

Diversiteit in de morele oordelen van scholieren: een onderzoek naar Kohlbergs structurele-eenheidspostulaat

L. DE MEY*

Vrije Universiteit Amsterdam

Samenvatting

Het structurele-eenheidspostulaat is een belangrijk fundament in Kohlbergs theorie van het morele oordelen. Het betreft hier zijn opvatting dat individuen in het algemeen consistent zijn in hun morele oordeelsniveaus. Maar hier is nogal wat kritiek op uitgeoefend, die voornamelijk van beschouwende aard is. In dit artikel worden de resultaten gerapporteerd van een empirisch onderzoek dat bedoeld is om een bijdrage te leveren aan de discussie over de houdbaarheid van het structurele-eenheidspostulaat. Onderzocht zijn leerlingen uit het Middelbaar Economisch en Administratief Onderwijs (MEAO-ers, n = 172, gemiddelde leeftijd 17 jaar) en leerlingen van scholen voor Zeer Moeilijk Opvoedbare Kinderen (ZMOK-ers, n = 64, gemiddelde leeftijd 16 jaar). Zij scoorden op vraagstukken uit Kohlbergs Heinz- en Belofstedilemma en nog een zogenaamd Zorgonderdeel. Om vast te kunnen stellen of er sprake is van consistentie van het niveau in de morele oordelen zijn multivariate variantie analyses (MANOVA, SPSS/PC+) en factoranalyses verricht. De uitkomsten wijzen uit dat de niveaus van morele oordelen van MEAO-ers niet consistent zijn, terwijl ZMOK-ers op bepaalde morele vraagstukken wel consistent scoren. In de discussie wordt hiervoor een interpretatie aangedragen.

1 Inleiding

Het onderzoeksprogramma dat Kohlberg ruim dertig jaar geleden startte, heeft een enorm aantal studies over de morele ontwikkeling van het individu opgeleverd. Centraal hierin staat Kohlbergs theorie van het morele oordelen, inhoudende dat het morele denken van individuen zich in zes stadia ontwikkelt (Kohlberg, 1981). De stadia zijn over drie niveaus van moreel redeneren verdeeld. Hoe hoger het stadium, hoe verder de morele ontwikkeling is gevorderd.

Moreel oordelen, getypeerd door een oriëntatie op straf en gehoorzaamheid, vermijding van straf en andere onaangename gevolgen, wordt aangeduid als stadium 1. In het tweede stadium kenmerkt het moreel oordelen zich door individualisme en instrumentele gerichtheid; menselijke relaties worden vooral gezien als ruilrelaties. Deze eerste twee stadia zijn subniveaus van het preconventioneel niveau. Moreel oordelen in het derde stadium betekent rekening houden met de gevoelens van anderen en de verwachtingen van de groep. In het vierde stadium is sprake van een gerichtheid op de handhaving van de sociale orde en het welzijn van de samenleving door gehoorzaamheid en loyaliteit aan de wet. Het derde en vierde stadium zijn samen subniveaus van het conventioneel niveau. In het vijfde en zesde stadium laat men zich in het moreel oordelen leiden door universele ethische principes. Deze zijn subniveaus van het postconventioneel niveau.

Een zeer belangrijke assumptie bij deze stadiumontwikkeling is het structurele-eenheidspostulaat. Colby en Kohlberg (1987, p. 6) omschrijven het postulaat als volgt: "A given stage-response on a task does not represent simply a specific response determined by knowledge and familiarity with that task or tasks similar to it; rather it represents an underlying thought-organization. The implication is that various aspects of stage structures should appear as a consistent cluster of responses in development". Het is duidelijk, dat Kohlberg en zijn collega's (Colby, Kohl-

* Met dank aan J.W. Steutel, G. Vreeke, J. Sturm, A. van der Leij en H. Ader voor hun commentaar en waardevolle adviezen.

berg, Gibbs & Liebermann, 1983) dit postulaat als een onderdeel van de 'hard core' van de theorie zien. Kohlbergs theorie, dat morele ontwikkeling zich volgens stadia voltrekt zou ernstig bedreigd worden, als aangetoond wordt dat er geen duidelijk van elkaar te onderscheiden stadia zijn, of als zou blijken, dat individuen niet in één bepaald stadium van morele ontwikkeling zijn in te delen¹. De Kohlberggroep (Colby et al., 1983) weet het postulaat met empirische onderzoeksresultaten te onderbouwen. Deze zijn:

1. over negen onafhankelijk gescoorde dilemma's valt tweederde van alle scores binnen het modale stadium van het subject, de overblijvende scores zijn toegekend aan een stadium grenzend aan het modale stadium²;
2. hoge correlaties tussen de scores op verschillende morele problemen, hetgeen op een hoge graad van interne consistentie wijst;
3. een factoroplossing, die een één-factorstructuur laat zien; deze verklaart een groot deel van de variantie over de verschillende morele dilemma's, te weten 46%; hiermee toont de Kohlberggroep aan dat "we have succeeded in defining a coherent moral domain united by a single underlying organizational structure" (Colby et al., 1983, p. 72).

Gelet op deze onderzoeksresultaten, hebben de onderzoekers kennelijk twee interpretaties van structurele eenheid: 1) het eerste resultaat betekent dat Kohlberg structurele eenheid tot uiting ziet komen in consistente scores in antwoorden op morele vraagstukken. Dit betekent bijvoorbeeld dat, als een individu op een vraagstuk uit het Heinzdilemma (zie 2.2) stadium vier scoort, dan scoort het individu in het algemeen op niveau vier, 2) het tweede en derde resultaat betekenen dat morele redeneringen voortkomen uit één moreel gebied.

Sommige onderzoekers wijzen op de mogelijkheid dat oordeelsniveaus met betrekking tot morele vraagstukken mede bepaald worden door situationele contexten en persoonlijke ervaringen en derhalve per moreel vraagstuk kunnen variëren. Zo brengt Van IJzendoorn (1981, p. 72), refererend aan het Heinz- en Beloftedilemma (zie 2.2), het volgende naar voren: "Het is niet onmogelijk dat een moreel dilemma over een doodzieke vrouw minder relevante persoonlijke ervaringen oproept dan een moreel dilemma over de verplichting zich te houden aan een eenmaal

gegeven belofte". Ook De Mey, Van der Draai en Spiecker (1989) vermoeden iets dergelijks. Zij vergelijken de totaalscores op twee verschillende soorten dilemma's, te weten het Heinzdilemma en het zogenaamde Zorgdilemma (zie 2.2). Op het Heinzdilemma, waarin het gaat om hypothetische conflicten tussen rechten, valt de totaalscore significant lager uit dan op het Zorgdilemma, waarin meer levenschte, relatiegerichte vraagstukken onderwerp zijn. De onderzoekers denken dat in opvoedingssituaties ouders meer aandacht vragen voor relatiegerichte problemen in zorg-situaties, dan voor problemen rond leven en dood of loyaliteit aan de wet. Een ander voorbeeld waarin individuen simultaan verschillende morele-oordeelstypen hanteren, is te vinden bij Sanvitale, Saltztein en Fish (1989). Deze onderzoekers tonen aan dat 14-jarige leerlingen onderscheid maken in hun scores op morele vraagstukken, wanneer deze variëren qua situatie. De morele dilemma's, weergegeven in een aantal korte verhalen op video of tape, zijn verdeeld in verhalen over twee personen die bepaalde activiteiten verrichten, waarbij persoon A persoon B dupeert. Maar in het ene geval kan A de gevolgen van zijn daden (niet bedoeld) wel voorzien en in het andere geval niet. De factor 'persoonlijke ervaringen' onderzoeken deze auteurs, door de scores van twee groepen scholieren, te weten 'normale' leerlingen en 'agressieve' leerlingen, met elkaar te vergelijken. 'Normale' leerlingen geven in hun morele oordelen een sterker contrast te zien tussen de zogenaamde 'te voorzien' - en 'niet te voorzien' - verhaaltjes, dan 'agressieve' leerlingen.

Naar aanleiding van deze vermoedens en observaties, luidt de probleemstelling voor dit artikel als volgt: *in hoeverre kenmerken de morele oordelen van scholieren zich door structurele eenheid? Wijken in dit opzicht scholieren, bij wie een normale sociaal-emotionele ontwikkeling verondersteld mag worden af van scholieren bij wie dit niet het geval is?*

Dit probleem is bestudeerd bij leerlingen uit het Middelbaar Economisch en Administratief Onderwijs (voortaan MEAO-ers) en leerlingen van scholen voor Zeer Moeilijk Opvoedbare Kinderen (voortaan ZMOK-ers). Aangenomen is dat de MEAO-ers zich in sociaal-emotioneel opzicht normaal ontwikkelen, terwijl voor de ZMOK-ers deze ontwik-

keling problematisch verloopt. Deze groepen oordelen in moreel problematische situaties op zeer uiteenlopende wijze (De Mey et al., 1989), hetgeen doet vermoeden dat ze niet even consistent in hun morele oordeelsniveau zijn. De observaties van Sanvitale en zijn collega's (Sanvitale et al., 1989) wijzen ook in deze richting. Kohlberg zelf deelt in dit opzicht niets mee. Hij baseert zich in zijn uitspraken alleen op waarnemingen van 'normale' kinderen. De scholieren scoorden op vraagstukken die betrekking hebben op drie morele dilemma's: twee bekende Kohlberg dilemma's, het Heinzdilemma en het Belofte-dilemma en een zogenaamd Zorgonderdeel (De Mey et al., 1989).

Eerst is nagegaan of er in de morele oordelen van scholieren, ongeacht de morele dilemma's waarin ze thuishoren, structurele eenheid is. De onderwerpen lopen namelijk per dilemma sterk uiteen. Verwacht wordt dan ook dat de morele oordelen van subjecten divergeren. Meer specifiek is vervolgens bestudeerd of er in de morele oordelen van het Heinzdilemma, het Belofte-dilemma en het zogenaamde Zorgonderdeel structurele eenheid is. Dit kan interessant zijn, omdat de vraagstukken per dilemma slechts gradueel in situatie variëren. Het ligt nu wel voor de hand structurele eenheid in de morele oordelen van elk dilemma te verwachten. Voorts is met betrekking tot structurele eenheid nagegaan op welke wijzen ZMOK-ers van MEAO-ers verschillen en overeenkomen. Hierbij zijn een aantal overeenkomstige verschillen tussen vraagstukken inhoudelijk bestudeerd. De vergelijkingsresultaten zijn mogelijk ondersteunend voor de observaties van Sanvitale et al. (1989) en leiden tot de formulering van enige hypothesen.

2 Methode

2.1 Onderzochte groepen

Aan het onderzoek namen drie groepen scholieren deel: groep 1 bestaande uit leerlingen van scholen voor Zeer Moeilijk Opvoedbare Kinderen en de groepen 2 en 3 met leerlingen afkomstig uit het Middelbaar Economisch en Administratief Onderwijs.

Groep 1 bestaat uit leerlingen ($n = 64$) van ZMOK-scholen uit Amsterdam, Haarlem, Utrecht en Rotterdam, met een gemiddelde

leeftijd van 16 jaar, minimum leeftijd 14, maximum leeftijd 18 jaar. Deze groep wijkt af van de andere groepen, waarbij een normale sociaal-emotionele ontwikkeling verondersteld wordt. Zeer moeilijk opvoedbare kinderen zijn doorgaans affectief verwaarloosd. Hun gevoelsleven is onvoldoende of niet 'gewekt' door een ouder of verzorger (Vedder, 1977). In moreel opzicht zijn deze kinderen aan hun lot overgelaten, doordat hun onvoldoende of geen normen zijn bijgebracht. Het zeer moeilijk opvoedbare kind onderscheidt zich dan ook voornamelijk van andere kinderen, niet in de sfeer van de intellectuele mogelijkheden, maar primair door een persoonlijkheids-ontwikkeling die de omgang met anderen in ernstige mate bemoeilijkt (Dumont, 1985). Dumont omschrijft het zeer moeilijk opvoedbare kind als volgt: "Kenmerkend voor de persoonlijkheids-ontwikkeling van deze kinderen is de uiterst egocentrische instelling. Omdat zij geen band met anderen kunnen krijgen, houden zij met andere kinderen geen rekening. . ." (Dumont, 1976, p. 45).

Groep 2 ($n = 83$) bestaat uit leerlingen³ uit de eerste klas van een MEAO in Hoorn. Hun gemiddelde leeftijd was 17, minimumleeftijd 16, maximumleeftijd 20 jaar (MEAO2).

Groep 3 ($n = 89$) is op dezelfde wijze samengesteld als groep 2. De gemiddelde leeftijd van deze groep is, evenals die van groep 2, 17 jaar, minimumleeftijd 16, maximumleeftijd 20 jaar (MEAO3).

De gemiddelde leeftijd van de MEAO-ers is hoger uitgevallen dan die van de ZMOK-ers. Een leeftijdseffect kan hier het gevolg van zijn. Om dit te controleren, is een groep MEAO-ers ($n = 35$) samengesteld met een gemiddelde leeftijd van 16 jaar. Deze groep heeft vraagstukken uit het Heinzdilemma en het Zorgonderdeel gescoord. Vergelijking met MEAO3 die een gemiddelde leeftijd van 17 jaar had, resulteerde niet in significante verschillen. Daar er geen leeftijdseffect kon worden vastgesteld is het onderzoek voortgezet met de grotere groepen MEAO2 en MEAO3.

2.2 Instrumentarium

Omschrijving van de dilemma's

Het gebruikte instrumentarium is rond de voornoemde drie dilemma's opgebouwd. Het Heinz- en Belofte-dilemma met bijbehorende

vraagstukken vormen de Sociomoral Reflection Objective Measure (S.R.O.M.). Het Heinzdilemma samen met het Zorgonderdeel vormen hier een variant op (De Mey et al., 1989). De S.R.O.M. is ontwikkeld uit Kohlbergs M.J.I. (Moral Judgement Interview) en de S.R.M. (Sociomoral Reflection Measure) (Gibbs, Arnold, Morgan, Schwartz, Gavaghan & Tappan, 1984). Er is hier met een vertaalde versie, afkomstig van Van IJzendoorn, gewerkt.

In het *Heinzdilemma* verkeert Heinz in de situatie, dat zijn vrouw komt te sterven als zij niet geholpen wordt met een net ontdekt medicijn. Heinz heeft niet genoeg geld voor het medicijn en de apotheker, die de ontdekking gedaan heeft, weigert het medicijn goedkoop aan Heinz te verstrekken. Wat moet Heinz doen? De wet overtreden door het medicijn te stelen, waarmee z'n vrouw gered wordt, of het naderende sterven van z'n vrouw aan te zien en daarmee de wet te respecteren?

In het *Beloftedilemma* is sprake van een situatie waarin een vader zijn zoon belooft op kamp te mogen gaan. Voorwaarde is dat de 14-jarige Jan hiervoor zelf het geld verdient. Deze neemt een baantje en spaart de benodigde honderd gulden bij elkaar. Maar vader verandert van gedachten. Hij kan mee met een vistocht, maar heeft niet genoeg geld. Dus eist hij de honderd gulden van zijn zoon op. Wat moet Jan doen? Niet gehoorzamen en op kamp gaan, of zijn vader helpen en afzien van het beloofde kamp?

In het *Zorg-onderdeel* is rekening gehouden met de kritiek van Gilligan op Kohlberg (Gilligan, 1977). Kohlbergs nadruk op rechtvaardigheid en het gebruik van hypothetische dilemma's zijn in dit verband de belangrijkste kritiekpunten. De vraagstukken in dit onderdeel gaan over situaties, waarin de respondent de relatie met anderen geproblematiseerd ziet. De situaties geven een discrepantie aan tussen het belang van de respondent en de relatie met anderen, en doen een beroep op de empathische rolname, zorg en verantwoordelijkheid van de respondent. De onderwerpen die in de situaties aan de orde komen, zijn toegesneden op de belevingswereld van scholieren.

Omschrijving van de antwoordmogelijkheden

Op het Heinz-, Belofte- en Zorgonderdeel wordt met een aantal vraagstukken gevarieerd

(Heinz 10, Belofte 6, Zorg 6). Ieder vraagstuk heeft zes antwoordmogelijkheden. Vijf ervan vertegenwoordigen de stadia van moreel oordelen, een zesde antwoordmogelijkheid is een zogenaamd pseudo-, een 'onzin'-alternatief dat een controlefunctie heeft voor het wel of niet serieus beantwoorden van de vragen door de respondent (zie ook procedure). Achtereenvolgens volgt ter verduidelijking een voorbeeld uit het Heinzdilemma en uit het Zorgonderdeel.

De respondent krijgt het vraagstuk voorgelegd, waarin de stervende persoon niet Heinz' vrouw is maar een vriend. Moet hij het medicijn stelen om z'n vriend te redden of zou hij het niet moeten stelen? De respondent moet van een aantal voorgegeven redenen aangeven of deze in de buurt komen van de redenen die hij/zij zou geven: a) omdat je vriend misschien dingen voor jou heeft gedaan zou je hem een dienst moeten bewijzen, zeker als je wilt dat je vriend je in de toekomst ook helpt, b) omdat een vriendschap gebaseerd moet zijn op wederzijds respect en samenwerking, c) omdat het je vriend is die een beroemd persoon kan zijn, d) omdat je je betrokken zou voelen bij je vriend en zou verwachten dat je vriend jou ook zou helpen, e) omdat jij en je vriend misschien volledig aan elkaar toegewijd zouden zijn geraakt, f) (*pseudo*) omdat een eerste vereiste van genegenheid een relatie is.

Uit het Zorgonderdeel het volgende voorbeeld. Beschreven wordt, dat de respondent met zijn/haar vriend(in) samen op vakantie gaat. Een paar weken voor de vakantie krijgt de respondent een uitnodiging van een klasgenoot om voor heel weinig geld een trip naar Amerika te maken. Deze buitenkans is zeer aantrekkelijk. Toch besluit de respondent om niet op het aanbod in te gaan, maar met de vriend(in) op vakantie te gaan. Welke redenen worden gekozen: a) omdat je dan na de vakantie zonder vriend(in) zit en je moet ook maar afwachten hoe het je op reis met die klasgenoot zou bevalen, b) (*pseudo*) omdat zonder het individu er geen verplichtingen in vriendschap kunnen bestaan, c) omdat je je beloften moet nakomen, d) omdat vertrouwen een van de belangrijkste waarden in onze samenleving is, e) omdat je vriend(in) anders waarschijnlijk kwaad zou worden, f) omdat je je kan voorstellen hoe rot je vriend(in) zich zou voelen; bovendien horen mensen zo niet met elkaar om te gaan.

Procedure

De verdeling van de morele dilemma's over de groepen scholieren gebeurde als volgt. Het Heinzdilemma is aan alle groepen aangeboden, het Zorgonderdeel aan de ZMOK-ers en MEAO3 en het Belofte-dilemma aan MEAO2. Redenen voor deze verdeling zijn de volgende. Het beantwoorden van alle vragen over de drie dilemma's door één groep scholieren was praktisch onmogelijk. Er stond maar één lesuur ter beschikking, terwijl het scoren op twee dilemma's precies in één lesuur mogelijk was. Toch werd inzicht in de scorepatronen op alle dilemma's belangrijk geacht. Om praktische redenen lag het voor de hand twee groepen MEAO samen te stellen. Gebruik kon worden gemaakt van reeds bestaande contacten met een MEAO-school. De gegevens zijn verzameld met behulp van klassikaal afgenomen vragenlijsten. De afname duurde ongeveer vijftig minuten. Voor iedere groep is onderzocht in hoeverre de verzamelde protocollen wel of niet voor verdere analyse in aanmerking kwamen. Een protocol met zeven of meer pseudo's deed niet mee, omdat dan aangenomen werd dat de rest van de ingevulde vragen onbetrouwbaar was (Gibbs et al., 1984).

Scoring en analyse

Een respondent kon uit de voorgegeven antwoordmogelijkheden, die elk een moreel stadium (1 tot 5) vertegenwoordigen in de theorie van Kohlberg, er meerdere kiezen. Voor iedere respondent is per vraagstuk de gemiddelde score berekend (Gibbs et al., 1984). Vervolgens is per groep voor ieder vraagstuk de gemiddelde score berekend.

De scores zijn eerst geanalyseerd met een multivariate variantieanalyse (MANOVA, SPSS/PC+) op de gelijkheid van gemiddelden, tweezijdig op 5% niveau. Daarna zijn factoranalyses verricht.

Metrische eigenschappen

Genoemd zijn drie morele dilemma's, die deel uitmaken van de S.R.O.M. en een door De Mey et al. (1989) ontwikkelde variant. De rapportages over de metrische eigenschappen van de S.R.O.M. zijn gunstig (Gibbs et al., 1984; Van IJzendoorn, 1986). De betrouwbaarheid van de S.R.O.M., nagegaan voor dit onderzoek, resulteerde in een alpha-coëfficiënt van .71. De betrouwbaarheid van de variant leverde een coëfficiënt van .85 op. Voor de Heinz-

Belofte- en Zorg-items zijn respectievelijk .79, .66, .75 berekend. Deze uitkomsten werden voldoende bevonden om verdere analyses te plegen.

3 Resultaten

Om de structurele eenheid in morele oordelen, ongeacht het dilemma te onderzoeken is eerst een simultane toets op de gemiddelde morele oordelen verricht. De uitkomsten staan in Tabel 1.

De toetsingswaarden van Hotellings T-kwadraat en de benaderende F-waarden geven aan dat de gemiddelde scores sterk significant van elkaar verschillen. Uitgaande van Kohlbergs eerste interpretatie (zie de inleiding) betekent dit dat de onderzochte scholieren in het algemeen over ieder moreel vraagstuk anders oordelen. Maar hiermee is nog niet gezegd dat de tweede interpretatie (zie de inleiding) van structurele eenheid: morele oordelen vormen één moreel domein, onhoudbaar is. De correlaties en factoranalytische oplossingen, die in dit verband door de Kohlberggroep gerapporteerd zijn, laten immers systematische scoreverschillen toe. Om ook deze interpretatie te onderzoeken zijn de scores, in navolging van de Kohlberggroep, aan een principale componentenanalyse onderworpen. De analyse op de scores van het grootst aantal mogelijke subjecten⁴ is geheel volgens de richtlijnen van de Kohlberggroep uitgevoerd. Factoren, die een eigenwaarde groter dan één hebben zijn geëxtraheerd⁵ en niet geroteerd. Ladingen boven de .40 zijn op 1% niveau als significant te beschouwen. De resultaten staan in Tabel 2.

Uit Tabel 2 blijkt dat 1) de meeste significante ladingen onder één factor vallen, 2) maar toch een aanzienlijk deel van de significante ladingen over de vier overige factoren varieert en 3) de door de eerste factor verklaarde variantie beduidend minder is dan de 46% die door de Kohlberggroep gerapporteerd is. Na een orthogonale rotatie werd dit factorprofiel alleen maar duidelijker. Daarom is vast te stellen dat in de morele oordelen van de hier onderzochte scholieren geen sprake is van één algemene factor.

De analysesresultaten samenvattend, is duidelijk geworden dat, wanneer geen onder-

Tabel 1 *Simultane toetsen (MANOVA) op de gemiddelde scores van vraagstukken over alle morele dilemma's: het Heinz-Belofte dilemma en het Zorgonderdeel. Tevens het gemiddelde (x) en de standaarddeviatie (sd) per vraagstuk*

moreel probleem	ZMOK (n = 64)	MEAO3 (n = 89)	MEAO2 (n = 83)	MEAO (n = 172)
Scores op Heinz				
	x sd	x sd	x sd	x sd
mb1	2.42 (1.49)	3.51 (1.37)	3.34 (1.59)	3.42 (1.45)
mb2	2.05 (1.63)	3.09 (1.11)	2.60 (1.58)	2.85 (1.37)
mb3	1.56 (1.64)	2.12 (1.50)	2.23 (1.53)	2.18 (1.51)
mb4	2.25 (1.63)	3.21 (1.24)	3.11 (1.30)	3.17 (1.26)
mb5	1.80 (1.50)	2.66 (1.33)	2.64 (1.43)	2.65 (1.38)
mb6	1.36 (1.50)	2.54 (1.39)	2.48 (1.40)	2.51 (1.39)
mb7	1.83 (1.58)	3.10 (1.07)	3.07 (1.06)	3.10 (1.06)
mb8	1.97 (1.61)	2.31 (1.37)	2.18 (1.45)	2.35 (1.41)
mb9	1.86 (1.53)	2.36 (1.47)	2.20 (1.54)	2.29 (1.50)
mb10	1.35 (1.47)	2.45 (1.55)	2.55 (1.52)	2.50 (1.53)
	Scores op Zorg		Scores op Belofte	
mb11	1.34 (1.66)	2.16 (1.99)	2.14 (1.41)	
mb12	1.78 (1.48)	2.90 (1.24)	2.29 (1.50)	
mb13	1.29 (1.70)	2.13 (1.80)	3.38 (1.46)	
mb14	2.02 (1.71)	2.85 (1.37)	2.95 (.96)	
mb15	1.72 (1.97)	3.09 (1.96)	2.40 (1.92)	
mb16	1.88 (1.47)	2.42 (1.17)	2.18 (1.67)	
Hotellings				
T-kwadr.	.85	1.86	1.32	.95
F-waarde	2.79	9.20	5.98	17.30
(df1, df2)	(15,49)	(15,74)	(15,68)	(9,163)
p-waarde	.004	.000	.000	.000

Tabel 2 *Factorstructuur verkregen door principale componenten analyse van scores uit het Heinz + Zorgonderdeel voor ZMOK en MEAO3 (n = 153), zonder rotatie*

Factor	1	2	3	4	5
Heinz					
mb1	.58	-.20	-.05	.10	-.16
mb2	.56	-.28	-.05	.09	.31
mb3	.18	-.54	.17	-.05	.20
mb4	.57	-.36	-.04	.10	-.38
mb5	.34	-.47	-.28	-.05	.09
mb6	.46	-.00	-.06	-.18	-.02
mb7	.49	.24	.36	-.15	-.43
mb8	.30	.41	-.47	-.13	-.11
mb9	.28	.26	.01	.76	-.13
mb10	.57	.09	-.11	-.01	.14
Zorg					
mb11	.43	.33	-.06	-.26	.57
mb12	.46	.24	-.14	-.37	-.26
mb13	.32	.53	.13	.38	.26
mb14	.43	.03	-.36	.05	.00
mb15	.46	-.26	.40	.15	.13
mb16	.33	.20	.65	-.27	.03
Eigenwaarden	3.07	1.60	1.24	1.13	1.04
Variantie (%)	19.2	10	7.7	7	6.5

Tabel 3 *Simultane toetsen (MANOVA, SPSS/PC+) op de gemiddelde scores van vraagstukken in het Heinz-dilemma, het Belofte-dilemma en het Zorgonderdeel*

	ZMOK (n = 64)		Heinz (n = 172)	MEO Zorg (n = 89)	Belofte (n = 83)
	Heinz	Zorg			
Hotellings T-kwadraat	.69	.20	.95	.31	.64
F-waarde	4.22	2.40	17.30	5.22	9.95
(df1, df2)	(9,55)	(5,59)	(9,163)	(5,84)	(5,78)
p-waarde	.000	.05	.000	.000	.000

Tabel 4 *Eigenwaarden (ew) van factoren en percentages verklaarde variantie (v) voor het Heinzdilemma, het Belofte-dilemma en het Zorgonderdeel*

Groepen en dilemma's	Factor 1		2		3		4	
	ew	v	ew	v	ew	v	ew	v
ZMOK (n = 64)								
Heinz	2.12	21%	1.70	17%	1.15	12%	1.13	11%
Zorg	1.70	28%	1.23	21%	—	—	—	—
MEO								
Heinz (n = 172)	1.42	14%	1.30	13%	1.25	13%	1.06	11%
Zorg (n = 89)	1.38	23%	1.16	19%	—	—	—	—
Belofte (n = 83)	1.37	23%	1.19	20%	1.10	18%	—	—

scheid gemaakt wordt tussen de morele dilemma's, er geen structurele eenheid is in de morele oordelen. Dit kan anders zijn indien elk dilemma afzonderlijk bestudeerd wordt. Nogmaals zijn simultane toetsen (MANOVA, SPSS/PC+) en principale componentenanalyses verricht. In Tabel 3 staan de resultaten van de simultane toetsen.

De uitkomsten in Tabel 3 geven met één uitzondering voor ZMOK-ers op het Zorgonderdeel, sterk significante verschillen aan tussen de gemiddelde scores.

Principale componenten analyses op de scores in elk dilemma leverden de in Tabel 4 vermelde resultaten op.

Tabel 4 laat zien dat in elk dilemma meerdere factoren met een eigenwaarde groter dan één zijn geëxtraheerd. Dit is voor zowel ZMOK-ers als MEO-ers het geval. Ook nu weer verklaart de eerste factor in elk dilemma beduidend minder dan de 46% van de Kohlberggroep.

Resumerend kan, in tegenstelling tot wat verwacht werd, gesteld worden dat in de morele oordelen van de meeste dilemma's geen

structurele eenheid te herkennen is, maar enige nuancering is mogelijk.

Bij vergelijking valt namelijk op, dat in de morele oordelen van ZMOK-ers in het algemeen geen structurele eenheid te bekennen is. Toch neigt deze groep tot meer eenheid in sommige van hun oordelen dan MEO-ers. In Tabel 3 is te zien dat op het Zorgonderdeel de Hotellings T-kwadraat en F-waarde voor ZMOK-ers niet significant zijn. Dit betekent dat er in het Zorgonderdeel geen significante verschillen zijn tussen de morele oordelen van de ZMOK-ers. Ook de resultaten van de principale componentenanalyses laten verschillen tussen beide groepen zien. Zowel op het Heinzdilemma als het Zorgonderdeel wordt bij ZMOK-ers door de factoren meer variantie verklaard dan voor de MEO-ers, in totaal respectievelijk 61% en 51%. Nu is het niet zo, dat de tweede interpretatie van structurele eenheid voor ZMOK-ers overeind blijft. Maar al met al lijkt de uitspraak gerechtvaardigd, dat de morele oordelen van ZMOK-ers naar meer structurele eenheid neigen dan de oordelen van MEO-ers. Terzijde zij opgemerkt dat er

overeenkomst is met de observaties van de eerder genoemde Sanvitale-groep (1988).

Tot nu toe is de vergelijking tussen ZMOK-ers en MEAO-ers vooral getalsmatig getrokken. Interessant is het ook een aantal verschillen tussen morele oordelen, die beide groepen gemeen hebben, inhoudelijk te presenteren. Hierdoor wordt begrijpelijk dat de beoordeling van morele problemen inderdaad situatie afhankelijk kan zijn. De problemen die geselecteerd zijn, hebben een wisselende affiniteit van de respondent met een persoon gemeen (zie Tabel 1).

Karakteristieken bij vraagstukken uit het Heinzdilemma

Opmerkelijk is dat zowel MEAO-ers (Hotellings T-kwadraat = .39, $F(2, 170) = 33.49$, $p = .000$) als ZMOK-ers (Hotellings T-kwadraat = .16, $F(2, 62) = 5.06$, $p = .009$) een verschil vertonen tussen de scores op mb1, mb2, mb3⁶, die respectievelijk de volgende vraagstukken vertegenwoordigen: 'waarom is het voor een echtgenoot belangrijk om te stellen (om het leven van zijn vrouw te redden)', 'waarom is het belangrijk voor jou alles te doen, zelfs het overtreden van de wet, om het leven van een vriend(in) te redden', 'waarom is het belangrijk voor jou alles te doen, zelfs het overtreden van de wet, om het leven van een vreemde te redden'. De variatie tussen deze vraagstukken voor de MEAO-ers, respectievelijk 3.42, 2.85, 2.18, laat zien dat het leven van een vreemde moreel minder belangrijk is. De morele oordeelsvorming verschuift van conventioneel naar preconventioneel en krijgt een instrumenteel karakter. Voor de ZMOK-ers is deze variatie 2.42, 2.05, 1.56. Zij komen niet boven het preconventionele niveau uit en laten een verschuiving in de richting van een straforiëntatie zien. Samengevat zijn deze bijzonderheden te formuleren in de volgende hypothese. *Hypothese 1*: Naarmate men minder affiniteit met een individu heeft, valt het morele oordeel over het recht op leven van dit individu lager uit.

Maar niet voor iedere morele waarde komt een gradatie in affiniteit op deze wijze in de morele oordelen tot uitdrukking. Hoewel in dit verband geen vergelijking met ZMOK-ers mogelijk is, is het in de voorgaande context toch de moeite waard, het volgende voorbeeld naar voren te brengen.

Karakteristieken bij vraagstukken uit het Belofte dilemma

Opvallend voor MEAO2 is het verschil ($F(1,82) = 23.53$, $p = .000$) tussen de scores op mb12 (2.29) en mb13 (3.38), respectievelijk 'waarom is het belangrijk je te houden aan een belofte, die je aan een vreemde hebt gedaan?' en 'waarom is het belangrijk je te houden aan een belofte aan een vriend? De belofte aan een vreemde wordt als moreel dwingender beoordeeld. Er vindt een verschuiving plaats van het conventionele naar het preconventionele niveau en het morele oordeel krijgt daarmee een instrumenteel karakter. Kennelijk kun je het met een vriend al gauw op een akkoordje gooien. Deze bevinding kan kortweg worden samengevat met de volgende hypothese. *Hypothese 2*: Het nakomen van verplichtingen in vriendschaps-relaties wordt vanuit een pragmatisch perspectief beoordeeld.

In het volgende voorbeeld komt tot uitdrukking dat verwantschap in vergelijking met een voor personen gevoelde affiniteit in moreel problematische situaties niet tot hogere oordelen leidt.

Karakteristieken bij vraagstukken uit het Zorgonderdeel

Het verschil tussen de scores op mb11 en mb14 voor ZMOK-ers (1.34 en 2.02) ($F(1,63) = 6.64$, $p = .012$) en MEAO3-ers (2.16 en 2.85) ($F(1,88) = 7.53$, $p = .007$) betreft de respectievelijke vraagstukken 'laten we zeggen dat je redenen moet geven om op een redelijke tijd naar huis te gaan uit de disco (als je ouders ongerust zijn)' en 'laten we zeggen dat je redenen moet geven waarom je niet met je klasgenoot op een vakantie gaat, wat een buitenkansje voor je is, maar met je beste vriend(in)'. Deze scoreverschillen zijn samen te vatten in de volgende hypothese. *Hypothese 3*: Gevoelens van zorg voor ouders geven aanleiding tot lagere morele oordelen dan gevoelens van zorg in vriendschapsrelaties.

4 Conclusies en discussie

Deze studie beoogt een bijdrage te leveren aan het geldigheidsvraagstuk met betrekking tot het structurele-eenheidspostulaat. De resultaten samenvattend is te constateren dat de morele oordeelsniveaus zeer divers gescoord zijn door de onderzochte scholieren. Niet alleen

geeft een algemene vergelijking tussen de scores op morele vraagstukken van verschillende morele dilemma's een grote diversiteit in oordeelsstadia te zien (hetgeen niet hoeft te verbazen, omdat de beschreven situaties sterk verschillen) maar ook de scores op oordeelsniveaus in één moreel dilemma zijn sterk uiteenlopend. Inhoudelijk gezien komen hierbij een aantal opvallende zaken naar voren. Een aantal significante scoreverschillen tussen de vraagstukken zijn door de verschillende groepen scholieren op dezelfde wijze geordend, tot uitdrukking komend in bijvoorbeeld de vraagstukken over het recht op leven in het Heinzdilemma. Hierbij valt op dat het recht op leven vanuit verschillende morele oordeelstructuren gescoord wordt als de situatie varieert. Wat misschien verwondert is, dat ZMOK-ers, die een beschadigde persoonlijke ontwikkeling hebben, de morele vraagstukken homogener beoordelen dan MEAO-ers. Een verklaring hiervoor kan gelegen zijn in het feit, dat ZMOK-ers op een lager moreel-redeneringsniveau functioneren dan de MEAO-ers. Dit zou erop kunnen duiden dat naarmate functioneren op een lager moreel-oordeelsniveau plaatsvindt, er minder morele echelons geactiveerd worden en dat derhalve de scores homogener zijn. Andersom, naarmate men op een hoger moreel redeneringsniveau functioneert kunnen er meer morele-argumentatieniveaus ingezet worden en zijn de scores bijgevolg diverser.

Enige terughoudendheid in de uitspraken over ZMOK-ers is misschien gepast. Tijdens de controle van de protocollen bleek dat ongeveer een derde van de tweeënegentig ondervraagde ZMOK-ers niet aan de gestelde scoringscriteria beantwoordde, hetgeen wil zeggen dat bij deze subjecten meer dan zeven pseudo's geteld werden. De scores van deze groep kwamen niet voor verdere analyse in aanmerking. Van de ondervraagde MEAO-ers waren dit er slechts vijf. Het is mogelijk dat hierdoor een selecte groep ZMOK-ers ($n = 64$) ontstaan is, die niet representatief is voor de doorsnee ZMOK-er. Mogelijk betekent dit ook dat juist deze geselecteerde groep ZMOK-ers homogeen scoort. Onderzoek naar het gegeven dat ZMOK-ers zoveel pseudo's scoren moet nog plaatsvinden. Het is mogelijk zinvol hierbij een vergelijking te maken met een groep delinquenten.

Niettemin kan worden gesteld, dat de bo-

vengenoemde resultaten het structurele-eenheidspostulaat ernstig bedreigen. Nochtans hoeft het niet direct te betekenen dat de theorie verlaten moet worden. Colby et al. (1983) en Colby en Kohlberg (1987) voeren een verdediging waarin zij er uitdrukkelijk op wijzen dat data, die het structurele-eenheidspostulaat steunen, verzameld moeten zijn met hun scoringssysteem: het Moral Judgment Interview (M.J.I.). Wanneer een ander instrument is gebruikt, (Colby en Kohlberg refereren dan aan de Defining Issues Test (D.I.T.) van Rest (1976)), kan dit, volgens hen, leiden tot afwijkende resultaten. Het is alleszins begrijpelijk dat zij deze verdedigingslijn volgen, omdat de D.I.T. inderdaad iets anders meet dan de M.J.I., hetgeen blijkt uit lage correlaties tussen beide instrumenten (Gibbs et al., 1984). Maar in deze studie zijn drie soorten morele dilemma's gemeten: het Heinz-, het Belofte- en het Zorg-dilemma, die wel overeenkomsten met de M.J.I. hebben. De eerste twee vormen de Sociomoral Reflection Objective Measure (S.R.O.M.), die bevredigende correlaties met de M.J.I. (Gibbs et al., 1984) vertoont. Het instrument waarmee het Zorg-onderdeel is onderzocht, correleert significant ($.59, p < .001$) met het Heinzdilemma. Toch ondersteunen de verkregen uitkomsten het structurele-eenheidspostulaat niet. Wat de door Colby en Kohlberg gehanteerde verdediging voorts afzwakt, is dat het scoringssysteem de toets der kritiek niet goed kan doorstaan. Hoewel het scoringssysteem verschillende keren gewijzigd is, blijft de wetenschappelijke toepassing twijfelachtig. Rest (1983, p. 583), doelend op het later (Colby & Kohlberg, 1987) gepubliceerde scoringssysteem, schrijft er bijvoorbeeld het volgende over: "For instance, the upper stage inclusion rule requires that whenever a subject gives material that matches the scoring criteria at a lower stage, it is not recorded on the scoring sheet if elsewhere in the discussion of that dilemma the subject gives more elaborated material that matches a higher stage". Dit soort coderingsregels zorgt er uiteraard voor dat scores een consistente indruk maken.

Hoewel het structurele-eenheidspostulaat in dit artikel, maar ook door andere onderzoekers, empirisch niet ondersteund wordt en er gaten zitten in de opgetrokken verdedigingslijn, is het niet aannemelijk dat Kohlbergs researchprogramma verlaten zal worden voor dat er een alternatief programma is.

Noten

1. Flavell (1985) in een bespreking van Piagets werk plaatste al in 1971 kanttekeningen bij het stadiumconcept met begrippen als 'structure', 'qualitative change', 'abruptness' en 'concurrency'.
2. In een eerdere publikatie (Kohlberg, 1976) hanteert Kohlberg een ander criterium voor de geldigheid van het 'structured whole'-postulaat, dat namelijk 50% van de scores van een individu, op één niveau valt. In deze studie wordt gereferreed aan recenter werk van Kohlberg (1981), waarin een strenger criterium wordt gehanteerd.
3. Met behulp van een computerprogramma is op een totaal van 172 eerste klasleerlingen MEAO een randomisering uitgevoerd, zodat qua leeftijd en sekse twee nagenoeg identieke groepen ontstonden, te weten groep 2 en groep 3.
4. Factoranalyses op de scores van kleine groepen, zoals de ZMOK-ers en MEAO-ers elk apart, is niet verantwoord (vgl. Johnson & Wichern, 1988). De Kohlberggroep voerde factoranalyses uit op de scores van 18 variabelen bij hun grootste groep (n = 40).
5. Hoewel extractie van factoren met een eigenwaarde groter dan één gebruikelijk is, is er op deze methode nogal wat aan te merken (vgl. Johnson & Wichern, 1988).
6. De contrastvergelijkingen tussen de scores op deze vraagstukken gaven voor iedere groep significante F-waarden.

Literatuur

- Colby, A., L. Kohlberg, *The Measurement of Moral Judgment*. Volume 1: Theoretical Foundations and Research Validation. Cambridge: University Press, 1987.
- Colby, A., L. Kohlberg, J. C. Gibbs & L. Lieberman, A longitudinal Study of Moral Judgment. *Monographs of the Society for Research in Child Development*, 1983, 48, nos 1-2. Chicago: University of Chicago Press.
- Dumont, J. J., *Leerstoornissen I*. Rotterdam: Lemniscaat, 1976.
- Dumont, J. J., *Leerstoornissen I*. Rotterdam: Lemniscaat, 1985.
- Flavell, J. H., *Cognitive development*. Englewood Cliffs, New Jersey: Prentice Hall Inc., 1985.
- Gibbs, J. C., K. D. Arnold, R. L. Morgan, E. S. Schwartz, M. P. Gavaghan & M. P. Tappan, Construction and validation of a Multiple-Choice Measure of Moral Reasoning. *Child Development*, 1984, 55, 527-536.
- Gilligan, C., In a Different Voice: Women's Conceptions of Self and of Morality. *Harvard Educational Review*, 1977, 47, 4, 481-517.

IJzendoorn M. H. van, Morele argumentaties en maatschappelijk protest. *Filosofie en Praktijk* 1981, 2, 66-79.

IJzendoorn M. H. van, Is Kohlbergs theorie van de morele ontwikkeling 'seksistisch'? *Kind en Adolescent*, 1986, 7, 34-39.

Johnson, R. A. & D. W. Wichern, *Applied Multivariate Statistical Analysis*. Englewood Cliffs, New Jersey: Prentice Hall Inc., 1988.

Kohlberg, L., The Cognitive-Development Approach to Moral Education. In: D. Purpel & K. Ryan (Ed.), ... *It Comes with the Territory*. Berkeley: Phi Delta Kappan, 1976.

Kohlberg, L., *The Philosophy of Moral Development, Moral Stages and the Idea of Justice*. San Francisco: Harper & Row Publishers, Volume 1, 1981.

Mey, J. R. P. B. de, J. van der Draai & B. Spiecker, Zorg en rechtvaardigheidsoordelen van scholieren, een empirisch onderzoek naar de controverse tussen Kohlberg en Gilligan. *Nederlands tijdschrift voor opvoeding, vorming en onderwijs*, 1989, 5, 338-349.

Rest, J. R., New Approaches in the Assessment of Moral Judgement. In: T. Lickona (Ed.), *Moral development and behavior. Theory, Research, and Social Issues*. New York: Holt, Rinehart & Winston, 1976.

Rest, J. R., Morality. In: J. H. Flavell, E. M. Markman (Ed.), *Cognitive Development. Manual of Child Psychology* (Vol. 3, pp. 556-629). New York: Wiley, 1983.

Sanvitale, D., H. D. Saltzstein & M. C. Fish, Moral Judgments by Normal and Conduct-Disordered Preadolescent and Adolescent Boys. *Merill-Palmer Quarterly*, 1989, vol. 35, 4, 463-481.

Vedder, R., *Kinderen met leer- en gedragsmoeilijkheden*. Groningen: Wolters-Noordhoff, 1977.

Curriculum vitae

L. de Mey (1950) studeerde sociologie (1978), en is als universitair docent verbonden aan de sectie Theoretische en Historische Pedagogiek van de Vrije Universiteit.

Adres: Vrije Universiteit, Sectie Theoretische en Historische Pedagogiek, Transitorium I, Van der Boechorststraat 1, 1081 BT Amsterdam.

Manuscript aanvaard 10-6-'91.

Summary

Mey, L. de. 'Variety in the moral judgments of students: an investigation of Kohlberg's structured whole postulate.' *Pedagogische Studiën*, 1991, 68, 265-275.

Structured wholeness is part of the hard core of Kohlberg's theory of moral judgments. In general, the given individual stage-responses on moral problems are consistent, according to Kohlberg. However, a lot of criticism has hit this statement, which has mostly a contemplative character. This article presents the results of an empirical study, aiming to contribute to the discussion on the validity of structured wholeness. Students from a College of Economic and Administrative Studies (MEAO-students, $n = 172$, mean age 17) and students from several schools for Children with Severe Conduct Disorders (ZMOK-students, $n = 64$, mean age 16) participated in this study. They rated moral problems of two of Kohlberg's dilemmas, the Heinzdilemma and the dilemma of Joe and his Father and a part with so-called Care-items. Multivariate analyses of variance and factor analyses are used to draw conclusions about the consistency of the moral judgments. The results show that no structured wholeness is found in the moral judgments of MEAO-students, in contrast to the ratings of certain moral problems for ZMOK-students.