingen. De transformationele leider verwacht effectiviteit en vernieuwingsgezindheid van de staf en zorgt voor intellectuele stimulans. Bij transformationeel (en gedistribueerd) leiderschap gaat het niet zozeer om de visie en strategieën van een "great man or woman" maar om ondersteuning, gedeelde verantwoordelijkheid en consensus (Mulford, Silins & Leithwood, 2004, p.14).

In hoeverre levert dit onderzoek nu aanwijzingen op, dat schoolleiders die ontwikkelingsgerichtheid bevorderen ook transformationele schoolleiders zijn? Uit de samenhang tussen ontwikkelingsgerichtheid en schoolleidersgedrag blijkt, dat een schoolleider die werkt aan duidelijke en gedeelde doelstellingen en die door contacten met belanghebbenden (buiten de school) zorgt voor voldoende ondersteuning en middelen om

eigen keuzes te maken, in belangrijke mate bijdraagt aan de ontwikkelingsgerichtheid van de school. Aandacht van de schoolleider voor de ondersteuning van docenten bij hun werk in de klas en voor een ordelijke sfeer op school horen hier echter ook bij. In deze handelingen zijn aspecten van transformationeel leiderschap te herkennen, maar omdat dit onderzoek niet specifiek was gericht op het analyseren van transformationeel leiderschap gaat het te ver om hier conclusies over te trekken. Opvallend is wel dat ‘ontwikkelingsgerichtheid’ in het structurele model is gekoppeld aan een externe focus van locatieleiders, terwijl uit de beschrijving van transformationeel leiderschap (zie boven) een overwegend interne gerichtheid naar voren komt. Op basis van de bijdragen van de verschillende oriëntaties van schoolleiders aan de ontwikkelingsgerichtheid van de school, zou een profiel kunnen worden opgesteld voor de ‘ontwikkelingsgerichte’ locatieleider, met een relatief hoge score op de rationele doel en open systeem oriëntatie, een gematigde score voor ‘human relations’ en een bescheiden score voor de interne proces oriëntatie. Een dergelijk profiel wijst volgens Quinn et al. (2003, p.399) op een “vindingrijke en productieve manager” die niet uitblinkt in het onderdeel interne processen. Dit wordt door Quinn c.s. beschouwd als een effectief profiel. Een ontwikkelingsgerichte school is dus waarschijnlijk gebaat bij een vindingrijke en productieve locatieleider. De taken van de schoolleiding zijn over het algemeen verdeeld over een team van schoolleiders. Verschillende functionarissen coördineren interne processen en oefenen onderwijskundig leiderschap uit. De eindverantwoordelijke locatieleider zal echter in belangrijke mate de richting moeten bepalen, onderhandelen met interne en externe ‘stakeholders’, voldoende middelen moeten verwerven en zorgdragen voor ontwikkeling en aanpassing. Een effectieve, vindingrijke en productieve, locatieleider kan de ontwikkelingsgerichtheid van een school bevorderen en daarmee het werkklimaat voor de leerlingen.

Het LOLSO-model voorspelt dat het werkklimaat in de klassen via de ‘participatie’ van leerlingen in schoolactiviteiten een direct effect heeft op de leerprestaties en via ‘identificatie’ van leerlingen met de school een indirect effect, door verminderde schooluitval. Daarbij zou de schoolleider, via het werk van docenten, met name invloed hebben op de mate waarin leerlingen zich identificeren met de school en niet (significant) op de participatie ((Silins & Mulford, 2003; Leithwood & Jantzi, 2000b). In dit onderzoek werd daarom alleen de mate waarin leerlingen zich identificeren met de school gemeten door middel van de variabele ‘waardering van de school’. Bij de toetsing van het causale model bleek echter dat de waardering van de school wel sterk samenhangt met de beoordeling door leerlingen van het werk van docenten (‘werkklimaat’), maar dat er geen significant (direct of indirect) effect is op de uitkomsten. De mate waarin leerlingen zich thuisvoelen op school blijkt dus in dit onderzoek geen intermediaire variabele te zijn voor de invloed van het werk van docenten op het doorstroompercentage of de leerprestaties. Deze uitkomst kan wellicht worden verklaard door de relatief hoge gemiddelde score van de havo-5 leerlingen voor de variabele ‘waardering van de school’. Deze leerlingen voelen zich meestal gewoon goed thuis op hun school. Zoals al eerder werd opgemerkt is het niet uit te sluiten dat leerlingen die zich minder thuis voelden al voor de vijfde klas zijn afgehaakt, waardoor de score van de vijfdeklassers hoger zou kunnen dan de gemiddelde score van alle leerlingen op een school. De havo-5 leerlingen zijn wat kritischer over het werkklimaat in de klassen, maar ook niet echt ontevreden. Het werkklimaat in de klassen, zoals beoordeeld door de vijfdeklassers, heeft een klein positief effect op het gemiddelde doorstroompercentage in de bovenbouw en via het doorstroompercentage een heel klein effect op het gemiddelde eindexamencijfer.

Door het bevorderen van de ontwikkelingsgerichtheid van de school kan de schoolleider het werkklimaat voor leerlingen verbeteren, waardoor het doorstroompercentage in de bovenbouw wordt vergroot. Een groter doorstroompercentage zorgt voor een hoger gemiddeld

eindexamencijfer, maar het indirecte effect van de schoolleider op het eindexamencijfer is uiteindelijk zo klein dat het niet significant meer is. Dit wordt onder meer veroorzaakt door het aantal intermediaire variabelen. Het indirecte effect van een (onafhankelijke) variabele op een andere (afhankelijke) variabele wordt kleiner, naarmate het aantal intermediaire variabelen toeneemt. Het grote aantal intermediaire variabelen voor het effect van schoolleiderschap op de leerprestaties houdt uiteraard verband met de (organisatorische) afstand tussen de locatieleider en de leerling in het voortgezet onderwijs. Het indirecte effect van de locatieleider op de leerprestaties is hierdoor eigenlijk per definitie klein.

Contextvariabelen spelen in het causale model een belangrijke rol. Katholieke scholen zijn prestatiegerichter dan andere scholen (met uitzondering van protestants-christelijke scholen) en hebben een groter rendement en hogere leerprestaties. Het verband tussen prestatiegerichtheid en de uitkomsten voor katholieke scholen vermindert echter wanneer contextvariabelen in beschouwing worden genomen. Hierbij speelt de samenstelling van de leerlingenpopulatie waarschijnlijk een rol. Scholen die veel concurrentie ondervinden van andere scholen kunnen bijvoorbeeld minder selectief zijn bij het aantrekken van leerlingen. Ook uit een onderzoek van Krüger, Witziers en Slegers (2007) naar de impact van schoolleiderschap bleek dat de wijze waarop de effecten van contextvariabelen in het model werden verdisconteerd, van grote invloed was op de uitkomsten. Effecten van de schoolcultuur en de schoolorganisatie op de betrokkenheid van leerlingen verdwenen, zodra rekening werd gehouden met de invloed van contextvariabelen. Leithwood en Levin (2005) wijzen erop dat 'modererende' (context-)variabelen het effect van leiderschap zouden kunnen verkleinen, neutraliseren of zelfs vergroten. Zij onderscheiden vijf categorieën van moderatoren: variabelen het niveau van leerlingen, docenten, schoolleiders, de schoolorganisatie en de omgeving van de school. Het gebrek aan aandacht voor deze modererende variabelen zou een belangrijke oorzaak kunnen zijn voor de tegenstrijdige resultaten in het onderzoek naar schoolleiderschap. In onderzoek waar modererende contextvariabelen wel werden meegenomen, blijkt de keuze bovendien vaak weinig theoretisch onderbouwd te zijn (Leithwood & Levin, 2005).

Tenslotte kan nog de vraag worden gesteld in hoeverre de keuze van contextvariabelen in dit onderzoek voldoende is afgewogen. Het was de bedoeling om in het causale model vooral contextvariabelen op te nemen die van invloed zouden kunnen zijn op leerprestaties en op schoolleiderschap. Het niet opnemen van variabelen die een effect hebben op twee of meer variabelen in een model kan leiden tot foutieve conclusies over de relaties tussen de variabelen die wel in het model zijn opgenomen (zie ook paragraaf 7.4). Contextvariabelen die slechts invloed hebben op één van de variabelen in het model, kunnen echter worden weggelaten (Verschuren, 1991, p.74). Variabelen zoals sekse of opleiding, die een effect hebben op het gedrag van schoolleiders, maar niet op de leerprestaties van leerlingen, zijn dus voor dit onderzoek niet van belang. De belangrijkste covariabelen voor leerprestaties hebben betrekking op de achtergrond van de leerlingen. Door het onderzoek te richten op de havo-opleiding, met relatief geringe verschillen in aanleg tussen de eindexamenkandidaten, werd ernaar gestreefd om de invloed van de achtergrond van leerlingen in het model te verkleinen. Verder zijn in het causale model zoveel mogelijk relevante achtergrondvariabelen op schoolniveau opgenomen. De grootte van het causale model dat getoetst kon worden, was echter beperkt door het relatief kleine aantal cases (103 scholen).

7.7 Aanbevelingen voor vervolgonderzoek

In dit onderzoek werd gebruik gemaakt van het generieke concurrerende waarden model om zowel schoolleiderschap als schoolcultuur en de relaties daartussen in kaart te brengen. Daarbij kwam een model naar voren voor een effectieve schoolleider, met kenmerken van transformationeel leiderschap, gekoppeld aan een duidelijke externe gerichtheid. Deze ‘vindingrijke en productieve’ schoolleider zou de ‘ontwikkelingsgerichtheid’ van een school kunnen bevorderen, waardoor de uitkomsten worden verbeterd. Het verdient aanbeveling om sterk ontwikkelingsgerichte scholen nader te onderzoeken. Wat zijn precies de kenmerken van deze scholen en van de schoolleiders? Sterk ontwikkelingsgerichte scholen zijn waarschijnlijk niet de scholen met de hoogste leerprestaties. Scholen met hoge leerprestaties lijken vooral te zijn gericht op het in stand houden van de bestaande situatie. In een vervolgonderzoek zou kunnen worden nagegaan in hoeverre er op ontwikkelingsgerichte scholen sprake is van transformationeel en gedistribueerd leiderschap. Hierbij zou de gehele schoolleiding moeten worden betrokken en niet alleen de locatieleider. Er zijn flinke verschillen tussen scholen in de verdeling van leiderschapstaken binnen de schoolleiding. Deze taakverdeling zou expliciet in het onderzoek meegenomen moeten worden, om verschillen tussen schoolleiders beter te kunnen analyseren. Overigens zijn er aanwijzingen dat ‘gedistribueerd leiderschap’, *total leadership*, of *collective leadership*, waarbij het gecombineerde effect van vele bronnen van leiderschap (niet alleen van de eindverantwoordelijke schoolleider, maar ook van andere leidinggevendenden, docenten, ouders, leerlingen en andere functionarissen, individueel of in teams) wordt bepaald, een grotere (indirecte) invloed heeft op de leerresultaten, dan leiderschap van de eindverantwoordelijke schoolleider alleen (vgl. Leithwood et al., 2006b, Leithwood & Mascall, 2008). *Collective leadership* zou tot vier procent van de totale variantie in de leerresultaten kunnen verklaren (Leithwood & Mascall, 2008). Aangezien verondersteld wordt dat het totale effect van schoolfactoren op de leerprestaties (*main school effect*, vgl. Luyten, 1994) niet meer dan vier tot negen procent van de variantie in de leerresultaten verklaart (zie Scheerens & Bosker, 1997, p.79), zou het grootste deel van het schooleffect moeten worden toegeschreven aan ‘gecombineerd leiderschap’. De vraag is echter wat dit gecombineerde leiderschap precies inhoudt. In het onderzoek van Leithwood en Mascall (2008) gaat het om een unidimensionele maat, die alleen bestaat uit een score voor de invloed op de besluitvorming. Verder wordt in dit onderzoek slechts in beperkte mate rekening gehouden met leerlingkenmerken en is niet duidelijk in hoeverre de schoolcompositie nog meespeelt in het effect van *collective leadership*. Tenslotte blijkt dat ook in het gecombineerde leiderschap de formele schoolleider nog steeds een belangrijke rol speelt. Leithwood en Mascall (2008) concluderen dan ook:

Recognizing that some form of leadership distribution has always been a necessary feature of school and other professional organizations, there is (as yet) no empirical justification for advocating more planful distribution of leadership as a strategy for organizational improvement beyond those important efforts to enlist the full range of capacities and commitments found within school organizations. (p. 557)

De uitkomsten van het onderzoek naar schoolleiderschap blijken slecht vergelijkbaar te zijn door een gebrek aan standaardisatie in de meetmethoden. Het zou daarom aanbeveling verdienen om te werken aan een gevalideerd gemeenschappelijk instrumentarium, bijvoorbeeld op basis van ‘best practices’.

De (ontbrekende) samenhang tussen prestatiegerichtheid en de leerprestaties, in relatie tot de samenstelling van de leerlingenpopulatie zou ook een onderwerp kunnen zijn voor nader

onderzoek. Daarbij zou vooral de situatie op katholieke scholen, die prestatiegerichter zijn dan andere scholen en hogere doorstroompercentages en leerprestaties behalen, nauwkeuriger geanalyseerd moeten worden.

De relatie tussen het doorstroompercentage (rendementsmaat) en het gemiddelde eindexamencijfer (effectiviteitsmaat) vormt eveneens een interessant onderwerp. Uit dit onderzoek blijkt dat het doorstroompercentage een klein positief effect heeft op het gemiddelde eindexamencijfer. Het is echter voor te stellen dat een hoger doorstroompercentage juist tot een lager gemiddeld eindexamencijfer zou leiden, wanneer meer leerlingen met matige cijfers ‘het voordeel van de twijfel’ zouden krijgen. Anderzijds zou een positieve atmosfeer op school en een goed werkklimaat in de klassen tot hogere leerprestaties kunnen leiden en minder uitval van leerlingen, waardoor het doorstroompercentage wordt verhoogd. Wellicht spelen hier wederkerige, positieve en negatieve relaties tussen het doorstroompercentage en het examencijfer door elkaar. Het gemiddelde eindexamencijfer kan worden beschouwd als een maat voor schooleffectiviteit en het doorstroompercentage van de derde klas naar het diploma als een maat voor efficiëntie (Maslowski, 2001, p. 107). Het doorstroompercentage is niet helemaal objectief, omdat het bijvoorbeeld kan afhangen van de houding van de school tegenover ‘zittenblijven’ (Maslowski, 2001, p.116). Het geeft echter wel een beeld van de selectiviteit van de school. Voor een juiste beoordeling van de prestaties van een school zijn beide uitkomstmaten dus van belang (vgl. Scheerens, 1989, p.157).

Leithwood en Levin (2005) geven aan dat er significante indirecte effecten zijn van transformationeel en gedistribueerd leiderschap op de leerprestaties. Miller en Rowan (2006) leiden uit de contingentiebenadering²² af dat ‘organische vormen van management’ vaker zouden ontstaan in een dynamische omgeving en daar ook effectiever zouden zijn, dan meer directieve en hiërarchische vormen van management. In een grootschalig onderzoek vinden zij hiervoor echter geen bewijs. Nader onderzoek van de verschillen tussen prestatiegerichte katholieke scholen die weinig concurrentie ervaren en andere scholen, die zich in een meer dynamische omgeving bevinden, zou hier licht op kunnen werpen. Barker (2007) concludeert uit een *case study* dat opvallende resultaten die werden verkregen onder leiding van een bijzondere transformationele schoolleider, voor een belangrijk deel moesten worden toegeschreven aan een veranderde samenstelling van de leerlingenpopulatie. Het aantrekken van meer kansrijke studenten kan overigens ook als een positief resultaat worden gezien.

Bij de toetsing van meetmodellen en structurele modellen in dit onderzoek vormde het aantal deelnemende scholen (103) een beperking voor de hoeveelheid variabelen in een model. Per variabele zijn zeker tien tot twintig cases nodig (zie Garson, z.j., b). Voor de toetsing van het structurele model met contextvariabelen, leiderschapsvariabelen, intermediaire variabelen en uitkomstvariabelen werd een oplossing gevonden, door het model in twee delen te optimaliseren en daarna samen te voegen. Deze strategie van het opsplitsen van modellen is wellicht nog verder door te voeren, waardoor de mogelijkheid ontstaat om de geldigheid van (deel-)modellen te testen door middel van kruisvalidering, omdat voor kleinere modellen minder cases nodig zijn. Hiermee vermindert het risico van kanskapitalisatie. Kanskapitalisatie kan tot gevolg hebben dat het resultaat alleen geldig is voor de steekproef en

²² Zie Scheerens, 1989, p. 141: “De contingentiebenadering komt erop neer dat de effectiviteit van een organisatiestructuur afhankelijk is van contextkenmerken (zoals de omgeving, de technologie, de grootte of de leeftijd van de organisatie. De basisgedachte van de contingentietheorie is in strijd met de idee dat er één in allerlei situaties toepasbaar effectiviteitsmodel zou kunnen worden geformuleerd.”

niet generaliseerbaar naar de gehele populatie (vgl. Kelloway, 1998; MacCallum, Roznowski & Necowitz, 1992; Stevens, 2002).

In dit onderzoek werd een aanvang gemaakt met meerniveau analyses. Hierbij bleek dat software en procedures voor meerniveau *structural equation modeling* nog volop in ontwikkeling waren. Vanwege de problemen die dit opleverde, werd ervoor gekozen om over te stappen op een éénniveau benadering, waarbij gegevens werden geaggregeerd naar schoolniveau. Dit leidt tot een verlies aan variantie. De betrouwbaarheid van de geaggregeerde variabelen bleek echter acceptabel te zijn, vanwege het grote aantal docenten en leerlingen per school. Bij de toetsing van het causale model werden de analyses verder vereenvoudigd, door aan te nemen dat alle latente variabelen zonder meetfouten werden gemeten. Dit zou (kleine) afwijkingen kunnen opleveren in de parameterschattingen (Jöreskog & Sörbom, 1993a, p. 37). Er zijn methoden beschikbaar om de meetfout(-variantie) te schatten, bijvoorbeeld aan de hand van de betrouwbaarheid van de variabelen. Vanuit methodologisch oogpunt zou het interessant kunnen zijn om de uitkomsten van de relatief eenvoudige benadering in dit onderzoek te vergelijken met die van een meer geavanceerde benadering.

Door gebruik te maken van meerniveau technieken bij de structurele analyse zouden de kenmerken van leerlingen beter verdisconteerd kunnen worden (zie bijv. De Maeyer et al., 2007). In een meerniveau model kunnen individuele leerprestaties en achtergrondkenmerken worden opgenomen. Het verwerven van individuele leerlingengegevens viel echter buiten het bestek van dit onderzoek. Voor het gebruik van dit soort gegevens moet toestemming worden gevraagd aan leerlingen of ouders. Bovendien is het een extra belasting voor scholen om deze gegevens aan te leveren en dit zou de bereidheid om deel te nemen aan het onderzoek niet hebben bevorderd. Tenslotte zou het een heel karwei zijn geweest om alle verschillende leerlingengegevens aan elkaar te koppelen. Door de invoering van het onderwijsnummer, vanaf 2002, worden nu gegevensbestanden opgebouwd in het *Basisregister Onderwijs* waarin individuele leerlingprestaties zijn gekoppeld aan achtergrondgegevens. Voor toekomstig onderzoek biedt dit uiteraard perspectieven. Het gebruik van gegevens uit het Basisregister Onderwijs is echter strikt gereguleerd in de *Wet Onderwijsnummer*. Slechts een aantal instanties, zoals de Inspectie van het Onderwijs en het Centraal Bureau voor de Statistiek, hebben toegang tot deze gegevens. Om gebruik te kunnen maken van gegevens uit het Basisregister Onderwijs, ligt samenwerking met deze instanties dus voor de hand. Het combineren van wederzijdse expertise zou daarbij voordelen kunnen opleveren. Overigens zijn er over de schoolbeleving van leerlingen in dit onderzoek individuele data verzameld, terwijl er eveneens individuele achtergrondvariabelen werden gemeten. Een verdergaande meerniveau benadering zou dus op basis van de beschikbare gegevens ook al mogelijk zijn.

Een belangrijke beperking van dit onderzoek is dat het *cross-sectioneel* is in plaats van *longitudinaal*. Dat betekent dat causale relaties alleen gebaseerd zijn op de theorie. Het onderzoek laat een ‘dwarsdoorsnede’ of een ‘snap shot’ zien van de situatie. Het is mogelijk dat toevallig tijdelijke verbanden zichtbaar zijn. Er is geen zekerheid over de causaliteit, omdat er slechts op één tijdstip is gemeten. De negatieve directe relaties tussen schoolleidersgedrag en de leerprestaties zouden bijvoorbeeld een gevolg kunnen zijn van de nadruk die de laatste jaren wordt gelegd op het afleggen van verantwoording door scholen over hun resultaten (*accountability*). Schoolleiders van slecht presterende scholen zijn geneigd om acties te ondernemen om de school te verbeteren, schoolleiders van goed presterende scholen minder. De gegevens die beschikbaar zijn via de opbrengstenkaarten voor het voortgezet onderwijs maken het in principe mogelijk om de resultaten over een aantal

jaren te bestuderen. Een beperkte longitudinale benadering is dus realiseerbaar. Verder zijn er ook gegevens beschikbaar over andere vakken of combinaties van vakken, zoals de moderne vreemde talen en de exacte vakken. Het zou nuttig zijn om het structurele model nogmaals te testen aan de hand van de eindexamencijfers voor deze vakkenclusters. De effecten van school- en leerlingkenmerken kunnen namelijk uiteenlopen voor verschillende vakken (vgl. Korobko, 2007). Leithwood en Levin (2005) benadrukken eveneens dat verschillende uitkomstmaten nodig zijn, inclusief lange termijn effecten.

7.8 Maken schoolleiders het verschil?

Een belangrijke vraag die hier nu nog beantwoord moet worden, is de vraag of schoolleiders ‘het verschil maken’. Op basis van de resultaten van dit onderzoek moet het antwoord zijn: “Nee, in het voorgezet onderwijs in Nederland hebben schoolleiders geen invloed op de *leerprestaties* van leerlingen.” Deze resultaten zijn in lijn met de uitkomsten van de kwantitatieve meta-analyse van Witziers, Bosker en Krüger (2003), waaruit bleek dat er in Nederland geen ‘directe’ effecten van schoolleiderschap konden worden aangetoond. Direct-effect-modellen van schoolleiderschap leveren in principe correcte informatie op over de ‘totale’ effecten, dat wil zeggen, als er voldoende rekening wordt gehouden met verschillen tussen scholen in de samenstelling van de leerlingenpopulatie. Niettemin bleek bij de toetsing van het ‘directe en indirecte’ effect-model in het hier gerapporteerde onderzoek (zie figuur 7.1), dat er wel een klein, maar significant, positief, indirect effect is van schoolleidersgedrag op het gemiddelde *doorstroompercentage* (rendement) in de bovenbouw. Andere schoolkenmerken, namelijk de ‘ontwikkelingsgerichtheid’ van een school en vooral het werkklimaat in de klassen, leveren eveneens een bijdrage aan het rendement, terwijl het werkklimaat ook nog een klein significant positief effect heeft op de leerprestaties. Tussen deze schoolkenmerken en schoolleidersgedrag bestaan relaties, maar het is niet met zekerheid te zeggen wat hierbij ‘oorzaak’ is en wat ‘gevolg’. Waarschijnlijk oefenen de schoolorganisatie en de schoolleider wederzijds invloed uit op elkaar. Om zekerheid te krijgen over causale verbanden, is experimenteel of longitudinaal onderzoek nodig. Uit het model voor schoolleidersgedrag dat in dit proefschrift wordt gepresenteerd, is op te maken dat er behalve een indirecte relatie tussen schoolleidersgedrag en de uitkomsten, ook een (groter) negatief direct verband bestaat tussen schoolleidersgedrag en de leerprestaties. Deze significante negatieve relatie tussen schoolleiderschap en de leerprestaties moet waarschijnlijk worden toegeschreven aan een reactie van schoolleiders op de schoolprestaties en niet aan een negatieve impact van schoolleiders op de resultaten. Verondersteld wordt, dat dit verband aangeeft dat schoolleiders hun gedrag afstemmen op de prestaties van de school. Schoolleiders van scholen met lagere leerprestaties ontwikkelen meer activiteiten om de school te verbeteren, dan schoolleiders van scholen met hoge leerprestaties. Schoolleiders van scholen met hoge leerprestaties richten zich vooral op het bewaken van de stabiliteit en continuïteit binnen hun school.

Tenslotte blijft de vraag over of schoolleiders in het voortgezet onderwijs wel het verschil zouden *kunnen* maken. Uit een overzichtsstudie van schooleffectiviteitsonderzoek door Marzano (2000) blijkt dat ongeveer 80 procent van de variantie in leerlingenprestaties wordt verklaard door achtergrondkenmerken van de leerling en van de resterende variantie bijna zeven procent door factoren op schoolniveau en dertien procent door factoren op het niveau van de klas of de docent (vgl. Luyten, 1994; Scheerens & Bosker, 1997). Scheerens en Bosker (1997, p. 79) laten zien, dat een dergelijk klein schooleffect tot gevolg kan hebben, dat leerlingen met vergelijkbare kenmerken na de basisschool op verschillende niveaus instromen in het vervolgonderwijs (bijvoorbeeld havo vs. vmbo). De effecten van het gedrag van de

schoonleider vormen over het algemeen echter niet meer dan een kwart van het schooleffect en in het voortgezet onderwijs in Nederland waarschijnlijk nog minder (vgl. Witziers et al., 2003). Het is daarom te verwachten dat het indirecte effect van schoolleiders op de leerresultaten te klein is om meetbaar (significant) te zijn. Dit is waarschijnlijk mede een gevolg van het feit dat de verschillen tussen schoolleiders in het voortgezet onderwijs in Nederland beperkt zijn. De beroepsgroep van schoolleiders in het voortgezet onderwijs in Nederland is overzichtelijk (750 directeuren in 2006), er zijn vele opleidingsmogelijkheden (EIM, 2007) en de 'Basiscompetenties schoolleider VO' zijn vastgelegd (ISISQ5/VO-raad, 2007). Bovendien zullen de effecten van het gedrag van schoolleiders in het voortgezet onderwijs ook nog worden 'gebufferd' door de (grote) schoolorganisaties, vanwege de verdeling van managementtaken en de autonomie van docenten. Over het algemeen zijn de effecten van factoren op klassen- of docentenniveau (bijvoorbeeld 'gelegenheid om te leren' en instructiestrategieën) op de leerprestaties hoger dan de effecten van factoren op schoolniveau (Marzano, 2000; Scheerens & Bosker, 1997). Hieruit zou kunnen worden geconcludeerd dat het zinvol zou zijn om juist deze effecten op klassenniveau te bevorderen (vgl. Rowe, 2003).

Hoewel schoolleiders in het voortgezet onderwijs in Nederland geen (meetbare) invloed hebben op de leerprestaties, is er wel een verband tussen het gedrag van de schoolleider en het rendement van de school. Uit het indirecte-effect-model dat in dit proefschrift is beschreven, blijkt dat de schoolleider, indirect, het werkklimaat voor de leerlingen kan bevorderen, als de schoolorganisatie voldoende ontwikkelingsgericht is. Dat wil zeggen, wanneer een zorgvuldige besluitvorming, betrokkenheid van docenten, samenwerking, professionele ontwikkeling en vernieuwing centraal staan in de schoolorganisatie. Een positief werkklimaat voor leerlingen leidt tot een beter doorstroompercentage en ook tot een (bescheiden) verbetering van de leerprestaties. Dit laatste effect is echter zo klein dat een eventuele bijdrage van schoolleidersgedrag, via de ontwikkelingsgerichtheid van de school, niet meer aantoonbaar is. Om het indirecte verband tussen schoolleidersgedrag en de schoolprestaties nauwkeuriger in kaart te kunnen brengen, dienen de kenmerken van ontwikkelingsgerichte schoolorganisaties nader onderzocht te worden. Om meer inzicht te krijgen in de relaties tussen schoolleidersgedrag, de ontwikkelingsgerichtheid van de school, het werkklimaat in de klassen en de uitkomsten zouden longitudinale gegevens bestudeerd moeten worden. In de praktijk is het echter moeilijk om deze gegevens te verkrijgen, vanwege de looptijd van dergelijk onderzoek en de belasting die het oplevert voor scholen. Standaardisatie van onderzoeksinstrumenten voor het meten van schoolleiderschap en andere schoolvariabelen, bijvoorbeeld aan de hand van een generiek model voor organisatie-effectiviteit zoals het concurrerende waarden model, zou kunnen leiden tot beter vergelijkbare uitkomsten van het onderzoek naar de effecten van schoolleiderschap op de schoolprestaties.

Literatuur

- Barker, B. (2007). The leadership paradox: Can school leaders transform student outcomes? *School effectiveness and School Improvement*, 18(1), 21-43.
- Blake, R.R., & Mouton, J.S. (1964). *The managerial grid*. Houston, Texas: Gulf Publications Company.
- Blake, R.R., Mouton, J.S., & Williams, M.S. (1981). *The academic managerial grid*. San Francisco: Jossey-Bass Publishers.
- Blase, J.J. (1987). Dimensions of effective school leadership: the teachers' perspective. *American Educational Research Journal*, 24(4), 589-610.
- Blom, S., Krüger, M.L. & Roozendaal, T.van (1990). Onderwijskundig leiderschap en cultuurmanagement. *Mesomagazine*, 11(53), 9-15.
- Boerman, P.L.J. (1998). *Decentrale besluitvorming en organisatie-effectiviteit: een organisatiekundige analyse van instellingen voor middelbaar beroepsonderwijs* (proefschrift). Amsterdam: Universiteit van Amsterdam.
- Bolman, L.G., & Deal, T.E. (1991). Leadership and management effectiveness: A multiframe, multi-sector analysis. *Human Resource Management*, 30, 509-534.
- Bosker, R.J. (1990). *Extra kansen dankzij de school? Het differentieel effect van schoolkenmerken op loopbanen in het voortgezet onderwijs voor lager versus hoger milieu leerlingen en jongens versus meisjes* (proefschrift). Nijmegen: ITS.
- Bosker, R.J. (1992). *Schoolgrootte, effectiviteit en de basisvorming*. Enschede: Universiteit Twente, OCTO.
- Bosker, R.J., Béguin, A., & Rekers-Mombarg, L. (2001). Hoe meten we de prestatie van een school? In A.B. Dijkstra, S. Karsten, R. Veenstra, & A. Visscher (Eds.), *Het oog der natie. Scholen op rapport. Standaarden voor de publicatie van schoolprestaties* (pp. 121-135). Assen: Van Gorcum.
- Bosker, R.J. & Hendriks, M.A. (1997). *Betrouwbaarheid, validiteit en bruikbaarheid van een instrumentarium ten behoeve van schoolzelfevaluatie*. Enschede: OCTO.
- Bosker, R.J., Vos, H. de, & Witziers, B. (2000). *Theories and models of educational effectiveness*. Enschede: Twente University Press.
- Bosker, R.J. & Witziers, B. (1996). The true size of school effects. Paper gepresenteerd op de Annual Meeting of the American Educational Research Association, New York.
- Bosker, R. J., Vos, H. J., Witziers, B., & Scheerens, J. (2000). Theories and models of educational effectiveness. Paper gepresenteerd op de Annual Meeting of the American Educational Research Association, New Orleans.
- Bossert, S.T., Dwyer, D.C., Rowan, R., & Lee, G.V. (1982). The instructional management role of the principal. *Educational Administration Quarterly*, 18 (3), 34-64.
- Brandsma, H. P. (1988). Onderwijskundig schoolleiderschap in het voortgezet onderwijs; een vergelijking tussen schoolleiderschap op scholen voor mavo en lbo. *Tijdschrift voor Onderwijswetenschappen*, 18, 271-283.
- Brandsma, H. P. (1993). *Basisschoolkenmerken en de kwaliteit van het onderwijs*. Academisch proefschrift. Groningen: RION.

- Brandsma, H.P., & Knuver, J.W.M. (1989). Effects of school and classroom characteristics on pupil progress in language and arithmetic. *International Journal of Educational Research*, 13, 777-788.
- Brandsma, H.P. & Stoel, W.G.R. (1987). Schoolleiders in het voortgezet onderwijs: een exploratieve studie naar de determinanten en effecten van schoolleiderskenmerken. In J. Scheerens & W.G.R. Stoel (Eds.), *Effectiviteit van onderwijsorganisaties. Bijdragen aan de onderwijsresearch no. 19* (pp. 99-113). Lisse: Swets & Zeitlinger.
- Bredschneider, C. (1993). A study of student participation in and identification with one Ontario secondary school. Toronto: OISE/UT, unpublished MA paper.
- Brookover, W., Beady, C., Flood, P., Schweitzer, J., & Wisenbaker, J. (1979). *School social systems and student achievement: Schools can make a difference*. New York: Praeger.
- Bryk, A.S., & Raudenbush, S.W. (1992). *Hierarchical linear models*. New York: Sage.
- Cameron, K.S. (1985). *Cultural Congruence Strength and Type: Relationships to Effectiveness*. Working paper, Business School, University of Michigan.
- Cameron, K.S., & Quinn, R.E. (1999). *Diagnosing and changing organizational culture*. Reading, MA: Addison-Wesley.
- Campbell, D.T., & Fiske, D. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Cavanagh, R.F., & Dellar, G.B. (1998). *The Development, Maintenance and Transformation of School Culture*. Paper gepresenteerd op de Annual Meeting of the American Educational Research Association, San Diego.
- CBS (2006). *Demografische kerncijfers per gemeente 2006*. Voorburg/Heerlen: Centraal Bureau voor de Statistiek.
- Chen, F., Bollen, K.A., Paxton, P., Curran, P.J., & Kirby, J.B. (2001). Improper solutions in structural equation models. Causes, consequences, and strategies. *Sociological methods & research*, 29(4), 468-508.
- Coleman, J.S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., & York, R. (1966). *Equality of educational opportunity*. Washington, D.C.: Government Printing Office.
- Damme, J. Van, De Fraine, B., Landeghem, G. Van., Opdenakker, M.-C, & Onghena, P. (2002). A new study on educational effectiveness in secondary schools in Flanders: An introduction. *School Effectiveness and School Improvement*, 13(4), 383-397.
- Davies, P., & Coates, G. (2005). Competing conceptions and values in school strategy. Rational planning and beyond. *Educational Management Administration & Leadership*, 33(1), 109-124.
- Deinum, J.F. (2000). *Schoolbeleid, instructie en leerresultaten* (proefschrift). Groningen: GION.
- Denison, D.R., Hooijberg, R., & Quinn, R.E. (1995). Paradox and performance: Toward a theory of behavioral complexity in managerial leadership. *Organization Science* 6(5), 524-540.
- Denison, D.R., & Mishra, A.K. (1995). Toward a Theory of Organizational Culture and Effectiveness. *Organization Science*, 6(2), 204-223.
- Driesen, G. (2002). Sociaal-etnische schoolcompositie en onderwijsresultaten. Effecten van posities, concentratie en diversiteit. *Pedagogische Studiën*, 79, 212-230.
- Driessen, G., Doesborgh, J., Ledoux, G., Veen, I. van der, & Vergeer, M. (2002). *Sociale integratie in het primair onderwijs. Een studie naar de relatie tussen de sociale, etnische, religieuze en cognitieve schoolcompositie en de cognitieve en niet-cognitieve poistie van verschillende groepen leerlingen. Analyses bij het PRIMA-cohortonderzoek, derde meting*. Nijmegen/Amsterdam: ITS/SCO-Kohnstamm Instituut.

- Driessen, G., & Slik, F. van der (2001). Religion, denomination, and education in the Netherlands: Cognitive and noncognitive outcomes after an era of secularization. *Journal for the Scientific Study of Religion*, 40, 561-572.
- Dronkers, J. (2004). Do public and religious schools really differ? Assessing the European evidence. In P.J. Wolf & S. Macedo (Eds.), *Educating citizens. International perspectives on civic values and school choice* (pp. 287-314). Washington, DC: Brookings Institution Press.
- Dukelow, G.A. (1993). A statistical analysis of educational variables that influence high school students' grades and participation. University of Victoria: unpublished masters thesis.
- Dwyer, D.C., Barnett, B.G., & Lee, G.V. (1987). The school principal: scapegoat or the last great hope? In *Leadership, examining the elusive* (pp. 30-46). ASCD Yearbook.
- Dwyer, D.C., Lee, G.V., Barnett, N.N., Filby, B., Rowan, B., Alpert, B. & Kojimoto, C. (1985). *Understanding the principal's contribution to instruction: Seven principals, seven stories* (Vols. 1-7 and Methodological Appendix). San Francisco, CA: Far West Laboratory for Educational Research and Development.
- Edmonds, R. (1979). Effective Schools for the urban Poor. *Educational Leadership*, vol. 37, pag. 15-24.
- Edwards, J.L., Green, K.E., & Lyons, Ch.A. (1996). *Factor and Rasch Analysis of the School Culture Survey*. Paper gepresenteerd op de Annual Meeting of the American Educational Research Association, New York.
- Eim (Panteia) (2007). *Schoolleiders in Nederland. Landendocument voor de OESO*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap.
- Engels, N., Aelterman, A., Schepens, A., & Petegem, K. Van (2003). Het welbevinden van leerlingen in het secundair onderwijs in Vlaanderen. *Pedagogische Studiën*, 80, 192-209.
- Enz, C.A. (1986). *Power and Shared Values in the Corporate Culture*. Ann Arbor, MI: UMI.
- Feinstein, L. (200). *The relative importance of academic, psychological and behavioural attributes developed in childhood*. London, Centre for Economic Performance: London School of Economics and Political Science.
- Fiedler, F.E. (1967). *A theory of leadership effectiveness*. New York, McGraw-Hill, Inc.
- Field, A. (2000). *Discovering statistics using SPSS for Windows. Advanced techniques for the beginner*. London: Sage Publications.
- Finn, J.D. (1989). Withdrawing from school. *Review of Educational Research*, 59(2), 117-143.
- Finn, J.D., & Cox, D. (1992). Participation and withdrawal among fourth-grade pupils. *American Educational Research Journal*, 29(1), 141-162.
- Firestone, W., & Riehl, C. (2005). *A new agenda for research in educational leadership*. New York: Teachers College Press.
- Fisher, D.L., & Fraser, B.J. (1990). *Validity and use of the school-level environment questionnaire*. Paper gepresenteerd op de Annual Meeting of the American Educational Research Association, Boston.
- Fitz-Gibbon, C. and Kochan, S. (2000) School effectiveness and education indicators. In C. Teddlie & D. Reynolds (Eds.), *The international handbook of school effectiveness research* (pp. 257-282). London/New York: Falmer Press.
- Garson, G.D. (z.j.). *Statnotes: Topics in multivariate analysis. a. Factor Analysis; b. Structural Equation Modeling; c. Reliability Analysis; d. Validity*. Ontleend aan: <http://faculty.chass.ncsu.edu/garson/PA765/statnote.htm> (a: 30-05-2006; b: 09-05-2006 + update 24-4-2008; c: 29-11-2006; d. 29-11-2006).
- Gaziel, H.H. (1997). Impact of school culture on effectiveness of secondary schools with disadvantaged students. *Journal of educational research*, 90(5), 310-318.

- Geijsel, F., Slegers, P., & Berg, R. van den (1999). Transformational leadership and the implementation of large-scale innovation programs. *Journal of Educational Administration*, 37(4), 309-328.
- Glaser, S.R., & Zamanou, S. (1987). Measuring and interpreting organisational culture. *Management Communication Quarterly*, 1(2), 173-198.
- Griffith, J. (2003). Schools as organizational models: implications for examining school effectiveness. *The Elementary School Journal*, 104(1), 29-47.
- Grift, W. van de (1985a). *Het meten van onderwijskundig leiderschap*. Amsterdam: SCO.
- Grift, W. van de (1985b). *Vragenlijst onderwijskundig leiderschap (OL)*. Amsterdam: SCO.
- Grift, W. van de (1986). Leerkracht percepties van onderwijskundig leiderschap. In F.J. van der Krogt (Ed.), *Schoolleiding en management, Bijdragen aan de onderwijsresearch no. 6* (pp. 33-42). Lisse: Swets & Zeitlinger.
- Grift, W. van de (1990). Educational Leadership and academic achievement in elementary education. *School effectiveness and School Improvement*, 1(3), 26-40.
- Grift, W. van de, & Akkermans, W. (1991). Onderwijskundig leiderschap en leerlingprestaties in het basisonderwijs. *Tijdschrift voor Onderwijsresearch*, 16(4), 244-257.
- Grisay, A. (1996). *Evolution des acquis cognitifs et socio-cognitifs des élèves au cours des années de collège* [Evolution of cognitive and social-affective achievement of students during secondary education]. Liège: Université de Liège.
- Hallinger, Ph. (1983). *Assessing the instructional management behavior of principals*. Unpublished Doctoral Dissertation, Stanford University, Stanford, CA.
- Hallinger, Ph. (1989). What makes a difference? School context, principal leadership and student achievement. *Elementary School Journal*, 96(5), 527-549.
- Hallinger, P. (1994). *A resource manual for the principal instructional management rating scale*. Nashville, TN: Vanderbilt University, Center for Advanced Study of Educational Leadership.
- Hallinger, Ph. (2003). Leading educational change. Reflections on the practice of instructional and educational leadership. *Cambridge Journal of Education*, 33(3), 329-351.
- Hallinger, Ph., Bickman, L., & Davis, K. (1990). *Modeling the effects of principal leadership on student achievement*. Paper gepresenteerd op het International Congress for School Effectiveness, Jerusalem.
- Hallinger, P., Bickman, L., & Davis, K. (1996). School context, principal leadership and student reading achievement. *Elementary School Journal*, 96(5), 527-549.
- Hallinger, Ph. & Heck, R.H. (1996a). Reassessing the principals role in school effectiveness: A review of empirical research, 1980-1995. *Educational Administration Quarterly*, 32, 5-44.
- Hallinger, Ph. & Heck, R.H. (1996b). The principal's role in school effectiveness: an assessment of methodological progress, 1980-1995. In K. Leithwood et al. (Eds.) *International Handbook of Educational Leadership and Administration*. Dordrecht: Kluwer Academic Publishers.
- Hallinger, Ph. & Heck, R.H. (1998). Exploring the principal's contribution to school effectiveness: 1980 – 1995. *School Effectiveness and School Improvement*, 9(2), 157-191.
- Hallinger, Ph., Leithwood, K. & Murphy, J. (Eds.) (1993). *Cognitive perspectives on educational leadership*. New York: Teachers College Press.
- Hallinger, Ph., & Murphy, J. (1985). Assessing the instructional management behavior of principals. *The Elementary School Journal*, 86, 2, 217-247.
- Heck, R.H., & Hallinger, Ph. (2000). Next generation methods for the study of leadership and school improvement. *Educational Administration Abstracts*, 35(1), 141-162.

- Heck, R.H., & Hallinger, Ph. (2005). The study of educational leadership and management. Where does the field stand today? *Educational Management Administration & Leadership* 33(2) 229-244.
- Heck, R.H., Larsen, T.J., & Marcoulides, G.A. (1990). Instructional leadership and school achievement: Validation of a causal model. *Educational Administration Quarterly*, 26, 94-125.
- Heck, R.H., & Marcoulides, G.A. (1996). School culture and Performance: Testing the invariance of an organizational model. *School Effectiveness and School Improvement*, 7(1), 76-95.
- Heck, R.H., Marcoulides, G.A., & Lang, P. (1991). Principal Instructional Leadership and School Achievement: The Application of Discriminant Techniques. *School Effectiveness and School Improvement*, 2, 115-135.
- Heck, R.H., & Thomas, S.L. (2000). *An introduction to multilevel modeling techniques*. Mahwah, N.J./London: Lawrence Erlbaum Associates.
- Heiser, D.A. (z.j.). *Note AB: Likert scale variables* (excerpts from messages on SEMNET (2005-2006). Ontleend aan: <http://www.daheiser.info/> (12-01-2006).
- 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.
- Hill, P.W., Rowe, K. J., & Holmes-Smith, P. (1995). *Factors affecting students' educational progress: Multilevel modeling of educational effectiveness*. Paper gepresenteerd op de annual meeting of the ICSEI, Leeuwarden.
- Hooijberg, R., & Choi, J. (2000). Which leadership roles matters to whom? An examination of rater effects on perceptions of effectiveness. *Leadership Quarterly*, 11(3), 341-364.
- Hooijberg, R., & Choi, J. (2001). The impact of organizational characteristics on leadership effectiveness models. An examination of leadership in a private and a public sector organization. *Administration Society*, 33(4), 403-431.
- Houtveen, Th., Vermeulen, C., & van de Grift, W. (1993). *Bouwstenen voor onderzoek naar de kwaliteit van scholen*. Utrecht: ISOR.
- Houtveen, A.A.M., Voogt, J.C., Vegt, A.L. van der, Grift, W.J.C.M. van de (1995). *Zo zijn onze manieren. Onderzoek naar de organisatiecultuur van scholen*. Utrecht: ISOR.
- Hoy, W.K., & Feldman, J.A. (1987). Organizational Health: The concept and its measure. *Journal of Research and Development in Education*, 20, 30-38.
- Hoy, W.K., Tarter, C.J., & Kottkamp, R.B. (1991). *Open schools / Healthy schools. Measuring organizational climate*. Newbury Park, CA: Sage.
- Hox, J.J. (2002). *Multilevel analysis: techniques and applications*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Hox, J.J., & Maas, J.M. (2001). The accuracy of multilevel structural equation modeling with pseudobalanced groups and small samples. *Structural equation modeling*, 8(2), 157-174.
- Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Imants, J. (1996). *Leiding geven aan onderwijs! Onderwijskundig leiderschap en taakdifferentiatie in basisscholen*. DSWO Press, Rijksuniversiteit Leiden.
- Immegart, G.L. (1988). Leadership and leader behaviour. In N.Boyan (Ed.), *Handbook of Research on Educational Administration* (pp 259-277). New York: Longman.
- ISISQ5/VO-raad (2007). *Basiscompetenties schoolleiders VO*. Utrecht: ISISQ5/VO-raad.
- Jencks, C.S., Smith, M., Ackland, H., Bane, M.J., Cohen, D., Gintis, H., Heyns, B., & Michelson, S. (1972). *Inequality: a reassessment of the effect of family and schooling in America*. New York: Basic.

- Jong, A.M. de, & van Doorne-Huiskes (1992). Leiderschapsstijl: verschillen tussen mannen en vrouwen? In M.I. Demenint & C.E. Disselen (Eds.), *Vrouwen, leiderschap en management* (pp. 23-36). Utrecht, Lemma BV.
- Jöreskog, K.G. (1999). *How large can a standardized coefficient be?* Versie June, 22, 1999. Lincolnwood, IL: Scientific Software International. Ontleend aan: <http://www.ssicentral.com/lisrel/techdocs/HowLargeCanaStandardizedCoefficientbe.pdf>
- Jöreskog, K.G. (2000). *Latent variable scores and their uses.* Versie July 10, 2000. Lincolnwood, IL: Scientific Software International. Ontleend aan: <http://www.ssicentral.com/lisrel/techdocs/lvscores.pdf> .
- Jöreskog, K.G. (2005). *Structural equation modeling with ordinal variables using LISREL.* Versie 26 April 2002, revised 10 February 2005. Lincolnwood, IL: Scientific Software International. Ontleend aan: <http://www.ssicentral.com/lisrel/techdocs/ordinal.pdf> .
- Jöreskog, K.G., & Sörbom, D. (1993a). *LISREL 8: Structural equation modeling with the SIMPLIS command language.* Chicago: Scientific Software International.
- Jöreskog, K.G., & Sörbom, D. (1993b). *LISREL 8: User's reference guide.* Chicago: Scientific Software International.
- Kalliath, Th. J., Bluedorn, A.C., & Gillespie, D.F. (1999). A confirmatory factor analysis of the competing values framework. *Educational and Psychological Measurement, 59*(1), 143-158.
- Kaplan, D., & Elliott, P.R. (1997). A model-based approach to validating education indicators using multilevel structural equation modelling. *Journal of Educational and Behavioural Studies, 22*(3), 323-347.
- Kim, H., & Yukl, G. (1995). Relationships of managerial effectiveness and advancement to self-reported and subordinate-reported leadership behaviors from the multiple linkage model. *Leadership Quarterly, 6*(3), 361-377.
- Kelloway, E.K. (1998). *Using LISREL for structural equation modelling: A researcher's guide.* Thousand Oaks, CA: Sage Publications, Inc.
- Kenny, D.A. (1998). Multiple factor models. Versie November 9, 1998. Ontleend aan: <http://davidakenny.net/cm/mfactor.htm> .
- Kerr, St. & Jermier, J. (1978). Substitutes for leadership: their meaning and measurement. *Organizational behaviour and Human Performance, 22*, 375-403.
- Kleintjes, F., & Kremers, E. (1992). School size, characteristics and outcomes in Dutch secondary education. In Tj. Plomp, J.M. Pieters & A. Feteris (Eds.), *Book of summaries ECER 1992* (volume I, pp. 25-28). Enschede: Universiteit Twente, Afdeling Toegepaste Onderwijskunde.
- Knuver, J.W.M. (1993). *De relatie tussen klas- en schoolkenmerken en het affectief functioneren van leerlingen* (proefschrift). Groningen: RION.
- Knuver, A.W.M., & Brandsma, H.P. (1993). Cognitive and affective outcomes in school effectiveness research. *School Effectiveness and School Improvement, 4*, 189-204.
- Korobko, O.B. (2007). *Comparison of examination grades using item response theory: a case study* (proefschrift). Enschede: Universiteit Twente.
- Krüger, M.L. (1994). *Sekseverschillen in schoolleiderschap.* Academisch proefschrift, UvA. Alphen aan den Rijn: Samsom H.D. Tjeenk Willink.
- Krüger, M.L., & Witziers, B. (2003). Ontwikkelingen in het denken over leiderschap. In B. Creemers (Ed.), *Handboek Schoolorganisatie en Onderwijsmanagement* (pp. E3000-1-E3000-28). Alphen aan den Rijn: Samsom.
- Krüger, M.L., Witziers, B., & Slegers, P.J.C. (2007). The impact of school leadership on school level factors: Validation of a causal model. *School Effectiveness and School Improvement, 18*(1), 1-20.

- Krüger, M.L., Witziers, B., Slegers, P.J.C., Imants J. (1999). Onderwijskundig leiderschap in moderne onderwijsinstellingen. In B.P.M Creemers, J.H.G.I. Giesbers, M.L Krüger & C.A. van Vilsteren (Eds.). *Handboek Schoolorganisatie en Onderwijsmanagement: Leiding geven in bestel, school en klas* (pp. C4120.1-C4120.26). Alphen aan den Rijn: Samsom.
- Lamond, D. (2001). The value of Quinn's competing values model in an Australian context. *Journal of Managerial Psychology*, 18(1), 46-59.
- Landeghem, G. Van., Damme, J. Van, Opdenakker, M.-C., De Fraine, B., & Onghena, P. (2002). The effect of schools and classes on noncognitive outcomes. *School Effectiveness and School Improvement*, 13(4), 429-451.
- Lawrence, K.A., Quinn, R.A., & Lenk, P. (2003). *Behavioural complexity in leadership: the psychometric properties of a new instrument*. Unpublished working paper, Ann Arbor, MI: University of Michigan Business School. Ontleend aan: <http://webuser.bus.umich.edu/plenk/Behavioral%20Complexity%20in%20Leadership.pdf>.
- Ledoux, G., Veen, I. van der, Vergeer, M, Driessen, G., & Doesborgh, J. (2003). *Sociale integratie in het primair onderwijs. Resultaten van een onderzoek naar de invloed van het leerlingenpubliek van scholen op prestaties en welbevinden van leerlingen*. Nijmegen/Amsterdam: ITS/SCO-Kohnstamm Instituut.
- Leithwood, K. (1994). Leadership for school restructuring. *Educational Administration Quarterly*, 30(4), 498-518.
- Leithwood, K. (1995). Cognitive perspectives on school leaders. *Journal of School Leadership*, 5, 155-135.
- Leithwood, K., Aitken, R., & Jantzi, D. (2000). *Making schools smarter. A system for monitoring school and district progress* (2nd ed). Thousand Oaks, CA: Corwin Press Inc.
- Leithwood, K., Begley, P.T. & Cousins, J.B. (1990). The nature, causes and consequences of principals' practice: an agenda for future research. *Journal of Educational Administration*, 28, 4, 5-30.
- Leithwood, K., Dart, B., Jantzi, D. and Steinbach, R. (1993). *Fostering Organisational Learning: a Study of British Columbia's Intermediate Sites 1990-1992* (Final report). Victoria, BC: British Columbia Ministry of Education.
- Leithwood, K., Day, C., Sammons, P., Harris, A., & Hopkins, D. (2006a). *Successful school leadership: what it is and how it influences student learning*. London: DfES, Research Report 800.
- Leithwood, K., Day, C., Sammons, P., Harris, A., & Hopkins, D. (2006b). *Seven strong claims about successful school leadership*. Nottingham: DfES/NCSL.
- Leithwood, K., & Duke, D. (1999). A century's quest to understand school leadership. In J. Murphy & K. Louis (Eds.). *Handbook of Research on Educational Administration* (2nd ed. pp 45-72). San Francisco: Jossey-Bass.
- Leithwood, K., & Jantzi, D. (1999a). The relative effects of principal and teacher sources of leadership on student engagement with school. *Educational Administration Quarterly*, 35(Supplemental), 679-706.
- Leithwood, K., & Jantzi, D. (1999b). Transformational school leadership effects: A replication. *School Effectiveness and School Improvement*, 10(4), 451-479.
- Leithwood, K., & Jantzi, D. (2000a). The effects of transformational leadership on organizational conditions and student engagement with school. *Journal of Educational Administration*, 38(2), 112-129.
- Leithwood, K., & Jantzi, D. (2000b). Principal and teacher leadership effects: A replication. *School leadership & Management*, 20(4), 415-434.

- Leithwood, K., & Jantzi, D. (2006). Transformational school leadership for large-scale reform: Effects on students, teachers, and their classroom practices. *School Effectiveness and School Improvement*, 17(2), 201-227.
- Leithwood, K., Jantzi, D., & Steinbach (1999). *Changing leadership for changing times*. Philadelphia: Open university press.
- Leithwood, K., & Levin, B. (2005). *Assessing school leader and leadership programme effects on pupil learning*. London: DfES, Research Report 662.
- Leithwood, K., & Mascall, B. (2008). Collective Leadership effects on student achievement. *Educational Administration Quarterly*, 44(4), 529-561.
- Leithwood, K., & Riehl (2003a). *What do we already know about successful school leadership?* Paper prepared for the AERA Division A Task Force on Developing Research in Educational Leadership.
- Leithwood, K., & Riehl (2003b). *What we know about successful school leadership*. A report by Division A of AERA, Autumn 2003.
- Leithwood, K., & Riehl (2005). What we know about successful school leadership. In W. Firestone & C. Riehl (Eds.), *A new agenda: directions for research on educational leadership* (pp.22-47). New York: Teachers College Press.
- Leithwood, K., Seashore Louis, K., Anderson, S., & Wahlstrom, K. (2004). *Review of research: How leadership influences learning*. New York: The Wallace Foundation.
- Leithwood, K., Tomlinson, D., & Genge, M. (1996). Transformational school leadership. In K. Leithwood, J. Chapman, D. Corson, P. Hallinger & H. Hart (Eds.), *International handbook of educational leadership and administration* (pp. 785-840). Dordrecht/Boston/London: Kluwer Academic Publishers.
- Lexicon Methoden en Technieken, Universiteit Leiden (z.j.). Ontleend aan: <http://www.ment.psychologie.leidenuniv.nl/lexicon/> (22-12-2006).
- Linden, R. van der, Teurlings, C., & Vermeulen, M. (2003). De school als professionele organisatie. Opbrengsten van vijf jaar kortlopend Onderwijsonderzoek (1998-2002). Tilburg: IVA.
- Lugthart, E., Roeders, P., Bosker, R., & Bos, K. (1989). *Effectieve schoolkenmerken in het voortgezet onderwijs: Deel I: een literatuuroverzicht*. Groningen: Instituut voor Onderwijsonderzoek (RION).
- Luyten, J.W. (1994). *School effects: stability and malleability* (proefschrift). Enschede: Universiteit Twente.
- Luyten, J.W., Scheerens, J., Visscher, A.J., Maslowski, R., Witziers, B., & Steen, R. (2005). *School factors related to quality and equity*. Results from Pisa 2000. Paris: OECD.
- MacCallum, R., Roznowski, M., & Necowitz, L.B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, 111(3), 490-504.
- Maeyer, S. de, & Rymenans, R. (2004). *Onderzoek naar kenmerken van effectieve scholen. Kritische factoren in een onderzoek naar schooleffectiviteit in het technisch en beroepssecundair onderwijs in Vlaanderen* (proefschrift). Gent: Academia Press.
- Maeyer, S. de, Rymenans, R., Petegem, P. Van, Bergh, H. van den, & Rijlaarsdam, G. (2007). Educational Leadership and pupil achievement: The choice of a valid conceptual model to test effects in school effectiveness research. *School effectiveness and School Improvement*, 18(2), 125-145.
- Marks, H.M., Louis & Printy, S. (2003). The capacity for organizational learning. In Leithwood, K. (Ed.), *Understanding schools as intelligent systems* (pp.239-265). Stamford, CT: JAI Press.

- Marks, H.M., & Printy, S. (2003). Principal leadership and school performance. An integration of transformational and instructional leadership. *Educational Administration Quarterly*, 29(3), 370-397.
- Marsh, D.D. (1997). Educational leadership for the 21st century: integrating three emerging perspectives. Paper gepresenteerd op de Annual Meeting of the American Educational Research Association, Chicago.
- Martin, M.O., Mullis, I.V.S., Gregory, K.D., Hoyle, C., & Shen, C. (2000). *Effective school in science and mathematics. IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: International Study Centre.
- Marwijk Kooy-von Baumhauer, L. van (1984). *Scholen verschillen* (proefschrift). Groningen.
- Marx, E.C.H. (1986). Schoolorganisatorische aspecten van waardengerichtheid in het onderwijs bij uiteenlopende schoolculturen. In B.P.M. Creemers, J. Giesbers, C.A. Van Vilsteren, & C. van der Perre (Eds.), *Handboek Schoolorganisatie en onderwijsmanagement*. Alphen aan den Rijn: Samsom.
- Marzano, R.J. (2000). *A new area of school reform: going where the research takes us*. Aurora, Colorado: Mid-continent Research for Education and Learning (McREL).
- Maslowski, R. (1997). Schoolcultuur: kenmerken en veranderingsmogelijkheden. In B.P.M. Creemers e.a. (Eds.), *Handboek Schoolorganisatie en Onderwijsmanagement* (pp. B1400/1-B1400/24). Alphen aan den Rijn: Samsom Tjeenk Willink.
- Maslowksi, R. (2001). *School culture and school performance. An explorative study into the organization culture of secondary schools and their effects* (proefschrift). Enschede: Twente University Press.
- Maslowski, R. (2006). A review of inventories for diagnosing school culture. *Journal of Educational Administration*, 44(1), 6-35.
- Meelissen, M. R. M. (2005). *ICT: meer voor Wim dan voor Jet? De rol van basisonderwijs in het aantrekkelijk maken van informatie- en communicatietechnologie voor jongens en meisjes* (proefschrift). Enschede: Universiteit Twente.
- Mels, G. (2004). *LISREL for Windows: getting started guide*. Lincolnwood, IL: Scientific Software International.
- Miller, R.J., & Rowan, B. (2006). Effects of organic management on student achievement. *American Educational Research Journal*, 43(2), 219-253.
- MinOCW (2007). *Nota werken in het onderwijs 2007*. Ontleend aan: <http://www.minocw.nl/documenten/WIO%202007.pdf>.
- Mintzberg, M. (2001). *The nature of managerial work*. New York: Prentice Hall.
- Morrison, D.F. (2002). *Multivariate Statistical methods*. Pacific Grove, CA: Brooks/Cole.
- Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1988). *School matters: The junior years*. Somerset, UK: Open Books.
- Muijen, J.J. van (1994). *Organisatiecultuur en organisatieklimaat: de ontwikkeling van een meetinstrument op basis van het 'competing values' model* (proefschrift). Ridderkerk: Ridderprint.
- Muijen, J. van, Koopman, P., & Witte, K. de (1996). *Focus op organisatiecultuur: het concurrerende waarden model en het meten en veranderen van organisatiecultuur*. Schoonhoven: Academic Service.
- Mulford, B. (2003). *School leaders: changing roles and impact on teacher and school effectiveness*. OECD, Paris, paper commissioned by the Education and Training Policy Division for the activity Attracting, Developing and Retaining Effective Teachers. Ontleend aan: <http://www.oecd.org/dataoecd/61/61/2635399.pdf>.
- Mulford, B. (2006). Leading change for student achievement. *Journal of Educational Change* 7, 47-58.

- Mulford, B., & Silins, H. (2003). Leadership for organisational learning and improved student outcomes-What do we know? *Cambridge Journal of Education*, 33(2), 175-195.
- Mulford, B., Silins, H., & Leithwood, K. (2003). *Educational leadership for organisational learning and improved student outcomes*. New York (etc.): Kluwer.
- Murphy, J. (1990). Principal instructional leadership. In P.W. Thurston & L.S. Lotto (Eds.), *Advances in Educational Administration, Vol. 1: Changing perspectives on schools* (pp. 163-200). Greenwich/London: Jai Press Inc.
- Muthén, B.O. (1990). *Means and covariance structure analysis of hierarchical data*. Los Angeles: UCLA Statistics series, no 62.
- Muthén, B.O. (1994). Multilevel covariance structure analysis. *Sociological Methods & Research*, 22, 376-398.
- Muthén, B.O. (1997). Latent variable modeling of longitudinal and multilevel data. *Sociological Methodology*, 27, 453-480.
- Muthén, L.K., & Muthén, B.O. (1998-2001). Mplus user's guide (2nd edition). Los Angeles, CA.: Muthén & Muthén.
- Muthén, B.O., & Satorra, A. (1995). Complex sameple data in structural equation modeling. *Sociological Methodology*, 25, 267-316.
- NCSL (2007). *What we know about school leadership*. Nottingham: National College for School Leadership.
- Norušis, M.J. (2004). *SPSS 13.0 Advanced Statistical Procedures Companion*. Englewood Cliffs: Prentice Hall, 2004.
- OECD (2001). *Knowledge and skills for life: First results from PISA 2000*. Paris: OECD.
- OECD (2003). *Manual for the PISA 2000 database*. Paris: OECD.
- Ogawa, R. & Bossert, S. (1995). Leadership as an organizational quality. *Educational Administration Quarterly*, 31, 224-243.
- Opdenakker, M.-C, & Van Damme, J. (2000). Effects of schools, teaching staff and classes on achievement and well-being in secondary education: similarities and differences between school outcomes. *School Effectiveness and School Improvement*, 11(2), 165-196.
- Opdenakker, M.-C, & Van Damme, J. (2001). Relationship between school composition and characteristics of school process and their effect on mathematics achievement. *British Educational Research Journal*, 27(4), 407-432.
- Opdenakker, M.-C, & Van Damme, J. (2006). Differences between secondary schools: A study about school context, Group composition, school practice, and school effects with special attention to public and Catholic schools and types of schools. *School Effectiveness and School Improvement*, 17(1), 87-117.
- Pang, N.S.K. (1996). School values and teachers' feelings: a LISREL model. *Journal of Educational Administration*, 34 (2), 64-83.
- Pang, N.S.K. (1998). *Organizational cultures of excellent schools in Hong Kong*. Paper gepresenteerd op de Annual Meeting of the American Educational Research Association, San Diego, CA.
- Payne, R., & Pugh, D.S. (1976). Organizational structure and climate. In M.D. Dunette, *Handbook of Industrial and Organizational Psychology* (pp. 1125-1174). Chicago: Rand McNally.
- Pitner, N.J. (1981). Hormones and harems: are the activiteit of superintending different for a woman? In P.A.Schmuck, W.W. Charters Jr., & Carlson, R.O. (Eds.), *Educational policy and management: Sex differentials* (pp. 221-234). New York: Academic Press.

- Pounder, D.G., Ogawa, R.T. & Adams, E.A. (1995). Leadership as an organization-wide phenomena: Its impact on school performance. *Educational Administration Quarterly*, 31(4), 564-588.
- Quinn, R.E. (1988). *Beyond rational management: mastering paradoxes and competing demands of high performance*. San Francisco, CA: Jossey-Bass.
- Quinn, R.E. (1998). *Persoonlijk meesterschap in management: voorbij rationeel management*. Schoonhoven: Academic Service
- Quinn, R.E. & Cameron, K.S. (1983). Organizational life cycles and shifting criteria of effectiveness: Some preliminary evidence. *Management Science*, 19, 33-51.
- Quinn, R., Faerman, S., & Dixit, N. (1987). Perceived Performance: Some Archetypes of Managerial Effectiveness and Ineffectiveness (working paper). Albany: Institute for Government and Policy Studies, Department of Public Administration, State University of New York.
- Quinn, R.E., Faerman, R.E., Thompson, M.P. & McGrath, M.R. (2003). *Handboek Managementvaardigheden* (3e druk). Schoonhoven: Academic Service.
- Quinn, R.E. & Rohrbaugh, J. (1983). A spatial model of effectiveness criteria: towards a competing values approach to organizational analysis. *Management Science*, (29)3, 363-377.
- Quinn, R.E., & Spreitzer, G.M. (1991). The psychometrics of the competing values culture instrument and an analysis of the impact of organizational culture on quality of life. *Research in Organizational Change and Development*, 5, 115-142.
- Raudenbush, S.W. (1994). Random effects models. In H. Cooper & L.V. Hedges (eds.), *The handbook of research synthesis* (pp. 231-323). New York: Russell Sage.
- Rekers-Mombarg, L.T.M., Lodewick, J.G.M.H., & Bosker, R.J. (2000). *Verschillen in examencijfers. De rol van initiële vaardigheden en de bijdrage van scholen: een onderzoek naar de onderwijsloopbanen van de gediplomeerde leerlingen uit het cohort VOCL '89*. Enschede: Universiteit Twente, OCTO.
- Reynolds, D., & Teddlie, C. (2000). The future agenda for school effectiveness research. In C. Teddlie & D. Reynolds (Eds.), *The international handbook of school effectiveness research* (pp. 322-343). London/New York: Falmer Press.
- Ribbins, P., & Gunter, H. (2002). Mapping leadership studies in education: Towards a typology of knowledge domains. *Educational Management & Administration* 30(4): 359-385.
- Richmon, M.J. & Allisin, D.J. (2003) Toward a Conceptual Framework for Leadership Inquiry. *Educational Management & Administration* 31(1), 31-50.
- Rigdon, E. (z.j.) *Not positive definite matrices—causes and cures*: Ontleend aan: Ed Rigdon's Structural Equation Modeling Page <http://www2.gsu.edu/~mkteer/> (06-11-2006).
- Rokeach, M. (1973). *The Nature of Human Values*. New York, NY: Free Press.
- Rowan, B., Raudenbush, S.W., & Kang, S.J. (1991). School climate in secondary schools. In S.W. Raudenbush & J.D. Willms (Eds.), *Schools, classrooms, and pupils: International studies of schooling from a multi-level perspective* (pp. 203-223). San Diego: Academic.
- Rowe, K. (1995). *Structural equation modeling. A thematic, Integrated Course*. ACSPRI winter program in social research methods and research technology. Adelaide: University of Adelaide.
- Rowe, K. (2002). *The measurement of latent and composite variables from multiple items or indicators: Applications in performance indicator systems*. Background paper to keynote address presented at the RMIT Statistics Seminar Series. Camberwell: Australian Council for Educational Research. Ontleend aan: <http://www.acer.edu.au/research/programs/learningprocess.html> .

- Rowe, K. (2003). *The importance of teacher quality as a key determinant of students' experiences and outcomes of schooling*. Background paper to keynote address presented at the ACER Research Conference. Melbourne: Australian Council for Educational Research. Ontleend aan:
http://www.acer.edu.au/documents/Rowe_ACER_Research_Conf_2003_Paper.pdf.
- Rowe, K.J., & Hill, P.W. (1998). Modeling educational effectiveness in classrooms: the use of multi-level structural equations to model students' progress. *Educational Research and Evaluation*, 4(4), 307-347.
- Rutter, M., Maugham, B., Mortimore, P., Ouston, J., & Smith, A. (1979). *Fifteen thousand hours: Secondary schools and their effects on children*. Cambridge, MA: Harvard University Press.
- Saphier, J., & King, M. (1985). Good Seeds Grow in Strong Cultures. *Educational Leadership*, 42 (6), 67-74.
- Saris, W.E., & Stronkhorst, L.H. (1984). *Causal modelling in nonexperimental research. An introduction to the LISREL Approach*. Amsterdam: Sociometric Research Foundation.
- Scheerens, J. (1989). *Wat maakt scholen effectief? Samenvatting en analyse van onderzoeksresultaten*. 's Gravenhage: Instituut voor Onderzoek van Onderwijs SVO.
- Scheerens, J. (1990). School effectiveness and the development of process indicators of school functioning. In *School effectiveness and school improvement* (pp. 61-80). Lisse: Swets & Zeitlinger.
- Scheerens, J. (1992). *Effective schooling, Research, theory and practice*. London: Cassell.
- Scheerens, J. (1993) Evaluatie. In P.J.J. Stijnen, J. Scheerens, A.L.M. Van Wieringen & H.G.W. Münstermann (Eds). *Transformatie van schoolorganisaties*. Alphen aan den Rijn/Heerlen: Samsom Tjeenk Willink/Open Universiteit Nederland.
- Scheerens, J. (1999). Concept and theories of school effectiveness. In A.J. Visscher (Ed.), *Managing schools toward high performance. Linking school management theory to the school effectiveness knowledge base* (pp. 37-70). Lisse: Swets & Zeitlinger.
- Scheerens, J. & Bosker, R. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.
- Scheerens, J. & Witziers, B. (2005). *Educational leadership and student performance*. Enschede: Universiteit Twente.
- Scheerens, J., Witziers, B., Steen, R. (nog niet gepubliceerd). A meta-analysis of school effectiveness studies.
- Schein, E.H. (1985). *Organizational culture and leadership*. Washington, London: Jossey Bass.
- Schmidt, G. (2009). (proefschrift).
- Schonrock-Adema, J. (2002). *De ontwikkeling en evaluatie van een zelfinstructieprogramma voor een training in basisgespreksvaardigheden* (proefschrift). Groningen: Rijksuniversiteit Groningen.
- Schreiber, J.B., Nora, A., Stage, F.L., Barlow, E.A., King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, 99(6), 323-337.
- Senge, P.M. (1990). *The fifth discipline, the art and practice of the learning organization*. New York: Doubledag Currency.
- Silins, H., & Mulford, B. (2002). Schools as learning organizations: The case for system, teacher and student learning. *The Journal of Educational Administration*, 40(5), 425-446.
- Silins, H., & Mulford, B. (2004). Schools as learning organizations: Effects on teacher leadership and student outcomes. *School Effectiveness and School Improvement*, 15(3-4), 443-446.

- Silins, H., & Mulford, B., & Zarins, S., (2002). Organizational Learning and School Change. *Educational Administration Quarterly*, 38(5), 613-642.
- Silins, H., Zarins, S., & Mulford, B. (2002). What characteristics and processes define a school as a learning organization? Is this a useful concept to apply to schools? *International Education Journal*, 3(1), 24-32.
- Sleegers, P.J.C. (1991). School en beleidsvoering: Een onderzoek naar de relatie tussen het beleidsvoerend vermogen van scholen en het benutten van de beleidsruimte door scholen (proefschrift). Nijmegen: Nijmegen University Press.
- Sleegers, P.J.C. (1999). *Leiding geven aan leren*. Nijmegen: Katholieke Universiteit Nijmegen.
- Sleegers, P.J.C., & Schooten, E. van (2006). *Onderwijskundige sturing en de effecten daarvan in schoolorganisaties*. Conceptversie rapportage NWO PROO. Amsterdam: SCO-Kohnstamm Instituut.
- Smart, J.C. (2003). Organizational effectiveness of 2-year colleges: The centrality of cultural and leadership complexity. *Research in Higher Education*, 44(6), 673-703.
- Snijders, T.A. & Bosker, R.J.(1999). *Multilevel analysis: an introduction to basic and advanced multilevel modeling*. London: Sage Publications.
- Snyder, K.J. (1988). *School Work Culture Profile*. Tampa, FL: School Management Institute.
- SSIcentral (z.j.). *Fitting structural equation model to complete continuous data*. Ontleend aan: <http://www.ssicentral.com/lisrel/techdocs/Session3.pdf> (23-01-2007).
- Staessens, K. (1990). *De professionele cultuur van basisscholen in vernieuwing: een empirisch onderzoek in V.L.O.-scholen* (proefschrift). Leuven: Katholieke Universiteit Leuven, Centrum voor onderwijsbeleid en vernieuwing.
- Staessens, K. (1991b). Ontwikkeling en validering van de Vragenlijst Professionele Cultuur van Basisscholen. *Pedagogische Studiën*, 68, 241-252.
- Stevens, J.P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Stoel, W.G.R. (1980). *De beleving van de school door leerlingen in het voortgezet onderwijs. De ontwikkeling van een schoolbelevingsschaal*. Haren: RION.
- Swanborn, P.G. (1994). *Methoden van sociaal-wetenschappelijk onderzoek* (Nieuwe editie). Meppel/Amsterdam: Boom.
- Sweeney, J. (1982). Research synthesis on effective school leadership. *Educational Leadership*, 39(5), 346-352.
- Tabachnick, B.G, & Fidell, L.S. (1983). *Using multivariate statistics*. New York: Harper & Row publishers.
- Teddlie, C., & Stringfield, S. (1993). *Schools make a difference. Lessons learned from a 10 year study of school effects*. New York: Teachers College Press.
- Teddlie, C., Stringfield, S. & Reynolds, D. (2000). Context issues with school effectiveness research. In C. Teddlie & D. Reynolds (Eds.), *The international handbook of school effectiveness research* (pp. 160-185). London/New York: Falmer Press.
- Teddlie, C., Reynolds, D. & Sammons, P. (2000). The methodology and scientific properties of school effectiveness research. In C. Teddlie & D. Reynolds (Eds.), *The international handbook of school effectiveness research* (pp. 160-185). London/New York: Falmer Press.
- Thompson, M.D. (2000). Gender, leadership orientation, and effectiveness: testing the theoretical models of Bolman & Deal and Quinn. *Sex Roles*, 42(11/12), 969-992.
- Toit, M. du, & Toit, S. du (2001). *Interactive LISREL: User's Guide*. Lincolnwood, IL: Scientific Software International.

- Toit, M. du, Toit, S. du, Mels, G. & Cheng, Y. (z.j.). *LISREL for Windows: PRELIS User's Guide*. Lincolnwood, IL: Scientific Software International. Ontleend aan: <http://www.ssicentral.com/lisrel/techdocs/IPUG.pdf> (12-01-2007).
- Toland, M.D., & Ayala, R.J. De (2005). A multilevel factor analysis of students' evaluations of teaching. *Educational and Psychological Measurement*, 65(2), 272-296.
- Tomarken, A.J., & Waller, N.G. (2005). Structural Equation Modelling: Strengths, limitations, and misconceptions. *Annual Review of Clinical Psychology*, 1, 31-65.
- Trochim, W.M. (2006). The multitrait multimethod matrix (versie 20-10-2006). In *The Research Methods Knowledge Base* (2nd Edition). Ontleend aan: <http://www.socialresearchmethods.net/kb/>.
- Veenstra, R. (1999). *Leerlingen, klassen, scholen. Prestaties en vorderingen van leerlingen in het voortgezet onderwijs*. Amsterdam: Thela.
- Vehkalahti, K., Puntanen, S., & Tarkkonen (2006). Estimation of reliability: a better alternative for Cronbach's alpha. Preprint submitted to Elsevier Science, 20 February 2006. Ontleend aan: <http://mathstat.helsinki.fi/reports/Preprint430.pdf>.
- Vermeulen, M. (1997). *De school als arbeidsorganisatie. Personeelsbeleid, organisatiekenmerken en arbeidsbeleving van leraren in scholen voor voortgezet onderwijs* (proefschrift). De Lier: ABC.
- Verschuren, P.J.M. (1991). *Structurele modellen tussen theorie en praktijk*. Utrecht: Het Spectrum.
- Vilsteren, C.A. van (1999). *Onderwijskundig leiderschap van schoolleiders in het voortgezet onderwijs: schoolleiding en vaksecties* (proefschrift). Enschede: Universiteit Twente.
- Visscher, A.J. (1999). Introduction to organizational and management aspects of schools. In A.J. Visscher (Ed.), *Managing schools toward high performance. Linking school management theory to the school effectiveness knowledge base* (pp. 3-36). Lisse: Swets & Zeitlinger.
- Voorbeeldfuncties FUWA-VO 2002. Ontleend aan: <http://www.besturenraad.nl/docs/fuwa-vo-2002-voorbeeldfuncties-24-03-2004.doc>.
- Walberg, H.J. (1984). Improving the productivity of America's schools. *Educational leadership*, 41(8), 19-27.
- Watson, T.J. (1986). *Management, Organisation, and Employment Strategy: New Directions in Theory and Practice*. London, Boston: Routledge, Kegan Paul International
- Weick, K.E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21, 1-19.
- Weick, K.E. (1982). Administering education in loosely coupled systems. *Phi Delta Kappan*, 63, 673-676.
- Weick, K.E. (1984). Management of educational change among loosely coupled elements. In P. Goodman (Ed.), *Change in organizations* (pp. 2-24). San Francisco: Jossey-Bass.
- Wieringen, A.M.L van, Ax, J., Karstanje, P.N., & Voogt, J.C. (2000). *Organisatie van scholen*. Leuven/Apeldoorn: Garant.
- Wilms, J.D., & Raudenbush, S.W. (1989). A longitudinal hierarchical linear model for estimating school effects and their stability. *Journal of Educational Measurement*, 26(3), 2009-232.
- Witziers, B. (1992). *Coördinatie binnen scholen voor voortgezet onderwijs*. Enschede: Universiteit Twente.
- Witziers, B. (1999). Coordination and control in education. In A.J. Visscher (Ed.), *Managing schools toward high performance. Linking school management theory to the school effectiveness knowledge base* (pp. 97-134). Lisse: Swets & Zeitlinger.

- Witziers, B., Bosker, R.J. & Krüger, M.L. (2003). Educational leadership and student achievement: The elusive search for an association. *Educational Administrative Quarterly*, 39(3), 398-425.
- Witziers, B. & de Groot, I.N. (1993). *Concurrentie tussen scholen. Een onderzoek naar de marketingmiddelen die scholen inzetten om hun concurrentiepositie te verbeteren en de wettelijke context waarbinnen scholen concurreren*. Enschede: Universiteit Twente, OCTO.
- Yu, C. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes*. A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Education. Ontleend aan: <http://www.statmodel.com/download/Yudissertation.pdf> (28-11-2006).
- Yukl, G.A. (1998). *Leadership in organizations*. Upper Saddle River, NJ: Prentice Hall.

Summary

Do school leaders make a difference with respect to school performance? Traditionally, school leadership is regarded as one of the key factors in school effectiveness (cf. Edmonds, 1979). However, little is known about the means by which school leaders achieve an impact on school performance. This study investigates the chain of effects between school leadership and student outcomes. The main research question is: “To what extent does educational leadership, directly or indirectly, influence student achievement?”

In the Netherlands, as well as in other countries, schools have been granted more autonomy in their decision making, due to a process of decentralisation. At the same time, schools are held more accountable for their performance. As a result, school leadership is increasingly defined by a demanding set of roles, including financial and human resource management and leadership for learning. These developments have made school leadership a priority in school effectiveness and school improvement research.

During the last decades, school leadership models evolved from *instructional* leadership, focused on teachers’ work, to broader models, like combined *transformational* and *distributed* leadership, which emphasizes commitment and capacities of teachers together with delegation of management tasks. In the same period, researchers concluded that a direct effects approach is not sufficient to determine the impact of school leadership on student outcomes (cf. Hallinger & Heck, 1998). Indirect effects studies revealed many potential mediating variables for school leadership effects, e.g. school culture, school organization, teacher capacities and teacher motivation. Moreover, several contextual and background variables are considered to be important moderators of leadership effects. In general, however, only small effect sizes are found. Less than 1% of the variation in student achievement seems to be associated with differences in educational leadership (Witziers, Bosker & Krüger, 2003).

This study was designed to overcome some of the ‘challenges’ that were identified in school leadership research (cf. Leithwood & Levin, 2005). The *competing values framework* (Quinn & Rohrbaugh, 1983) was used as a generic model for measuring school leadership as well as school cultural and school organisational variables. Furthermore, a mediated effects model was constructed, which included school culture, school organisation, teachers’ work and student engagement as mediating variables. Several context variables were taken into account, like school size, degree of urbanisation of the school environment, family educational culture and the percentage of ethnic minorities in the school population.

The competing values framework encompasses four perspectives on organisations: the *rational goal* orientation aiming for productivity and efficiency, the *internal process* orientation emphasizing stability and control, the *human relations* orientation, aiming for development of human resources and commitment, and the *open systems* orientation, focusing on innovation and resource acquisition. These four perspectives are seen as sub-domains of a larger construct: organizational and managerial effectiveness. Based on the competing values framework, several instruments were developed for measuring school leader behaviour and

school cultural and organisational characteristics, each composed of four scales, representing the four organisation perspectives in the framework.

Data were collected from a representative sample of 103 schools for secondary education in the Netherlands. In total, questionnaires were filled out by 103 school leaders (principals), 998 teachers in senior general education ('havo') and 4336 havo-students in their pre-examination (5th) year. Students in senior general secondary education are considered to be a relatively homogeneous population. In general, schools that took part in the study had a 'mixed' schooltype, like many schools for secondary education in the Netherlands: senior general education is combined with pre-vocational and pre-university education. The mean number of students per school was 1200 and the mean number of teachers was 90.

The psychometric quality of the instruments for measuring school leadership, school culture, school organisation, teachers' work, student engagement and family educational culture was tested using confirmatory factor analysis. Some items were deleted or moved to another scale on the basis of the results of this test and a thorough analysis of the content and meaning of the items. Factor scores were calculated and aggregated at the school level. Because of high correlations between some of the school culture and school organisation variables, these variables were factor analysed. Two new variables were constructed, representing an *achievement orientation* (combined rational goal en and internal process perspective) and a *development orientation* (combined human relations and open system perspective) in school organisations, respectively.

A causal model, examining the effects of school leadership on school performance, was tested using structural equation modeling techniques. School performance was measured by mean national examination scores and mean promotion rates (estimates of the probability that students will be promoted to the next grade) in upper secondary school. These measures were averaged over three subsequent years (2003-2005). The results of the structural analyses indicate that there are significant, but very small, positive indirect effects of school leadership on mean promotion rates, mediated by a development orientation with an emphasis on cooperation, professionalization and innovation, and by the quality of instruction as perceived by students ('teachers' work'). Student engagement, although influenced by teachers' work, is not a mediating variable for leadership effects on achievement and promotion rates. No relation was detected between an achievement oriented and orderly school atmosphere and school performance. Testing an alternative model, however, without contextual variables, revealed a significant positive effect of an achievement orientation on promotion rates, for catholic schools only. This effect 'disappeared' when contextual variables were reintroduced. It was concluded that high expectations and an orderly atmosphere might be school composition effects rather than a cause of high achievement. Surprisingly, in addition to the above-mentioned small indirect effects of school leadership, significant 'direct' effects were detected between the rational goal and open system orientations of school leaders and mean examination scores. In contrast to the indirect effects, the direct effects are negative and they are larger in size. These findings might indicate that school leaders in lower performing schools tend to be more goal-oriented and focused on adaptation and innovation. On the other hand, it might also reflect the attitude of school leaders in high performing schools, who are aiming for stability and continuity and who could be less inclined to improve their schools.

The development oriented school organisation that was described in this study shows similarities to a *learning organisation*. Organisational learning was found to involve a trusting and collaborative climate, a shared and monitored mission, taking initiatives and risks and

ongoing relevant professional development (Mulford & Silins, 2003). It is expected to lead to improved student outcomes. In this study, school leaders appear to contribute to a development orientated school organisation by focusing on a clear mission and shared goals and by having an external and innovative orientation. Productive and inventive school leaders might enhance the performance of their schools. However, their influence seems to be restricted to promotion rates. There is no evidence that school leaders have a significant positive influence on student achievement in secondary education in the Netherlands. It is possible that the differences between school leaders are too small to detect effects of their behaviour on student achievement. In general, school effects on student achievement are very small (cf. Luyten, 1994). The relatively large school size and the autonomy of teachers in secondary education in the Netherlands might also 'buffer' the effects of school leadership.

As pointed out by several researchers, contextual variables play an important role in school leadership models, possibly reducing, neutralising or enhancing leadership effects. Moreover, positive and negative, direct and indirect effects of school leadership might cancel each other out to some extent. The cross sectional approach that was used in this study has strong limitations in revealing the direction of causal effects. It would therefore be recommendable to use longitudinal research designs in future leadership effect studies. Finally, more consistency in the use of models and standardization of instruments is considered to be important for furthering this specific domain of school effectiveness research. The competing values framework, which was used in this study, could be a candidate for this purpose, as a generic leadership model.

BIJLAGEN

Bijlage 1A

Vragenlijst *Handelingen van schoolleiders*

42 items (alpha .81); antwoord: nooit (1) – heel vaak (4)

Rationele Doel Model, 8 items (alpha .63)

- rd1 Ik maak gebruik van gegevens over leerlingprestaties bij het ontwikkelen van de onderwijsdoelen voor de school
- rd2 Ik werk aan het ontwikkelen van doelen die door docenten gemakkelijk vertaald kunnen worden naar operationele doelen
- rd3 In vergaderingen spreek ik met docenten over de onderwijsdoelen van de school
- rd4 Wanneer er beslissingen over curriculumzaken genomen moeten worden, verwijst ik naar de onderwijsdoelen van de school
- rd5 Ik probeer zoveel mogelijk duidelijkheid te scheppen over wat wij met ons onderwijs willen
- rd6 Ik werk eraan dat de onderwijsdoelen van de school eenduidig geïnterpreteerd worden
- rd7 Ik vergelijk onze leerlingresultaten met resultaten van andere scholen (kwaliteitskaarten) bij het nemen van beslissingen
- rd8 Ik zorg dat ik op de hoogte blijf van de onderwijsprogramma's op klasse-niveau om er over te waken dat de onderwijsdoelen van de school gedekt worden

Interne Proces Model, 11 items (alpha .69)

- ip1 Ik werk eraan dat er op school een taakgerichte sfeer heerst
- ip2 Ik werk op school aan het tot stand brengen van een op leren gerichte sfeer
- ip3 Ik werk op school aan het tot stand brengen van een ordelijke sfeer
- ip4 Ik zorg ervoor dat er duidelijke regels worden opgesteld voor het gedrag van leerlingen
- ip5 Ik let er op of het onrustig is in de klassen
- ip6 Als het onrustig is in een klas, tref ik maatregelen
- ip7 Ik werk eraan dat de functies en taken helder omschreven zijn en bekend aan docenten
- ip8 Ik zie er op toe dat de schoolregels door docenten nageleefd worden
- ip9 Ik ben zichtbaar voor docenten (bijvoorbeeld door regelmatig op de gang rond te wandelen)
- ip10 Ik zie er op toe dat de activiteiten en procedures vastgelegd worden in verslagen
- ip11 Ik vraag docenten om mij informatie te bezorgen over wat er in hun lessen gebeurt

Human Relations Model, 14 items (alpha .70)

- hr1 Ik geef op individueel niveau hulp en begeleiding aan docenten
- hr2 Ik laat docenten blijken dat ze met vragen en problemen bij mij terecht kunnen
- hr3 Als een docent problemen heeft in de klas, neem ik het initiatief om met die docent te praten
- hr4 Ik zorg ervoor dat er functioneringsgesprekken met docenten gehouden worden
- hr5 In docentenvergaderingen probeer ik een sfeer van onderling vertrouwen te realiseren
- hr6 Ik werk eraan om de docenten te integreren tot een hecht schoolteam
- hr7 Ik schep structuren en randvoorwaarden ter optimalisering van onderlinge contacten tussen de docenten op school
- hr8 Ik neem tijd om met leerlingen en docenten te praten tijdens pauzes
- hr9 Ik zorg ervoor dat uitzonderlijke prestaties van docenten bekend worden bij de andere docenten
- hr10 Ik complimenteer docenten persoonlijk met hun inzet of prestaties in hun taakuitoefening
- hr11 Ik zorg ervoor dat docenten invloed hebben op de besluitvorming
- hr12 Ik complimenteer tijdens docentenvergaderingen de docenten bij wie de lessen goed verlopen
- hr13 Ik stimuleer docenten om deel te nemen aan nascholingsactiviteiten
- hr14 Ik informeer docenten over mogelijkheden tot nascholing

Open Systeem Model, 9 items (alpha .77)

- os1 Ik praat over de missie of visie van de school met contactpersonen buiten school (bestuur, gemeente, andere scholen, ouders e.d.)
- os2 Ik bezoek andere scholen, conferenties e.d. om nieuwe ideeën op te doen
- os3 Ik creëer extra budget zodat we als school financiële armslag hebben om eigen keuzes te maken
- os4 Ik zorg voor goed contact met de andere leden van de schoolleiding c.q. de centrale directie
- os5 Ik zorg ervoor dat er op deze school altijd minstens één of twee onderwijskundige experimenten draaien
- os6 Ik weet belangrijke externe partijen te overtuigen van de keuzes die wij op onze school maken
- os7 Ik organiseer activiteiten om ervoor te zorgen dat de school een goed imago heeft in de buurt
- os8 Ik probeer om de ouders zoveel mogelijk bij de beslissingen die in deze school genomen worden te betrekken
- os9 Ik onderneem actie n.a.v. trends die ik voorzie

Bijlage 1B

Vragenlijst *Waarneming van handelingen van schoolleiders*

42 items (alpha .94); antwoord: nooit (1) – heel vaak (4)

Rationele Doel Model, 8 items (alpha .84)

- rd1 Gebruik maken van gegevens over leerlingprestaties bij het ontwikkelen van de onderwijsdoelen voor de school
- rd2 Werken aan het ontwikkelen van doelen die door docenten gemakkelijk vertaald kunnen worden naar operationele doelen
- rd3 In vergaderingen spreken met docenten over de onderwijsdoelen van de school
- rd4 Wanneer er beslissingen over curriculumzaken genomen moeten worden, naar de onderwijsdoelen van de school verwijzen
- rd5 Proberen zoveel mogelijk duidelijkheid te scheppen over wat wij met ons onderwijs willen
- rd6 Ervoor zorgen dat de onderwijsdoelen eenduidig geïnterpreteerd worden
- rd7 Vergelijken van onze leerlingresultaten met resultaten van andere scholen (kwaliteitskaarten) bij het nemen van beslissingen
- rd8 Zorgen om op de hoogte te blijven van de onderwijs programma's op klasse-niveau om er over te waken dat de onderwijsdoelen van de school gedekt worden

Interne Proces Model, 11 items (alpha .85)

- ip1 Eraan werken dat er op school een taakgerichte sfeer heerst
- ip2 Werken aan het tot stand brengen van een op leren gerichte sfeer op school
- ip3 Werken aan het tot stand brengen van een ordelijke sfeer op school
- ip4 Ervoor zorgen dat er duidelijke regels worden opgesteld voor het gedrag van leerlingen
- ip5 Erop letten of het onrustig is in de klassen
- ip6 Als het onrustig is in een klas, maatregelen treffen
- ip7 Eraan werken dat de functies en taken helder omschreven zijn en bekend aan docenten
- ip8 Erop toezien dat de schoolregels door docenten nageleefd worden
- ip9 Zichtbaar zijn voor docenten (bijvoorbeeld door regelmatig op de gang rond te wandelen)
- ip10 Erop toezien dat de activiteiten en procedures vastgelegd worden in verslagen
- ip11 Docenten vragen om informatie te verschaffen over wat er in hun lessen gebeurt

Human Relations Model, 14 items (alpha .88)

- hr1 Op individueel niveau hulp en begeleiding aan docenten geven
- hr2 Docenten laten blijken dat ze met vragen en problemen bij hem of haar terecht kunnen
- hr3 Als een docent problemen heeft in de klas (bv. orde-problemen of problemen met lesgeven), initiatief nemen om met die docent te praten
- hr4 Ervoor zorgen dat er functioneringsgesprekken met docenten gehouden worden
- hr5 In docentenvergaderingen proberen om een sfeer van onderling vertrouwen te realiseren
- hr6 Eraan werken om de docenten te integreren tot een hecht schoolteam
- hr7 Structuren en randvoorwaarden scheppen ter optimalisering van onderlinge contacten tussen de docenten op school
- hr8 Tijd nemen om met leerlingen en docenten te praten tijdens pauzes
- hr9 Ervoor zorgen dat uitzonderlijke prestaties van docenten bekend worden bij de andere docenten
- hr10 Docenten persoonlijk met hun inzet of prestaties in hun taakuitoefening complimenteren
- hr11 Ervoor zorgen dat docenten invloed hebben op de besluitvorming
- hr12 Tijdens docentenvergaderingen de docenten bij wie de lessen goed verlopen complimenteren
- hr13 Docenten stimuleren om deel te nemen aan nascholingsactiviteiten
- hr14 Docenten over mogelijkheden tot nascholing informeren

Open Systeem Model, 9 items (alpha .76)

- os1 Over de missie of visie van de school praten met contactpersonen buiten school (bestuur, gemeente, andere scholen, ouders, e.d.)
- os2 Andere scholen, conferenties e.d. bezoeken om nieuwe ideeën op te doen
- os3 Extra budget creëren zodat de school financiële armslag heeft om eigen keuzes te maken
- os4 Zorgen voor goed contact met de andere leden van de schoolleiding c.q. de centrale directie
- os5 Ervoor zorgen dat er op deze school altijd minstens één of twee onderwijskundige experimenten draaien
- os6 Belangrijke externe partijen weten te overtuigen van de keuzes die wij op onze school gemaakt hebben
- os7 Activiteiten organiseren om ervoor te zorgen dat de school een goed imago heeft in de buurt
- os8 Proberen om de ouders zoveel mogelijk bij de beslissingen die in deze school genomen worden te betrekken
- os9 Actie ondernemen n.a.v. trends die zich voordoen

Bijlage 1C **Vragenlijst Schoolcultuur**

40 items (alpha .93); antwoord: niet of nauwelijks (1) – in zeer sterke mate (4)

Bij ons op school...

Rationele Doel Model, 10 items (alpha .81)

- rd1 hebben wij hoge verwachtingen van de leerprestaties van alle leerlingen
- rd2 zijn wij erop gericht om alle leerlingen hun uiterste best te laten doen
- rd3 zijn wij vooral gericht op de cognitieve ontwikkeling van leerlingen
- rd4 vinden wij dat de leerprestaties centraal moeten staan
- rd5 streven wij naar slagingspercentages boven het landelijk gemiddelde
- rd6 zijn wij gericht op het stellen van concrete doelen
- rd7 hechten wij aan efficiënte besluitvorming
- rd8 hechten wij aan actieplannen voor het realiseren van doelstellingen
- rd9 hechten wij aan een taakgerichte sfeer
- rd10 zijn wij gericht op effectief gebruik van onze tijd

Interne Proces Model, 10 items (alpha .83)

- ip1 streven wij naar een ordelijke en rustige werksfeer tijdens de lessen
- ip2 zijn wij gericht op een duidelijke structuur in alle onderwijsactiviteiten
- ip3 hechten wij waarde aan heldere richtlijnen en procedures
- ip4 verwachten wij dat ieder zich aan de regels en procedures van de schoolorganisatie conformeert
- ip5 zijn wij gericht op continuïteit in de schoolorganisatie
- ip6 zijn wij gehecht aan planmatig werken
- ip7 streven wij naar een optimale interne informatievoorziening
- ip8 hechten wij waarde aan het nauwgezet registreren van de ontwikkeling van leerlingen
- ip9 achten wij zelfevaluatie van groot belang
- ip10 hechten wij bij besluitvorming aan een zorgvuldige analyse van gegevens

Human Relations Model, 10 items (alpha .87)

- hr1 vinden wij investeren in het personeel van groot belang
- hr2 zijn wij erop gericht om te werken aan onze eigen professionele ontwikkeling
- hr3 staat de professionaliteit van het personeel centraal
- hr4 streven wij naar een optimaal gebruik van ieders specifieke kwaliteiten
- hr5 hechten wij eraan dat iedereen zich verantwoordelijk voelt voor de school
- hr6 heerst een 'wij-gevoel'
- hr7 zijn wij erop gericht om met elkaar samen te werken
- hr8 hechten wij aan een actieve inbreng van docenten bij de beleidsontwikkeling
- hr9 verwachten wij van docenten dat zij geïnteresseerd zijn in elkaars werk
- hr10 hechten wij eraan om open te kunnen discussiëren over verschillen van mening

Open Systeem Model, 10 items (alpha .80)

- os1 hechten wij aan een duidelijke profilering van de school
- os2 zijn wij gericht op het zorgvuldig onderhouden van externe contacten
- os3 zijn wij erop gericht om extra middelen en faciliteiten voor onze school te verwerven
- os4 zijn wij erop gericht om de ouders zoveel mogelijk te betrekken bij de school
- os5 achten wij evaluatie door externe instanties van belang
- os6 zijn wij erop gericht om in te spelen op actuele maatschappelijke ontwikkelingen
- os7 vinden wij een positieve houding tegenover onderwijsvernieuwingen van belang
- os8 wordt van ons verwacht dat wij in staat zijn om vernieuwingen in te voeren
- os9 hechten wij er waarde aan om bekend te staan als innovatieve school
- os10 verwachten wij van iedereen een flexibele houding ten opzichte van veranderingen

Bijlage 1D

Vragenlijst Schoolorganisatiepraktijk

40 items (alpha .92); antwoord: niet of nauwelijks (1) – in hoge mate (4)

Bij ons op school...

Rationele Doel Model, 10 items (alpha .78)

- rd1 is er een selectief toelatingsbeleid voor leerlingen
- rd2 staan hoge leerprestaties centraal in de doelstellingen
- rd3 vormen de resultaten van de leerlingen het uitgangspunt bij de ontwikkeling van het onderwijsbeleid
- rd4 wordt aan de leerlingen duidelijk gemaakt dat hoge prestaties verwacht worden
- rd5 worden de resultaten van leerlingen systematisch geanalyseerd
- rd6 worden streefgetallen gehanteerd voor de doorstroom van leerlingen binnen de school
- rd7 is het onderwijsbeleid gebaseerd op concrete onderwijskundige doelstellingen
- rd8 worden de onderwijsdoelstellingen als uitgangspunt gebruikt bij de besluitvorming
- rd9 worden beleidsvoorstellen voorzien van een gedetailleerde planning
- rd10 wordt het onderwijsbeleid vertaald in operationele doelen voor de docenten

Interne Proces Model, 10 items (alpha .77)

- ip1 ligt in de onderwijsvisie de nadruk op een ordelijk en veilig leerklimaat
- ip2 zijn er duidelijke procedures om absentie van leerlingen tegen te gaan
- ip3 wordt erop toegezien dat leerlingen zich aan de gedragsregels houden
- ip4 is de leerlingenzorg en -begeleiding structureel ingebed in de organisatie
- ip5 zijn er heldere taak- en functieomschrijvingen voor docenten
- ip6 zijn er procedures om ervoor te zorgen dat iedereen tijdig van de juiste informatie wordt voorzien
- ip7 zijn er indicatoren vastgesteld om te meten of de doelen worden bereikt
- ip8 is er een geavanceerd leerlingvolgsysteem
- ip9 wordt er gewerkt met geautomatiseerde informatiesystemen
- ip10 zijn er procedures, planningen en instrumenten voor zelfevaluatie

Human Relations Model, 10 items (alpha .74)

- os1 ligt in het personeelsbeleid de nadruk op competentieontwikkeling
- os2 is er een uitgewerkt nascholingsplan
- os3 houden docenten een persoonlijk scholings- of ontwikkelingsplan bij
- os4 worden systematisch gesprekken gevoerd met docenten over hun loopbaan
- os5 vindt taakdifferentiatie plaats op basis van de aanwezige deskundigheden en ervaringen
- os6 wordt voor buitengewone prestaties van docenten een bijzondere vorm van waardering gehanteerd
- os7 zijn de docenten actief betrokken bij de besluitvorming
- os8 worden nieuwe docenten gecoacht
- os9 maken docenten gebruik van intervisie om van elkaar te leren
- os10 regelen docenten allerlei zaken met elkaar door informeel overleg

Open Systeem Model, 10 items (alpha .71)

- os1 worden marketingactiviteiten ondernomen om de concurrentiepositie te verbeteren
- os2 zijn er richtlijnen voor een professionele aanpak van de externe communicatie
- os3 worden activiteiten ondernomen om het imago van de school te onderzoeken en verbeteren
- os4 vindt systematisch overleg plaats met vervolgoopleidingen over de aansluiting
- os5 is er beleid om de ouders zoveel mogelijk bij het onderwijs te betrekken
- os6 wordt bij de ontwikkeling van verbeterplannen het advies van derden (buiten de school) verworven
- os7 is er beleid om de toepassing van nieuwe technologie in het onderwijs te bevorderen
- os8 worden docenten ondersteund bij het invoeren van onderwijsvernieuwingen
- os9 is het werken aan onderwijsvernieuwing in de organisatie ingebed
- os10 worden docenten in de gelegenheid gesteld om initiatieven te nemen

Bijlage 1E **Vragenlijst Schoolbeleving van leerlingen**

42 items (alpha .90)

antwoord: helemaal mee oneens (1)–helemaal mee eens (4)

Zich thuisvoelen op school - A sense of Belonging (alpha .78)

Omgang met medeleerlingen

- llomln1 Op deze school kun je echte vrienden maken
- llomln2 Ik kan mijn klasgenoten vertrouwen
- llomln3 Ik heb voldoende vrienden op school
- llomln4 Ik kan goed opschieten met de meeste andere leerlingen die ik ken op school
- llomln5 Ik kan op mijn klasgenoten rekenen als ik iets niet begrijp

Omgang met leraren

- llomler1 Ik kan goed opschieten met de meeste leraren
- llomler2 De meeste leraren behandelen mij met respect
- llomler3 De meeste leraren zijn in mij als persoon geïnteresseerd
- llomler4 De meeste leraren lijken mij te begrijpen
- llomler5 De meeste leraren zorgen ervoor dat ik mij op mijn gemak voel in de klas

Waardering - Valuing (alpha .77)

Waardering van onderwijs

- llwaond1 Iedereen zou zoveel mogelijk onderwijs moeten krijgen
- llwaond2 Ik vind mijn schoolwerk echt belangrijk
- llwaond3 Ik vind het heel belangrijk om op school kennis op te doen en vaardigheden te ontwikkelen
- llwaond4 Ik doe mijn best om goede cijfers te halen
- llwaond5 De dingen die ik op school leer zijn nuttig voor later

Waardering van de school

- llwasch1 Ik ga graag naar school
- llwasch2 Ik vind het meestal leuk op school
- llwasch3 Ik vind het prettig om de pauze en de tussenuren op school door te brengen
- llwasch4 De sfeer bij ons op school is goed
- llwasch5 Ik heb het gevoel dat ik thuishoor op deze school

Kwaliteit van de Instructie - Quality of Instruction (alpha .81)

Resultaatgerichtheid / Uitdagendheid

- llresul1 De meeste leraren verwachten dat ik altijd mijn best doe
- llresul2 De meeste leraren laten merken dat zij goede resultaten verwachten van iedereen
- llresul3 De meeste vakken zijn voldoende uitdagend voor mij

Ondersteuning door leraren

- llstler1 De meeste leraren moedigen mij aan als ik iets goed doe in de klas
- llstler2 Van de meeste leraren krijg ik de gelegenheid om mijn mening te geven in de klas
- llstler3 De meeste leraren praten regelmatig met mij over mijn schoolwerk
- llstler4 De meeste leraren doen erg hun best om leerlingen te helpen

Klassenklimaat

- llwklim1 Tijdens de lessen is er meestal een rustige werksfeer
- llwklim2 De meeste leraren hebben duidelijke regels over wat wel of niet kan tijdens de lessen
- llwklim3 Tijdens de meeste lessen wordt de tijd goed besteed

Organisatie van het onderwijs

- llorond1 Voor de meeste vakken zijn er duidelijke leerdoelen (welke kennis en vaardigheden je moet leren)
- llorond2 De meeste lessen hebben een goede structuur
- llorond3 In de meeste vakken wordt op een goede manier getoetst

Invulling van de lessen

- llinvul1 De meeste leraren zorgen voor afwisselende activiteiten tijdens de lessen
- llinvul2 De meeste leraren proberen in hun lessen aan te sluiten bij wat leerlingen weten en beleven
- llinvul3 De meeste leraren leggen bij het begeleiden van schoolwerk de nadruk meer op de aanpak dan op de oplossing
- llinvul4 De meeste leraren geven de gelegenheid om zelfstandig te werken

Onderwijsondersteuning van thuis - *Family Educational Culture* (alpha .59)

- llstthu1 Mijn ouders/verzorgers zijn altijd bereid om mij te helpen met mijn huiswerk
- llstthu2 Ik heb thuis een rustige plek om mijn huiswerk te doen
- llstthu3 Mijn ouders/verzorgers letten erop dat ik mijn huiswerk doe
- llstthu4 Ik bespreek mijn schoolwerk vaak met mijn ouders/verzorgers
- llstthu5 Mijn ouders/verzorgers gaan meestal naar ouderavonden en andere speciale gebeurtenissen op school

Bijlage 2 Methoden

A. Factoranalyse

Met behulp van een *factoranalyse* kan worden onderzocht wat de onderliggende latente structuur is van een set variabelen (Garson, z.j., a). Uitgaande van de correlaties (of de covarianties) tussen variabelen wordt nagegaan, welke variabelen aspecten meten van dezelfde onderliggende dimensie (Field, 2000, p. 423). Als er op basis van theorie een hypothetisch model is opgesteld voor de relaties tussen de indicatoren en de onderliggende dimensies, dan kan een *confirmatieve* factoranalyse uitgevoerd worden (Jöreskog & Sörbom, 1993a). Als er geen vooraf opgesteld model beschikbaar is, dan is er sprake van *exploratieve* factoranalyse.

Met de meeste standaard statistische pakketten kan een factoranalyse uitgevoerd worden. Feitelijk gaat het hier meestal om een exploratieve factoranalyse, omdat er niet een vooraf opgesteld model wordt ingevoerd. De uitkomsten van de standaard factoranalyse geven aan in welke mate de indicator variabelen ‘laden’ op een aantal factoren. Deze uitkomsten worden echter traditioneel ook vaak gebruikt voor confirmatieve doeleinden (Garson, z.j., a). Er wordt dan bekeken of verbanden die gevonden worden tussen de indicator variabelen en de latente variabelen overeenkomen met de relaties in het hypothetische model. Rowe (1995, 2002) beschrijft een aantal bezwaren tegen de exploratieve *data-fishing* methode. De procedures die worden gebruikt er gaan bijvoorbeeld vanuit dat de variabelen continu zijn, zonder meetfout gemeten en normaal verdeeld met homogene variantie. Dit is meestal niet het geval, waardoor de uitkomsten onvoldoende betrouwbaar zijn.

Voor het uitvoeren van een echte confirmatieve factoranalyse, waarbij een vooraf opgesteld model wordt getest op de data, wordt gebruik gemaakt van *Structural Equation Modeling (SEM)* methoden (zie verder). Zowel Rowe (1995, p. 15), als Hox (2002, p. 233) en Garson (z.j., b), geven echter aan dat een exploratieve factoranalyse nuttig kan zijn als voorloper van een confirmatieve factoranalyse, om na te gaan of het juiste aantal factoren wordt onderscheiden en of de indicatoren inderdaad laden op de bijbehorende factoren. Op basis van de uitkomsten van een exploratieve factoranalyse kunnen verschillende alternatieve modellen worden opgesteld, die vervolgens in een confirmatieve factoranalyse worden getest (Garson, z.j., b).

B. Structural equation modeling in LISREL

Om de validiteit van een model voor de relaties tussen de indicator variabelen en de latente variabelen vast te stellen, dient in een confirmatieve factoranalyse te worden nagegaan in hoeverre het model past bij de geobserveerde data. Dit kan worden uitgevoerd met behulp van *structural equation modeling*. Voor structural equation modeling is gespecialiseerde statistische software nodig, zoals LISREL of Mplus. Voor het hier beschreven onderzoek werd LISREL (versie 8.8) gebruikt. Een volledig LISREL model bestaat uit twee delen: het *meetmodel* waarin wordt gespecificeerd hoe de geobserveerde variabelen (indicatoren) gekoppeld zijn aan de latente variabelen en het *structurele model* waarin de causale relaties tussen de latente variabelen worden beschreven (Jöreskog & Sörbom, 1993b). In LISREL kunnen onderdelen van het volledige model (zoals het meetmodel) apart getest worden.

In een standaard LISREL meetmodel voor een confirmatieve factoranalyse wordt ervan uitgegaan dat elke indicator op slechts één factor laadt, dat de factoren onderling correleren en dat de meetfouten voor de geobserveerde variabelen niet correleren (Garson, z.j., b). Doordat de meetfouten ook worden meegenomen in het meetmodel, kan een zuivere schatting verkregen worden van de factorlading. Als de meetfout groot is, dan is de factorlading echter toch minder betrouwbaar.

Er zijn verschillende *schattingmethoden* beschikbaar om de parameters in een LISREL model te beschrijven (Jöreskog & Sörbom, 1993b). Een veel gebruikte methode is *maximum likelihood estimation* (MLE). Deze methode probeert de parameters zodanig te schatten, dat de kans het hoogste is dat de geobserveerde data gereproduceerd worden. Hoewel du Toit en du Toit (2001, p. 199) aangegeven dat voor de analyse van ordinale data de WLS (*Weighted Least Squares*) methode

gebruikt zou moeten worden, werkt MLE volgens Garson (z.j., b) ook goed wanneer er vier of meer Likert categoriën zijn, bij een redelijk grote steekproef.

C. Analyseren van een model

Voor het analyseren van modellen in structural equation modeling zijn er drie benaderingen (Jöreskog & Sörbom, 1993a, p. 115):

- *Strictly confirmatory* (SC): hierbij wordt slechts één model getest en dit wordt geaccepteerd of verworpen.
- *Alternative models* (AM): hierbij worden verschillende modellen getest op dezelfde dataset en dan wordt het beste model geselecteerd. .
- *Model generating* (MG): hierbij wordt een voorlopig model opgesteld, dat wordt getest op de data en wordt bijgesteld als het niet voldoende past. Bijstellingen kunnen zowel op basis van theoretische overwegingen of op basis van de structuur in de geobserveerde data plaatsvinden.

De *model generating* benadering komt het meeste voor. Een probleem bij deze benadering kan zijn dat het model teveel wordt afgestemd op de beschikbare data en dat het mogelijk niet past op nieuwe data. Er is dan sprake van *kanskapitalisatie* (capitalization on chance, MacCallum, Roznowski & Necowitz, 1992). Het model wordt afgestemd op toevallige karakteristieken van de data. Het is daarom beter om de uiteindelijke oplossing ook nog te kruisvalideren (*cross-validation*) aan de hand van een andere dataset (Garson, z.j., b).

D. Beoordelen van de geschiktheid van een model

Om te bepalen hoe goed een meetmodel bij de data past, kan volgens Jöreskog en Sörbom (1993a, p. 121) de volgende procedure gevolgd worden:

1. Eerst wordt nagegaan of de parameterschattingen kloppen volgens het vooraf gespecificeerde model. Hierbij wordt bijvoorbeeld onderzocht of de indicator variabelen voldoende laden op de bijbehorende factor. Als de kritieke waarde (*critical ratio*) voor een factorlading groter is dan 1.96 ($\alpha = 0.05$) dan is de parameter significant (Garson, z.j., b.). De kritieke waarde is gelijk aan de ongestandaardiseerde parameterwaarde gedeeld door de standaardfout. De sterkte van de relatie wordt weergegeven in R^2 (*squared multiple correlation*). Dit is het percentage van de variantie in de indicator variabele dat wordt verklaard door de latente variabele.
2. Vervolgens wordt de *overall fit* bekeken. Hiermee wordt aangegeven in hoeverre het model als geheel bij de data past. Op de beoordeling van de overall fit wordt hieronder ingegaan.
3. Als het model niet goed past, kunnen vervolgens de residuen en de door LISREL gegenereerde modificatie indices worden beschouwd. Een residu is het verschil tussen de geobserveerde en de geschatte covariantie voor een tweetal variabelen. De modificatie indices geven aan hoe het model verbeterd kan worden door extra parameters te schatten. Als een indicator variabele blijkt te laden op twee latente variabelen kan bijvoorbeeld een extra relatie toegevoegd worden. Voor de modificatie indices wordt uitgegaan van veranderingen die een vermindering van chi-kwadraat (zie verder) opleveren.

Volgens Garson (z.j., b) kan ook eerst stap 2 worden uitgevoerd en dan stap 1. De overall fit (modelpassing) wordt gebruikt als een eerste indicatie voor de geschiktheid van het model en vervolgens wordt aan de hand van de parameterschattingen onderzocht, of het model terecht wordt geaccepteerd of verworpen. Een goede overall fit betekent bijvoorbeeld nog niet dat alle relaties in het model voldoende sterk zijn. Als de correlaties in een set variabelen klein zijn, wordt sneller een goede modelpassing gevonden. Bij lage correlaties zullen de factorladingen echter ook laag zijn en mogelijk niet significant.

E. Chi-kwadraat

Een veelgebruikte test voor de modelpassing is *chi-kwadraat*. Als de waarde van chi-kwadraat significant is, wijkt de covariantie structuur van het model significant af van de geobserveerde covariantie matrix (Garson, z.j., b.). Bij een significante chi-kwadraat wordt het model verworpen. Chi-kwadraat is een *goodness-of-fit* maat, maar feitelijk geeft het de *badness-of-fit* aan. Jöreskog en

Sörbom (1993a) geven aan dat het in de praktijk (in de model generating benadering) zinvoller is om chi-kwadraat als een maat voor de passing van een model te gebruiken en niet zozeer als een test. Hoge waarden van chi-kwadraat wijzen op een slechte fit, lage waarden op een goede fit.

Chi-kwadraat is gevoelig voor de steekproefomvang (Garson, z.j., b.). Hoe groter de steekproefomvang, des te groter de kans dat chi-kwadraat significant is. Ook de complexiteit van het model is van invloed op chi-kwadraat. In het algemeen levert een complex model namelijk een betere modelpassing op dan een eenvoudiger (*parsimonious*) model, waarin minder parameters worden geschat ten opzichte van het aantal variabelen.

F. Goodness-of-fit indices

Naast chi-kwadraat bestaan er nog een groot aantal andere *goodness-of-fit indices*. Deze goodness-of-fit indices hebben volgens Hox (2002, p. 239) allemaal de chi-kwadraat test als uitgangspunt, maar ze zijn gebaseerd op verschillende principes. Welke passingsmaten precies worden gegenereerd bij structural equation modeling, kan afhangen van het softwarepakket en van de statistische methode die wordt gekozen. Over de keuze van de passingsmaten die vermeld moeten worden bij de beschrijving van een model wordt regelmatig gepubliceerd. Door middel van simulatieonderzoek wordt nagegaan welke indices het beste inzicht geven in de passing van een model, onder bepaalde omstandigheden. Hieronder volgt een korte beschrijving van de ontwikkelingen in het gebruik van passingsmaten.

Jöreskog en Sörbom (1993b) stellen in een handleiding bij LISREL voor om naast chi-kwadraat de *Goodness-of-Fit Index* (GFI), de *Adjusted Goodness-of-Fit Index* (AGFI) en *Root Mean square Residual* (RMR) weer te geven. Deze passingsmaten zijn allemaal gebaseerd op het verschil tussen de door het model voorspelde en de geobserveerde covarianties (Garson, z.j., b.). GFI en AGFI variëren beide van 0 tot 1. Hoe hoger de waarde, hoe beter het model past. De waarde van GFI wordt beïnvloed door het aantal vrijheidsgraden in het model, in de AGFI wordt hiervoor gecorrigeerd. Beide passingsmaten zijn gevoelig voor de steekproefomvang. GFI heeft de neiging om groter te worden als de steekproef omvang toeneemt, terwijl AGFI de modelpassing onderschat bij een kleine steekproefomvang. RMR wordt berekend op basis van de residuen (zie boven). Hoe kleiner RMR is, hoe beter het model past.

Hox (2002 p. 240) geeft in zijn leerboek over meerniveau analyse aan dat in (meerniveau) structural equation modeling in het algemeen geadviseerd wordt om een aantal goodness-of-fit indices te kiezen die zijn gebaseerd op verschillende principes. Hox kiest voor de *Goodness-of-Fit Index* GFI, de *Comparative Fit Index* CFI en de *Root Mean Square Error of Approximation*, RMSEA. CFI is gebaseerd op de vergelijking van het model met een alternatief model (i.c. het nul model, waarin de latente variabelen niet correleren). CFI varieert van 0 tot 1. Hoe groter CFI is, hoe beter het model past. RMSEA is evenals (A)GFI en RMR gebaseerd op het verschil tussen voorspelde en geobserveerde covarianties, maar in RMSEA wordt daarbij gecorrigeerd voor de complexiteit van het model. Voor RMSEA geldt dat hoe kleiner de waarde is, hoe beter de fit.

Een recent overzicht van het gebruik van passingsmaten in rapportages van de resultaten van structural equation modeling wordt gegeven door Schreiber, Nora, Stage, Barlow en King (2006). De auteurs vermelden voor een groot aantal indices de regels die in het algemeen gehanteerd worden om te beslissen of de modelpassing acceptabel is. Daarbij verwijzen zij onder meer naar het werk van Hu en Bentler (1999) en Yu (2002), die op basis van simulaties de vuistregels onderzochten voor de *cutoff criteria* van een aantal veelgebruikte passingsmaten. Hu en Bentler (1999) waren er daarbij op gericht om met behulp van een combinatie van twee indices een goede inschatting van de modelpassing te kunnen maken. In de '2-index' presentatie strategie wordt SRMR gecombineerd met een andere fit index. SRMR is de gestandaardiseerde versie van RMR (zie boven). Schreiber et al. (2006) concluderen dat CFI, RMSEA en TLI geschikte maten zijn om de modelpassing te beoordelen. TLI is de *Tucker-Lewis Index*, die ook wel de *Non-Normed Fit Index* (NNFI) wordt genoemd. Deze index is gebaseerd op hetzelfde principe als CFI. TLI corrigeert voor de complexiteit van het model en is minder gevoelig voor de steekproefomvang dan andere indices (Garson, z.j., b.). Een combinatie van SRMR met CFI, RMSEA of TLI (of nog een aantal andere indicatoren) levert goede resultaten op.

Alleen als de steekproefomvang beperkt is ($N \leq 250$) dan geldt dat TLI en RMSEA minder geschikt zijn.

Voor het vergelijken van de passing van verschillende modellen wordt AIC aanbevolen (Schreiber et al., 2006), vooral als op basis van de andere passingsmaten blijkt dat de verschillen tussen de modellen klein zijn. AIC staat voor *Akaike Information Criterion* (Garson, z.j., b). Als de AIC waarden voor twee modellen vergeleken worden, geldt dat het model met de laagste waarde het beste is. AIC kan gebruikt worden voor zowel geneste als niet geneste modellen. Een model is genest ten opzichte van een ander model als het evenveel variabelen bevat, terwijl er meer of minder relaties tussen die variabelen gespecificeerd zijn. Voor geneste modellen wordt vooral de chi-kwadraat test gebruikt (Schreiber et al., 2006).

Uit het onderzoek van Hu en Bentler (1999) blijkt dat GFI en AGFI minder geschikt zijn om de passing van een model te beoordelen. De reden waarom GFI en AGFI hier toch behandeld worden is dat deze indices in wat oudere publicaties regelmatig voorkomen (zie bijvoorbeeld Kaplan & Elliot, 1997; Rowe & Hill, 1998) en ook in recente publicatie nog opduiken (zie bijvoorbeeld Toland & De Ayala, 2005). Bovendien nemen deze indices in bepaalde LISREL output een prominente plaats in. Hieronder volgt een overzicht van de besproken passingsmaten en vuistregels die daarbij gehanteerd worden (gebaseerd op Schreiber et al., 2006, p. 330, tabel 2).

Goodness-of-fit index	Afkorting	Vuistregel voor beoordelen model fit
Chi-kwadraat	χ^2	$X^2/df \leq 2$ of 3, vergelijking geneste modellen
Akaike information criterion	AIC	hoe kleiner hoe beter, vergelijking niet-geneste modellen
Comparative fit index	CFI	$\geq .95$ voor accepteren van het model
Tucker-Lewis index	TLI	$\geq .95$ voor accepteren van het model
Root mean square residual	RMR	hoe kleiner hoe beter; = 0 bij perfecte fit
Standardized RMR	SRMR	$\leq .08$ voor accepteren van het model
Root mean square error of approximation	RMSEA	$< .06$ (tot $.08$) voor accepteren van het model
(Adjusted) goodness-of-fit index	(A)GFI	$\geq .95$ niet aanbevolen vanwege slechte prestatie in simulatiestudies

Zoals hierboven al vermeld is, betekent een goede modelpassing niet dat elke relatie binnen het model ook voldoende sterk is. Bovendien is een goed passend model niet noodzakelijk het enige of het beste model dat past op de data (Tomarken & Waller, 2005; Garson, z.j., b.).

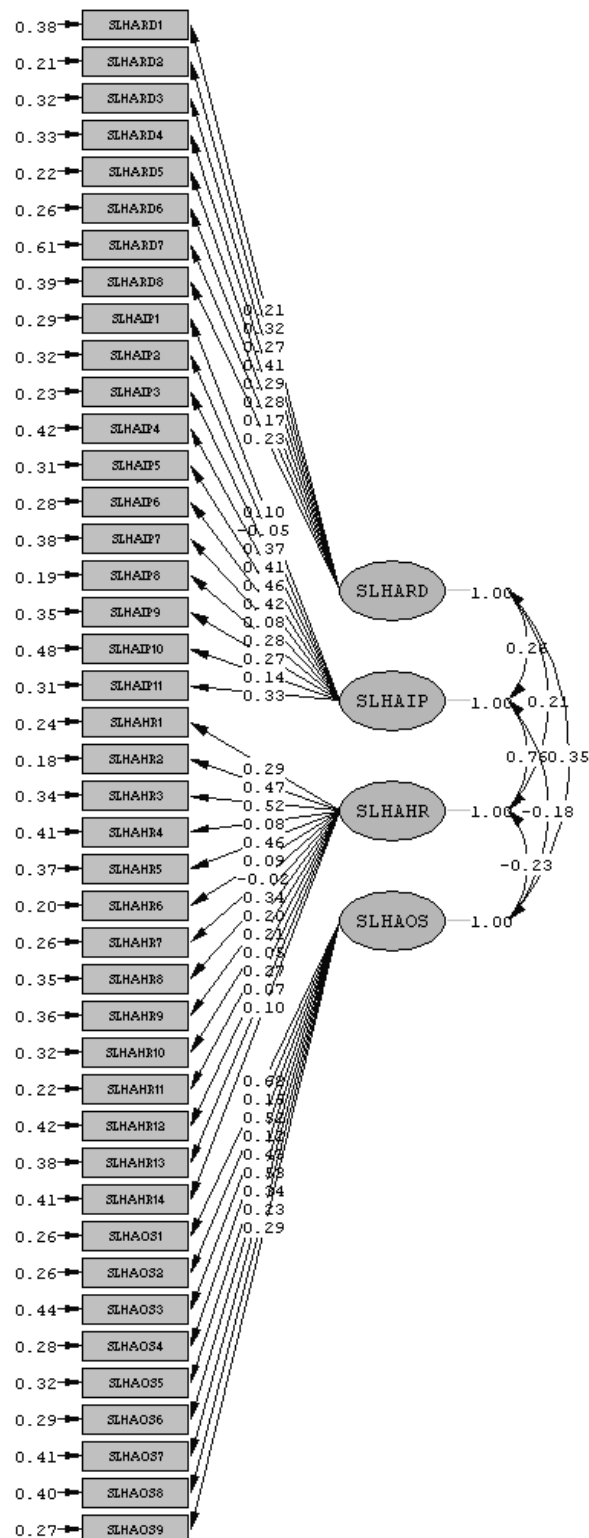
G. Factor scores en latente variabele scores

Factoranalyse wordt niet alleen gebruikt worden om de validiteit van een model te bepalen, maar ook om het aantal variabelen in een onderzoeksmodel te reduceren (Field, 2000, p. 424). Om voor vervolganalyses de *factorscore* te berekenen voor een bepaalde case kan de gestandaardiseerde score op elke variabele worden vermenigvuldigd met de factorlading voor die variabele, waarna de producten worden opgeteld (Garson z.j., a). Een betere benadering van de factorscore wordt echter verkregen wanneer de *factor score coëfficiënten* als gewicht worden gebruikt in plaats van de factorladingen (Field, 2000, p. 429 e.v.). Er zijn verschillende methoden bekend om factor score coëfficiënten te berekenen, zoals de regressiemethode, de Bartlett methode of de Anderson-Rubin methode, die beschikbaar zijn in SPSS (Field, 2000, p. 431, Morrison, 2002).

Voor LISREL modellen is door Jöreskog (2000) een nieuwe methode ontwikkeld om *latente variabele scores* te berekenen. Deze scores zijn niet vergelijkbaar met de factorscores. De latente variabele is geen (lineaire) combinatie van de indicatoren. De latente variabele score bootst het gedrag na van de latente variabele die door structural equation modeling (LISREL) bepaald wordt (John von Briesen, SSI, persoonlijke communicatie²³). De procedure die in LISREL gebruikt moet worden om de latente variabele scores te genereren is beschreven door Mels (2004).

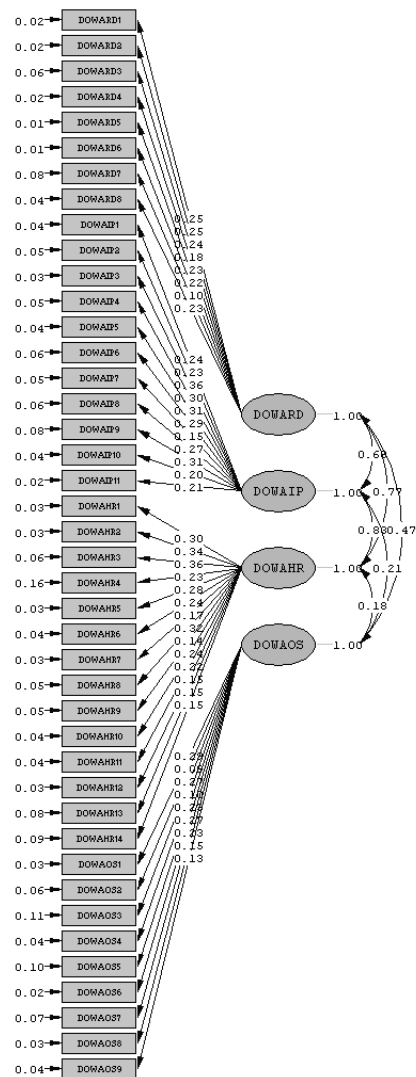
²³ SSI = Scientific Software International Inc., de producent van LISREL.

Bijlage 3 Meetmodellen en parameterschattingen voor de handelingen van schoolleiders

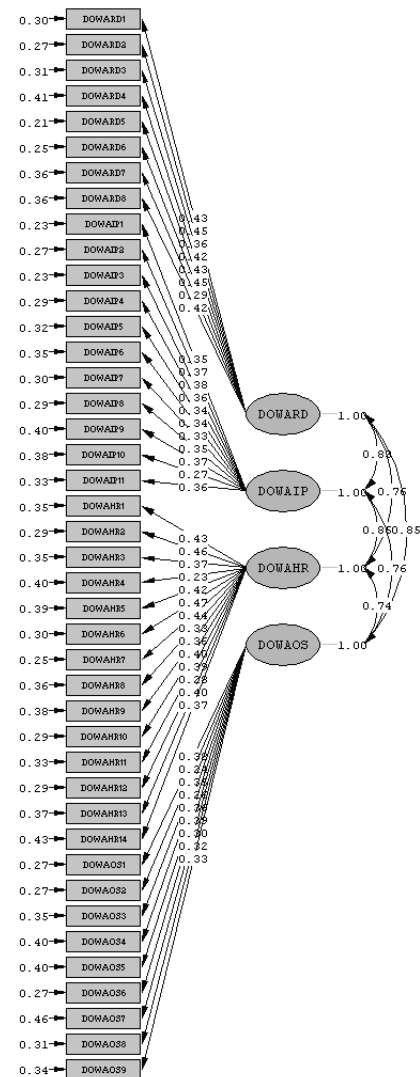


Chi-Square=1198.29, df=813, P-value=0.00000, RMSEA=0.068

Bijlage 3A: Parameterschattingen bij een éénniveau confirmatieve factoranalyse voor het totale model van schoolleidersgedrag, aan de hand van schoolleidersdata (n= 103) in de standardized solution (latente variabelen gestandaardiseerd). SLHA = schoolleiders handelingen; RD = Rationele doel schaal, IP= interne proces schaal, HR = human relations schaal, OS = open systeem schaal.



Chi-Square=3796.76, df=1626, P-value=0.00000, RMSEA=0.052



Chi-Square=3796.76, df=1626, P-value=0.00000, RMSEA=0.052

Bijlage 3B: Parameterschattingen bij een meerniveau confirmatieve factoranalyse met Muthén's approximate Maximum Likelihood solution (MUML), voor het totale model van schoolleidersgedrag, aan de hand van docentendata. Links het schoolniveau (between groups) en rechts het docentenniveau (within groups) in de common metric standardized solution (latente variabelen gestandaardiseerd). DOWA = docent waarneming van handelingen; RD = Rationele doel schaal, IP= interne proces schaal, HR = human relations schaal, OS = open systeem schaal.

Bijlage 4A **Handelingen van schoolleiders: nieuw model**

26 items (schoolleiders alpha .72, docenten alpha .91)

Rationele Doel Model, 7 items (schoolleiders alpha .67; docenten alpha .85)

- rd1 Gebruik maken van gegevens over leerlingprestaties bij het ontwikkelen van de onderwijsdoelen voor de school
- rd2 Werken aan het ontwikkelen van doelen die door docenten gemakkelijk vertaald kunnen worden naar operationele doelen
- rd3 In vergaderingen spreken met docenten over de onderwijsdoelen van de school
- rd4 Wanneer er beslissingen over curriculumzaken genomen moeten worden, naar de onderwijsdoelen van de school verwijzen
- rd5 Proberen zoveel mogelijk duidelijkheid te scheppen over wat wij met ons onderwijs willen
- rd6 Ervoor zorgen dat de onderwijsdoelen eenduidig geïnterpreteerd worden
- ip2 Werken aan het tot stand brengen van een op leren gerichte sfeer op school

Interne Proces Model, 5 items (schoolleiders alpha .70; docenten alpha .76)

- ip3 Werken aan het tot stand brengen van een ordelijke sfeer op school
- ip4 Ervoor zorgen dat er duidelijke regels worden opgesteld voor het gedrag van leerlingen
- ip5 Erop letten of het onrustig is in de klassen
- ip8 Erop toezien dat de schoolregels door docenten nageleefd worden
- ip10 Erop toezien dat de activiteiten en procedures vastgelegd worden in verslagen

Human Relations Model, 8 items (schoolleiders alpha .77; docenten alpha .84)

- ip11 Docenten vragen om informatie te verschaffen over wat er in hun lessen gebeurt
- hr1 Op individueel niveau hulp en begeleiding aan docenten geven
- hr2 Docenten laten blijken dat ze met vragen en problemen bij hem of haar terecht kunnen
- hr3 Als een docent problemen heeft in de klas (bv. orde-problemen of problemen met lesgeven), initiatief nemen om met die docent te praten
- hr5 In docentenvergaderingen proberen om een sfeer van onderling vertrouwen te realiseren
- hr8 Tijd nemen om met leerlingen en docenten te praten tijdens pauzes
- hr10 Docenten persoonlijk met hun inzet of prestaties in hun taakuitoefening complimenteren
- hr12 Tijdens docentenvergaderingen de docenten bij wie de lessen goed verlopen complimenteren

Open Systeem Model, 6 items (schoolleiders alpha .76; docenten alpha .73)

- os1 Over de missie of visie van de school praten met contactpersonen buiten school (bestuur, gemeente, andere scholen, ouders, e.d.)
- os3 Extra budget creëren zodat de school financiële armslag heeft om eigen keuzes te maken
- os6 Belangrijke externe partijen weten te overtuigen van de keuzes die wij op onze school gemaakt hebben
- os7 Activiteiten organiseren om ervoor te zorgen dat de school een goed imago heeft in de buurt
- os8 Proberen om de ouders zoveel mogelijk bij de beslissingen die in deze school genomen worden te betrekken
- os9 Actie ondernemen n.a.v. trends die zich voordoen

Bijlage 4B

Schoolcultuur: nieuw model

29 items (alpha .91)

Bij ons op school...

Rationele Doel Model, 6 items (alpha .80) > *Prestatiegerichtheid*²⁴

- rd1 hebben wij hoge verwachtingen van de leerprestaties van alle leerlingen
- rd2 zijn wij erop gericht om alle leerlingen hun uiterste best te laten doen
- rd3 zijn wij vooral gericht op de cognitieve ontwikkeling van leerlingen
- rd4 vinden wij dat de leerprestaties centraal moeten staan
- rd5 streven wij naar slagingspercentages boven het landelijk gemiddelde
- ip1 streven wij naar een ordelijke en rustige werksfeer tijdens de lessen

Interne Proces Model, 7 items (alpha .80) > *Prestatiegerichtheid*

- ip2 zijn wij gericht op een duidelijke structuur in alle onderwijsactiviteiten
- ip3 hechten wij waarde aan heldere richtlijnen en procedures
- ip4 verwachten wij dat ieder zich aan de regels en procedures van de schoolorganisatie conformeert
- ip5 zijn wij gericht op continuïteit in de schoolorganisatie
- ip6 zijn wij gehecht aan planmatig werken
- ip7 streven wij naar een optimale interne informatievoorziening
- ip10 hechten wij bij besluitvorming aan een zorgvuldige analyse van gegevens

Human Relations Model, 9 items (alpha .86) > *Ontwikkelingsgerichtheid*²⁵

- ip9 achten wij zelfevaluatie van groot belang
- hr1 vinden wij investeren in het personeel van groot belang
- hr2 zijn wij erop gericht om te werken aan onze eigen professionele ontwikkeling
- hr4 streven wij naar een optimaal gebruik van ieders specifieke kwaliteiten
- hr5 hechten wij eraan dat iedereen zich verantwoordelijk voelt voor de school
- hr6 heerst een "wij-gevoel"
- hr7 zijn wij erop gericht om met elkaar samen te werken
- hr9 verwachten wij van docenten dat zij geïnteresseerd zijn in elkaars werk
- hr10 hechten wij eraan om open te kunnen discussiëren over verschillen van mening

Open Systeem Model, 7 items (alpha .78) > *Ontwikkelingsgerichtheid*

- os1 hechten wij aan een duidelijke profilering van de school
- os3 zijn wij erop gericht om extra middelen en faciliteiten voor onze school te verwerven
- os6 zijn wij erop gericht om in te spelen op actuele maatschappelijke ontwikkelingen
- os7 vinden wij een positieve houding tegenover onderwijsvernieuwingen van belang
- os8 wordt van ons verwacht dat wij in staat zijn om vernieuwingen in te voeren
- os9 hechten wij er waarde aan om bekend te staan als innovatieve school
- os10 verwachten wij van iedereen een flexibele houding ten opzichte van veranderingen

²⁴ Nieuwe schaal Prestatiegerichtheid (op basis van factoranalyse)

²⁵ Nieuwe schaal Ontwikkelingsgerichtheid (op basis van factoranalyse)

Bijlage 4C **Schoolorganisatiepraktijk: nieuw model**

26 items (alpha .89)

Bij ons op school...

Rationele Doel Model, 7 items (alpha .74) > Prestatiegerichtheid

- rd2 staan hoge leerprestaties centraal in de doelstellingen
- rd3 vormen de resultaten van de leerlingen het uitgangspunt bij de ontwikkeling van het onderwijsbeleid
- rd4 wordt aan de leerlingen duidelijk gemaakt dat hoge prestaties verwacht worden
- rd5 worden de resultaten van leerlingen systematisch geanalyseerd
- ip1 ligt in de onderwijsvisie de nadruk op een ordelijk en veilig leerklimaat
- ip2 zijn er duidelijke procedures om absentie van leerlingen tegen te gaan
- ip3 wordt erop toegezien dat leerlingen zich aan de gedragsregels houden

Interne Proces Model, 8 items (alpha .80) > Ontwikkelingsgerichtheid

- rd8 worden de onderwijsdoelstellingen als uitgangspunt gebruikt bij de besluitvorming
- rd9 worden beleidsvoorstellen voorzien van een gedetailleerde planning
- ip6 zijn er procedures om ervoor te zorgen dat iedereen tijdig van de juiste informatie wordt voorzien
- hr7 zijn de docenten actief betrokken bij de besluitvorming
- os7 is er beleid om de toepassing van nieuwe technologie in het onderwijs te bevorderen
- os8 worden docenten ondersteund bij het invoeren van onderwijsvernieuwingen dooros8
- os9 is het werken aan onderwijsvernieuwing in de organisatie ingebed
- os10 worden docenten in de gelegenheid gesteld om initiatieven te nemen

Human Relations Model, 7 items (alpha .78) > Ontwikkelingsgerichtheid

- ip7 zijn er indicatoren vastgesteld om te meten of de doelen worden bereikt
- ip10 zijn er procedures, planningen en instrumenten voor zelfevaluatie
- hr1 ligt in het personeelsbeleid de nadruk op competentieontwikkeling
- hr2 is er een uitgewerkt nascholingsplan
- hr3 houden docenten een persoonlijk scholings- of ontwikkelingsplan bij
- hr9 maken docenten gebruik van intervisie om van elkaar te leren
- os4 vindt systematisch overleg plaats met vervolgopleidingen over de aansluiting

Open Systeem Model, 4 items (alpha .71)

- os1 worden marketingactiviteiten ondernomen om de concurrentiepositie te verbeteren
- os2 zijn er richtlijnen voor een professionele aanpak van de externe communicatie
- os3 worden activiteiten ondernomen om het imago van de school te onderzoeken en verbeteren
- os5 is er beleid om de ouders zoveel mogelijk bij het onderwijs te betrekken

Bijlage 4D

Schoolbeleving van leerlingen: nieuw model

37 items (alpha .89)

Waardering van de school (alpha .83)

Omgang met medeleerlingen

- llomln1 Op deze school kun je echte vrienden maken
- llomln2 Ik kan mijn klasgenoten vertrouwen
- llomln3 Ik heb voldoende vrienden op school
- llomln4 Ik kan goed opschieten met de meeste andere leerlingen die ik ken op school
- llomln5 Ik kan op mijn klasgenoten rekenen als ik iets niet begrijp

Waardering van de school

- llwasch1 Ik ga graag naar school
- llwasch2 Ik vind het meestal leuk op school
- llwasch3 Ik vind het prettig om de pauze en de tussenuren op school door te brengen
- llwasch4 De sfeer bij ons op school is goed
- llwasch5 Ik heb het gevoel dat ik thuishoor op deze school

Waardering van onderwijs (alpha .70)

Waardering van onderwijs

- llwaond1 Iedereen zou zoveel mogelijk onderwijs moeten krijgen
- llwaond2 Ik vind mijn schoolwerk echt belangrijk
- llwaond3 Ik vind het heel belangrijk om op school kennis op te doen en vaardigheden te ontwikkelen
- llwaond4 Ik doe mijn best om goede cijfers te halen
- llwaond5 De dingen die ik op school leer zijn nuttig voor later

Family Educational Culture – Onderwijsondersteuning van thuis

- llstthu1 Mijn ouders/verzorgers zijn altijd bereid om mij te helpen met mijn huiswerk
- llstthu2 Ik heb thuis een rustige plek om mijn huiswerk te doen
- llstthu3 Mijn ouders/verzorgers letten erop dat ik mijn huiswerk doe
- llstthu4 Ik bespreek mijn schoolwerk vaak met mijn ouders/verzorgers
- llstthu5 Mijn ouders/verzorgers gaan meestal naar ouderavonden en andere speciale gebeurtenissen op school

Werkklimaat (alpha .87)

Omgang met leraren

- llomler1 Ik kan goed opschieten met de meeste leraren
- llomler2 De meeste leraren behandelen mij met respect
- llomler3 De meeste leraren zijn in mij als persoon geïnteresseerd
- llomler4 De meeste leraren lijken mij te begrijpen
- llomler5 De meeste leraren ervoor dat ik mij op mijn gemak voel in de klas

Ondersteuning door leraren

- llstler1 De meeste leraren moedigen mij aan als ik iets goed doe in de klas
- llstler2 Van de meeste leraren krijg ik de gelegenheid om mijn mening te geven in de klas
- llstler3 De meeste leraren praten regelmatig met mij over mijn schoolwerk
- llstler4 De meeste leraren praten doen erg hun best om leerlingen te helpen

Klassenklimaat

- llwklim1 Tijdens de lessen is er meestal een rustige werksfeer
- llwklim2 De meeste leraren hebben duidelijke regels over wat wel of niet kan tijdens de lessen
- llwklim3 Tijdens de meeste lessen wordt de tijd goed besteed

Organisatie van het onderwijs

- llorond1 Voor de meeste vakken zijn er duidelijke leerdoelen
(welke kennis en vaardigheden je moet leren)
- llorond2 De meeste lessen hebben een goede structuur
- llorond3 In de meeste vakken wordt op een goede manier getoetst

Invulling van de lessen

- llinvol1 De meeste leraren zorgen voor afwisselende activiteiten tijdens de lessen
- llinvol2 De meeste leraren proberen in hun lessen aan te sluiten bij wat leerlingen weten en beleven

Bijlage 5A Parameterschattingen schoolcultuur (nieuw model)

Volledig gestandaardiseerde parameterschattingen voor het nieuwe meetmodel van schoolcultuur, bij éénniveau confirmatieve factoranalyse aan de hand van docentendata (n=972) en meerniveau confirmatieve factoranalyse met 'Muthén's approximate Maximum Likelihood solution' (MUML), aan de hand van docentendata (N=100, n = 997).

item	éénniveau		meerniveau MUML			
	css	R ²	Between		Within	
	css	R ²	css	R ²	css	R ²
Rationele doel						
rd1	.64	.41	.74	.56	.60	.36
rd2	.71	.51	.57	.32	.77	.60
rd3	.56	.31	.89	.78	.48	.23
rd4	.69	.47	.95	.91	.63	.39
rd5	.55	.30	.81	.66	.46	.21
ip1	.63	.40	.75	.56	.63	.40
ip2	.68	.46	.93	.86	.65	.42
Interne proces						
ip3	.62	.38	.87	.75	.60	.36
ip4	.55	.30	.83	.69	.51	.26
ip5	.57	.32	.70	.49	.54	.29
ip6	.55	.30	.81	.65	.53	.28
ip7	.63	.40	.69	.48	.63	.40
ip10	.64	.41	.39	.15	.68	.44
Human relations						
ip9	.68	.45	.85	.73	.66	.46
hr1	.63	.39	.72	.52	.62	.39
hr2	.65	.42	.85	.73	.62	.38
hr4	.72	.51	.86	.73	.70	.49
hr5	.66	.44	.78	.60	.65	.42
hr6	.57	.33	.72	.52	.53	.28
hr7	.67	.45	.92	.84	.62	.38
hr9	.66	.44	.96	.92	.62	.38
hr10	.58	.34	.57	.33	.59	.35
Open systeem						
os1	.56	.32	.56	.31	.58	.34
os3	.43	.19	.45	.20	.46	.21
os6	.61	.38	.80	.63	.58	.34
os7	.65	.42	.85	.72	.59	.35
os8	.63	.40	.79	.63	.60	.36
os9	.57	.33	.89	.79	.51	.26
os10	.59	.35	.98	.96	.51	.26

Css = factorlading; volledig gestandaardiseerde oplossing;

Schuingedrukte factorladingen zijn niet significant (kritieke ratio < 1.96, p < .05);

R² = squared multiple correlation = percentage verklaarde variantie in geobserveerde variabele;

MUML: meerniveau analyse; 'between' = between groups, 'within' = within groups;

Bijlage 5B**Parameterschattingen schoolorganisatie (nieuw model)**

Volledig gestandaardiseerde parameterschattingen voor het nieuwe meetmodel van schoolorganisatie, bij éénniveau confirmatieve factoranalyse aan de hand van docentendata (n=955) en meerniveau confirmatieve factoranalyse met 'Muthén's approximate Maximum Likelihood solution' (MUML), aan de hand van docentendata (N=100, n = 997).

item	éénniveau		meerniveau MUML			
	css	R ²	between		within	
	css	R ²	css	R ²	css	R ²
Rationele doel						
rd2	.51	.26	.71	.51	.47	.22
rd3	.51	.26	.79	.62	.50	.25
rd4	.58	.34	.69	.48	.57	.32
rd5	.55	.30	.56	.31	.54	.30
ip1	.55	.30	.76	.57	.49	.24
ip2	.51	.26	.60	.35	.51	.26
ip3	.58	.34	.79	.63	.53	.28
Interne proces						
rd8	.61	.37	.85	.72	.57	.32
rd9	.58	.33	.74	.55	.54	.29
ip6	.56	.32	.76	.57	.54	.29
hr7	.55	.31	.61	.38	.54	.29
os7	.45	.20	.50	.25	.43	.18
os8	.64	.41	.82	.67	.62	.38
os9	.65	.42	.86	.74	.60	.36
os10	.52	.27	.78	.61	.47	.22
Human relations						
ip7	.73	.54	.79	.62	.72	.51
ip10	.68	.49	.86	.74	.61	.38
hr1	.54	.29	.85	.73	.45	.21
hr2	.58	.34	.75	.57	.56	.31
hr3	.54	.29	.63	.39	.53	.28
hr9	.58	.34	.72	.52	.54	.29
os4	.42	.18	.49	.24	.42	.18
Open systeem						
os1	.56	.32	.81	.66	.54	.29
os2	.51	.26	.33	.11	.53	.28
os3	.64	.41	.86	.74	.60	.36
os5	.49	.24	.50	.24	.45	.20

Css = factorlading: volledig gestandaardiseerde oplossing;

Schuingedrukte factorladingen zijn niet significant (kritieke ratio < 1.96, p < .05);

R² = squared multiple correlation = percentage verklaarde variantie in geobserveerde variabele;

MUML: meerniveau analyse; 'between' = between groups, 'within' = within groups;

Bijlage 5C

Parameterschattingen schoolbeleving (nieuw model)

Volledig gestandaardiseerde parameterschattingen voor het nieuwe meetmodel van schoolbeleving bij (éénniveau) confirmatieve factoranalyse aan de hand van leerlinggegevens (n=3760).

Item	css	R ²
Waardering van de school LLWASCH		
<i>Omgang met medeleerlingen</i>		
llomln1	.60	.36
llomln2	.48	.23
llomln3	.42	.18
llomln4	.48	.23
llomln5	.48	.23
<i>Waardering van de school</i>		
llwasch1	.69	.47
llwasch2	.73	.53
llwasch3	.51	.26
llwasch4	.60	.36
llwasch5	.70	.49
Waardering van Onderwijs LLWAOND		
<i>Waardering van Onderwijs</i>		
llwaond1	.45	.20
llwaond2	.68	.46
llwaond3	.56	.31
llwaond4	.49	.24
llwaond5	.45	.20
<i>Family Educational Culture</i>		
llstthu1	.38	.15
llstthu2	.32	.10
llstthu3	.36	.13
llstthu4	.45	.20
llstthu5	.29	.08
Werkklimaat LLWKLIM2		
<i>Omgang met leraren</i>		
llomler1	.59	.34
llomler2	.64	.41
llomler3	.55	.30
llomler4	.62	.39
llomler5	.66	.43
<i>Ondersteuning door leraren</i>		
llstler1	.46	.22
llstler2	.55	.30
llstler3	.40	.16
llstler4	.61	.37
<i>Klassenklimaat</i>		
llwklim1	.42	.18
llwklim2	.41	.17
llwklim3	.53	.29
<i>Organisatie van het onderwijs</i>		
llorond1	.50	.25
llorond2	.61	.37
llorond3	.50	.25
<i>Invulling van de lessen</i>		
llinvul1	.52	.27
llinvul2	.50	.25

Css = factorlading: volledig gestandaardiseerde oplossing;
R² = squared multiple correlation = percentage verklaarde
variantie in geobserveerde variabele

