

'Diagnostische' criteria, essentieel maar controversieel: overwegingen vanuit een kinderpsychiatrische praktijk

door G.F.M.G. Berden

Samenvatting

Door de introductie van de DSM-III is er een vernieuwde interesse ontstaan voor het gebruik van diagnostische criteria bij het afgrenzen van classificatiecategorieën. In dit artikel wordt door middel van een op de kinderpsychiatrische praktijk gericht exploratief onderzoek, waarin de DSM-III-criteria voor Major Depression Disorder vergeleken worden met de Weinberg-criteria, aangetoond hoe belangrijk zowel een zorgvuldige keuze van criteria als van scoringsvoorwaarden wel is. Kleine verschillen resulteren namelijk onmiddellijk in het selecteren van een totaal andere populatie. Uitgaande van de classificatiecategorie Depressie bij kinderen worden er ten slotte een aantal suggesties gedaan om tot een meer verantwoorde keuze van criteria en scoringsvoorwaarden te komen, om alzo betrouwbaarheid en validiteit van een dergelijk concept te verhogen.

Inleiding

Het systematisch hanteren van 'diagnostische' criteria bij het classificeren van psychopathologie, voor het eerst consequent toegepast in het DSM-III-classificatiesysteem (APA 1980), wordt door diverse auteurs (Van Praag 1982, Verhulst 1983) een wezenlijke vooruitgang binnen ons vakgebied genoemd. Het is namelijk door de introductie van deze criteria dat men bij het classificeren tot een hogere tussenbeoordelaarsbetrouwbaarheid kan komen, wat een essentiële vereiste is wil een classificatiesysteem bruikbaar en betekenisvol, valide, worden (Spitzer en Fleiss 1974). Meerdere onderzoekers moeten er namelijk, onafhankelijk van elkaar, met behulp van een dergelijk model in slagen om psychiatrische problematiek vrij eenduidig te classificeren. Vroegere classificatiemodellen beantwoordden onvoldoende aan deze basisvoorwaarde (Beck 1962, Kreitman 1961, Zubin 1968, Spitzer en Fleiss 1974, Kendell 1975). Ward en medewerkers (1962) toonden aan dat te vage classificatiecriteria, met als gevolg een te hoge criteriumvariantie, daarbij een belangrijke rol speelden.

In dit artikel worden, uitgaande van een welbepaalde classificatieca-

tegorie, namelijk depressie bij kinderen, een aantal problemen rond de keuze en de scoring van 'diagnostische' criteria nader belicht. Als kinderpsychiater ging mijn voorkeur uit naar het onderwerp depressie omdat met name rond de depressie-categorie in de literatuur nogal wisselende criteria worden vermeld, die hun voedingsbodem vinden in de zeer uiteenlopende standpunten die er rond het al of niet bestaan en de uitdrukkingsvorm van deze categorie bij kinderen beschreven zijn (Rochlin 1958, Rie 1966, Frommer 1968, Poznanski en Zroll 1970, Weinberg 1973).

Keuze en scoring van classificatiecriteria

Een verkenning: In hoeverre beantwoorden bestaande classificatiemodellen aan de klinische praktijk? – De diversiteit aan criteria binnen de veelheid aan depressiemodellen illustreert hoe moeilijk het is om klinische ervaringsgegevens, die steeds als de belangrijkste bron van herkomst genoemd worden, te expliciteren in welomschreven criteria.

Toetsing van de diverse keuzes en groeperingen van criteria is beperkt omdat vooralsnog de nodige variabelen ontbreken om de diverse klinisch-hypothetische constructies extern te valideren. Zo zijn er bij voorbeeld nog geen biologische waarden voorhanden die specifiek gerelateerd zijn aan één van de veelal descriptief bepaalde depressiemodellen.

Het ontbreken van een duidelijk extern criterium maakt het des te noodzakelijker dat een descriptief georiënteerde categorie zo nauw mogelijk aansluit bij haar klinische bron. Daarbij komt nog dat de herkenbaarheid in de klinische praktijk een essentiële vereiste is wil een dergelijk concept niet enkel de wetenschapper maar ook de clinicus boeien.

In een exploratief onderzoek werd nagegaan in hoeverre twee onderscheiden depressieclassificatiecategoriën, namelijk Major Depressive Disorder (MDD) uit de DSM-III en Kinderdepressie volgens Weinberg-criteria (1973), overeenstemden met de klinische praktijk.

Voor de categorie MDD werd geopteerd omdat depressies op de kinderleeftijd volgens de DSM-III-richtlijnen aldus gerangschikt worden. De Weinberg-constructie werd gekozen omdat deze categorie herhaaldelijk opduikt in de literatuur en enige weerklank vindt binnen de research. Hieronder volgen de classificatiecriteria voor beide kategoriën.

I. Criteria voor Major Depressive Disorder (MDD) volgens DSM-III:

A. *Dysfore stemming*, verlies van interesse of plezier in alle of bijna alle activiteiten. De dysfore stemming is gekenmerkt door de volgende symptomen: gedeprimeerd, bedroefd, neerslachtig, hopeloos, in de put, niet monter. De stemmingsstoornis moet op de voorgrond staan en

een zekere persistentie hebben, maar hoeft niet noodzakelijkerwijs het meest belangrijke symptoom te zijn. Plotselinge veranderingen van de ene dysfore stemming naar de andere, bij voorbeeld van verdriet naar woede, weer gevolgd door dysfore opwinding, zijn niet inbegrepen. Voor *kinderen jonger dan 6 jaar* kan de dysfore stemming afgeleid worden uit de persisterende *sombere gelaatsuitdrukking*.

B. Ten minste vier van de volgende symptomen moeten voor een periode van tenminste *twee weken*, bijna *elke dag*, aanwezig zijn. Voor *kinderen onder 6 jaar* ten minste drie van de eerstvolgende vier symptomen:

1. *verlies van eetlust of gewichtsverlies van enige betekenis* (zonder het volgen van dieet) of *toegenomen eetlust en gewichtstoename van betekenis*; voor *kinderen beneden 6 jaar* eveneens het *uitblijven van verwachte gewichtstoename* erin betrekken;
2. *slapeloosheid of slaperigheid*;
3. *psychomotorische opwinding of bewegingsarmoede* (maar niet het subjectieve gevoel van rusteloosheid of traagheid); bij *kinderen beneden 6 jaar*: *hypoactiviteit*;
4. *verlies van interesse of plezier in de gebruikelijke activiteiten of vermindering van aandrift*, niet beperkt tot een periode met wanen of hallucinaties; bij *kinderen beneden 6 jaar* tekenen van *apathie*;
5. *vermindering van energie*, vermoeidheid;
6. gevoelens van *niet de moeite waard zijn*, waardeloos zijn, zelfverwijten of overmatig of inadequaat schuldgevoel (eventueel met waanvorming);
7. klachten of aanwijzingen van een *verminderd vermogen om te denken* of zich te concentreren, zoals vertraagd denkproces of besluiteloosheid, niet verbonden met verhoogd associatief denken of incoherent denken;
8. terugkerende gedachten over de *dood, suïcidale gedachten of suïcidepoging*.

C. *Geen* van de volgende symptomen domineert het klinisch beeld als een affectieve stoornis (volgens bovenvermelde criteria A en B) niet aanwezig is, dat wil zeggen voor deze zich ontwikkeld heeft of erna:

1. *preoccupatie* met een stemmingsincongruente waan of hallucinatie;
2. *bizar gedrag*.

D. *Niet* gesuperponeerd op *schizofrenie, schizofreniforme stoornis of een paranoïde stoornis*.

E. *Niet* te wijten aan enige *organische mentale stoornis* of een *ongecomplieerd zwaar verlies* (bijv. sterfgeval).

II. Criteria voor kinderdepressie volgens Weinberg:

1. er moet sprake zijn van een *dysfore stemming*, wat zich onder

meer kan uiten in een droef uiterlijk, zich eenzaam en ongelukkig voelen, zich hopeloos voelen, geïrriteerd zijn en hypergevoelig zijn;

2. uitingen van een *laag zelfgevoel*, uitingen van gevoelens van waardeloosheid, zich dom ervaren of zich schuldig voelen, doodwensen en suïcidepogingen.

Aan *beide* criteria moet *voldaan* zijn. Vervolgens wordt gevraagd dat *minstens aan twee van de volgende acht* criteria voldaan wordt:

3. *agressief gedrag*, twistziek gedrag, vijandig en geagiteerd gedrag, excessief veel vechten en woedeaanvallen;
4. *slaapstoornissen* zoals moeilijk inslapen, onrustig slapen of moeilijk wakker worden 's morgens;
5. *verminderde schoolprestaties*, dagdromen, verminderde interesse voor school, onaf inleveren van huiswerk;
6. minder sociaal zich inzetten, minder vriendelijk zijn, *zich terugtrekken*, verlies van de interesse in de gewone sociale contacten;
7. *wijzigingen in de houden ten opzichte van school*, er geen plezier meer in hebben, school weigeren;
8. *lichamelijke klachten*, niet migraineuze hoofdpijn, buikpijn, spierpijn en andere;
9. *vermoeidheid*, verlies van persoonlijke interesses en hobby's, mentale en fysische vermoeidheid;
10. *wijzigingen in eetlust en/of gewicht*, anorexie of polyfagie, ongewone gewichtsveranderingen.

Een symptoom werd aanwezig geacht zo het bij de ouders of bij het kind een *bron van zorg* is of zo het ervaren wordt als een *wijziging in het gebruikelijke functioneren* van het kind. Een symptoom moet ook minstens gedurende *1 maand* aanwezig zijn.

Methode – Over een periode van 6 maanden werden de genoemde criteria systematisch geïnventariseerd bij alle kinderen die werden aangemeld op een polikliniek*. Dit gebeurde door de aldaar werkzame kinderpsychiaters, al of niet in opleiding, te vragen of ze op een vragenlijst wilden aankruisen of de genoemde items al of niet aanwezig waren bij elk kind dat ze in onderzoek kregen. Uiteraard kwamen op deze lijst ook alle uitsluitingscriteria voor. Eenmaal ingevuld werden de vragenlijsten, voorzien van het dossiernummer van de betrokken kinderen, ingeleverd. Vervolgens werd bekeken of in de aldus verkregen informatie de nodige en voldoende criteria aanwezig waren om te kunnen besluiten tot een MDD-classificatie en/of een kinderdepressie beantwoordend aan de Weinberg-richtlijnen. Vervolgens werden door mij

* Met dank aan alle medewerkers van de polikliniek kinderpsychiatrie van het Academisch Medisch Centrum-Amsterdam (Hoofd prof. dr. D.J. de Levita).

van alle aangemelde kinderen de onderzoeksverslagen doorgelezen om na te gaan of de kinderpsychiater die hen onderzocht had van oordeel was dat een depressieve stemmingsstoornis centraal stond in hun problematiek, en of een depressie expliciet vermeld werd in de einddiagnose.

Resultaten – De onderzoekspopulatie omvatte in het totaal 75 kinderen. Bij vier van hen werd door de kinderpsychiater in de einddiagnose expliciet genoemd dat er sprake was van een depressieve stoornis. Op basis van de vragenlijstinformatie scoorden 9 van de 75 kinderen positief voor MDD en 14 kinderen, uit de groep van 75, voldeden aan de door Weinberg gestelde voorwaarden.

Bij nadere analyse bleek dat de vier kinderen, bij wie er op basis van het kinderpsychiatrisch onderzoek gesproken werd van een depressie, allen deel uitmaakten van de MDD-groep, terwijl de nodige Weinberg-criteria slechts aanwezig waren bij 3 ervan. Dus 5 kinderen uit de MDD-groep en 11 uit die van Weinberg voldeden wel aan de respectievelijk gestelde categorisatierichtlijnen, doch werden uitgaande van het kinderpsychiatrisch oordeel niet als depressief gediagnostiseerd. We hebben deze onderzoeksresultaten gevisualiseerd in tabel 1.

Hanteren we nu het klinisch oordeel van de kinderpsychiater als, geenszins onfeilbare, maat, dan geeft zowel het gebruik van de MDD-richtlijnen als van de Weinberg-criteria aanleiding tot een vrij hoge 'vals-positieve' score. Dit betekent onder meer dat het volgens genoemde classificatieregels samenstellen van een onderzoekspopulatie voor een researchproject, waarin men bij voorbeeld op klinische inzichten verkregen hypothesen poogt te toetsen, tot gevolg heeft dat te veel individuen in de groep opgenomen worden die er niet in thuis horen. Relevante klinische inzichten kunnen aldus onterecht verworpen worden.

Weer met het eindoordeel van de kinderpsychiater als referentiepunt was er uitgaande van de Weinberg-criteria sprake van 1 'vals-negatieve' score. Daardoor kan bij grotere onderzoeksgroepen belangrijke informatie verloren gaan doordat te veel individuen die in de onderzoekspopulatie thuishoren er niet in worden opgenomen.

Het feit dat van een aantal kinderen het gedrag wel beantwoordde aan de nodige depressiecriteria, zonder dat er bij hen door de kinderpsychiater een depressie werd gediagnostiseerd, was voor mij aanleiding om wat verder op hun problematiek in te gaan. Aan de kinderpsychiaters die hen onderzocht hadden werd de vraag gesteld of ze van oordeel waren of de betrokken kinderen ook manifest depressief waren. Dit was volgens hen wel het geval voor de 5 kinderen uit de MDD-categorie bij wie in de einddiagnose geen depressie vermeld werd, doch gold niet voor 7 van de 11 kinderen uit de Weinberg-groep waarbij een soortgelijke discrepantie bestond. Op de vraag of voor deze 7 kinderen soms het concept 'gemaskeerde' depressie (Toolan 1960, Glaser 1967) gebruikt

kon worden, werd eveneens ontkennend geantwoord.

In die gevallen waar er sprake was, volgens de desbetreffende kinderpsychiater, van een manifest depressief kind, zonder dat de term depressie expliciet genoemd werd in de einddiagnose, bleek telkens dat het klinisch-descriptief beeld ondergeschikt gemaakt was aan een etiologisch-psychodynamisch verklaringsmodel. Zo werd het depressieve aspect gevangen onder concepten als een fobisch-hysterische of angst-hysterische neurose.

Overigens: bij slechts 5 kinderen was er een overlap tussen de MDD en de Weinberg-groep. In onderstaande tabel worden alle resultaten nogmaals samengevat:

Tabel 1: De onderlinge relaties tussen het kinderpsychiatrisch oordeel, de MDD-classificatie en de Weinberg-classificatie

	totaal	het kinderpsychiatrisch oordeel	
		depressie	geen depressie
MDD	9	4	5*
Weinberg	14	3	11*
MDD + Weinberg	5	3	2
enkel MDD	4	1	3
enkel Weinberg	9	0**	9

* 'vals-positieve scores'

** 'vals-negatieve score'

Het gegeven dat de Weinberg-criteria duidelijk een andere, minder aan de kinderpsychiatrische praktijk beantwoordende, problematiek afgrensden dan de MDD-richtlijnen, was voor mij een extra stimulans om wat nader te blijven stilstaan bij de keuze en de scoringsvoorwaarden van, in dit geval, depressiecriteria.

Keuze van criteria—Om tot een selectie van criteria te komen die nodig geacht moeten worden wil men een welbepaalde problematiek kunnen aflijnen, kan men onder meer gebruik maken van op de klinische praktijk gerichte toetsingsprocedures en van factoranalytische onderzoeksmethoden.

a. *Klinische toetsingsprocedures.* Een voorbeeld daarvan is het door ons verrichte exploratieve onderzoek. Wanneer we de MDD-instructies met deze van Weinberg vergelijken, dan valt onmiddellijk de vrij grote mate van overeenkomst op. De MDD-criteria 'verlies van plezier in de gebruikelijke activiteiten' en 'vertraagd denkproces en concentratieproblemen' zijn ons inziens enigszins vergelijkbaar met wat bij Weinberg genoemd wordt 'verminderde interesse in school' en 'verminderde schoolprestaties en dagdromen', zodat het enige criterium dat daadwerkelijk verschilt is, het, enkel bij Weinberg genoemde al of niet aanwezig zijn van lichamelijke klachten.

Is het zo belangrijk dat *lichamelijke klachten* als obligaat criterium opgenomen worden in de lijst die beoogt de categorie depressie af te lijnen? In de klinische praktijk wordt herhaaldelijk gewezen op het voorkomen van somatische klachten, zonder onderliggend somatisch lijden bij kinderen met een depressieve stoornis. Ling (1970) en Kashani (1980, 1982) vermelden zelfs dat bij deze kinderen dergelijke klachten vaak de directe aanmeldingsreden vormen. Om tot de classificatie depressie te besluiten maakte Ling daarbij gebruik van de Weinberg-criteria en Kashani van de DSM-III-richtlijnen.

Men kan zich echter afvragen of lichamelijke klachten wel zo specifiek gerelateerd zijn aan een depressieve stoornis en of ze niet eerder een maat zijn voor psychisch dysfunctioneren in zijn algemeenheid. In ons exploratieve onderzoek bleek namelijk dat dergelijke klachten bij 25 van de in totaal 75 kinderen uit de onderzoekspopulatie voorkwamen.

Het onderkennen van criteria met een lage graad van specificiteit is belangrijk omdat het voornamelijk deze criteria zijn die het risico vergroten dat een aldus omlijnd concept te vaag blijft om nog klinische en/of wetenschappelijke relevantie te kunnen hebben. Onderzoek naar de relatieve specificiteit van de diverse criteria lijkt me dan ook aangewezen te zijn. Ik denk in dit verband aan slaap- en eetproblemen waarvan we op klinische gronden weten dat ze bij diverse vormen van psychopathologie bij kinderen kunnen voorkomen.

b. Empirische benaderingswijze. Research die refereert naar de klinische praktijk gaat na in hoeverre een geconstrueerde classificatiecategorie een aanvaardbare representant is van de klinische concepten die als zijn inspiratiebron gefungeerd hebben. Hoe belangrijk een dergelijke benadering ook is, door de kring-redenering die er in vevat ligt blijft de bewijskracht ervan beperkt.

Een waardevolle aanvulling kan geboden worden door van factoranalytische technieken gebruik makend onderzoek, waarbij er niet uitgegaan wordt van a priori-concepten maar van een empirische benadering. Na een uitgebreide inventarisatie te hebben gemaakt van een groot aantal gedragsitems, gaat men namelijk na in hoeverre de diverse items tenderen om samen voor te komen. Indien er een welomschreven depressiecategorie bestaat moet het door middel van dergelijk onderzoek mogelijk zijn, in een kinderpsychiatrische populatie, een aantal items te vinden die in vrij hoge mate tenderen om samen voor te komen en die onder de noemer depressie gerangschikt kunnen worden. Deze gedragsitems kunnen dan fungeren als classificatiecriteria.

Wat deze vorm van research betreft is er binnen de kinderpsychiatrie onder meer onderzoek verricht door T.M. Achenbach (1978-1979). Na gebruik makend van oudervragenlijsten, een aantal gedragingen te hebben verzameld bij kinderen die aangemeld waren bij GGZ-instellingen, kon hij door middel van factoranalyse een aantal empirische syndromen omlijnen. Binnen de leeftijdsgroep 6- tot 11-jarigen was er, zowel bij jongens als bij meisjes, sprake van een aantal items die in vrij

hoge mate samen voorkwamen en die onder de term depressie gebundeld konden worden. Deze items waren: zich waardeloos/schuldig voelen, zorgelijk/bedroefd zijn, suïcidale uitingen vertonen, zich isoleren/terugtrekken, perfect willen zijn, angstig zijn voor eigen impulsen, zich achtervolgd voelen, zenuwachtig zijn en pruilen. Er is dus slechts ten dele overeenstemming met de MDD-criteria; enkel de eerste vier items vinden we terug.

De overeenstemming wordt nog teleurstellender als men de leeftijdscategorie 12-tot 16-jarigen bekijkt. Items die in voldoende mate tenderen om samen voor te komen, en die onder het begrip 'depressie' gerangschikt kunnen worden, ontbraken bij de jongens uit deze leeftijdsgroep en kwamen slechts ten dele bij de meisjes voor. Tevens onderstrepen deze bevindingen dat men terdege rekening moet houden met mogelijke sekse- en leeftijdsinvloeden. De ontwerpers van de DSM-III zijn er bij de depressie categorie nochtans van uitgegaan dat deze aspecten te verwaarlozen waren. De MDD-categorie maakt zelfs geen deel uit van het specifiek kinderpsychiatrische onderdeel van de DSM-III, maar hoort bij de sectie voor volwassenen. Of dit wel zo verantwoord is blijft zeer de vraag, en de uitspraak van Puig-Antich (1982) dat de criteria geen manifeste veranderingen meer zouden ondergaan na de leeftijd van 6-7 jaar vraagt zeker om bevestiging.

Ook deze factoranalytische benaderingswijze heeft echter zijn beperkingen. Samenstelling van de vragenlijsten, wijze van inventarisieren van de gedragsitems, verwerking van het materiaal e.d. heeft invloed op de uiteindelijke onderzoeksresultaten. Achenbach zelf benadrukt dan ook dat verder onderzoek op dit vlak zonder meer noodzakelijk is. Hoe belangrijk dit is blijkt al als we de bevindingen van deze auteur vergelijken met de onderzoeksresultaten van Arnold, die in 1974 een gelijkwaardig onderzoek verrichtte. Eveneens gebruik makend van een oudervragenlijst inventariseerde deze het al dan niet aanwezig zijn van een aantal gedragsitems bij 351 kinderen, aangemeld op een kinderpsychiatrische polikliniek. Ook hier kon door middel van factoranalyse een categorie 'depression-withdrawal' afgebakend worden, dit zowel in de leeftijdsgroep onder de 12 jaar als bij de 12- tot 18-jarigen. In tegenstelling tot Achenbachs onderzoek werd er hier geen opsplitsing jongens-meisjes gemaakt, wat uiteraard van invloed kan zijn op de verkregen resultaten. De gedragsitems die voor de categorie 'depression-withdrawal' in aanmerking kwamen waren: verdrietig zijn, interesseverlies, pruilend gedrag, zich niet geliefd/eenzaam voelen, verward denken, erg gesloten zijn en blijvend herhalen van handelingen. Deze reeks stemt slechts ten dele overeen met die van Achenbach. Overigens het criterium 'denkstoornissen' dat in het onderzoek van laatstgenoemde niet aan bod kwam, doch dat wel gehanteerd wordt in de MDD-categorie, krijgt nu wel bevestiging. Op te merken valt dat de gedragsitems lichamelijke klachten, eet- en slaapstoornissen ontbreken in de depressie categorieën van beide onderzoe-

ken. Dit suggereert ook nu weer dat deze drie criteria waarschijnlijk erg specifiek zijn wat betreft het depressieconcept. De vraag duikt dan ook op of deze items wel deel moeten uitmaken van een depressieconstruct.

Als we in het door ons verricht exploratief onderzoek de 3 genoemde criteria buiten beschouwing lieten, dan bleek dat alsnog de problematiek van 8 van de 9 kinderen, die tot de MDD-groep behoorden, aldus geclassificeerd konden worden op basis van de resterende 7 criteria. De 4 kinderen in het exploratief onderzoek waarbij op klinische gronden gesproken werd van een depressieve stoornis behoorden tot die groep van acht.

Scoringsvoorwaarden – Niet enkel de keuze van criteria maar ook de scoringsvoorwaarden die daarbij gehanteerd worden zijn erg belangrijk. Zo bleek in ons onderzoek dat het stellen van 2 *essentiële voorwaarden* door Weinberg tot gevolg had dat de problematiek bij 1 kind, waarbij de kinderpsychiater wel een depressie diagnostiseerde, niet voldeed aan de geformuleerde vereisten om tot een eensluidende categorisatie te besluiten. Deze overeenstemming was er wel met het MDD-concept, dat slechts 1 essentiële voorwaarde kent.

Men kan vanuit klinische ervaring beargumenteren dat het stellen van meerdere absolute criteria grote risico's in zich draagt. Een classificatiecategorie is namelijk een ideaalconcept dat de werkelijkheid slechts ten dele benadert. Kanner (1969) geeft treffend uitdrukking aan dit onderscheid tussen theorie en praktijk als hij stelt 'children haven't read those books'. Wat geconstrueerd wordt is een 'prototype' (Rosch, Mavis, Gay e.a. 1976), waarvan de meest gangbare criteria genoemd worden doch waarbij men nooit verwachten kan dat steeds aan alle criteria voldaan is. De heterogeniteit die aldus binnen eenzelfde classificatiecategorie mogelijk blijft, wanneer men het aantal essentiële voorwaarden beperkt, beantwoordt meer aan de realiteit.

Ook het *minimum aantal criteria* dat noodzakelijk geacht wordt speelt een belangrijke rol. Voor het scoren van MDD moet er minstens voldaan worden aan 5 criteria, bij Weinberg kan men volstaan met 4. Dit verschil kan één der factoren zijn die maakt dat voor de eerstgenoemde categorie een grotere depressie-intensiteit vereist is dan voor de tweede. Tot deze bevinding kwamen namelijk Carlson en Cantwell (1982). Deze onderzoekers zijn niet enkel nagegaan hoeveel kinderen uit een kinderpsychiatrische populatie problematiek vertoonden die voldeed aan de voor beide categorieën gestelde voorwaarden, maar ze hebben tevens een schatting van de depressie-intensiteit gemaakt door middel van een vragenlijst, de Children Depression Inventory, CDI (Kovacs 1978). Het resultaat van hun onderzoek was dat op een totaal van 102 kinderen er 27% voldeed aan de MDD-richtlijnen, 57% aan de Weinberg-criteria en daarbij vond men dat de gemiddelde depressie-intensiteit in de eerste groep tweemaal groter was dan in de tweede.

Verder onderzoek is echter noodzakelijk, en bepalen van deze minimumgrens blijft voorsnog een vrij arbitrair gegeven. Toetsing aan de klinische praktijk is ook hier, zolang externe validerende factoren ontbreken, de beste richtlijn.

Scoringsbetrouwbaarheid – Hoe betrouwbaar is elk criterium te scoren? Indien men door het gebruik van criteria de betrouwbaarheid van een classificatiesysteem wil verhogen dan is het noodzakelijk dat de tussenbeoordelaarsbetrouwbaarheid voor elk van de geselecteerde criteria vrij hoog is.

Het is echter niet zo eenvoudig om te voorspellen welke gedragsitems vermoedelijk met een vrij grote mate van betrouwbaarheid te scoren zullen zijn. Zo is het geen algemene regel dat men subjectieve, globale gedragsitems minder betrouwbaar scoren kan dan meer objectieve en specifiekere. Rutter en Graham (1968) vonden bij voorbeeld dat men tijdens een kinderpsychiatrisch onderzoek een eenduidiger score bereikte bij een item als 'emotionele responsiviteit' dan bij 'snel schrikken of beven'. Op hun beurt constateerden Wing (1974) en Luria (1979) dat psychomotore opwindning of geremdheid, één der MDD-criteria, tot de minst betrouwbaar te scoren items uit de Present State Examination (Wing 1967) gerekend moest worden.

Een hoge betrouwbaarheidsscore op een globaal item is anderzijds niet altijd een argument om dit te verkiezen boven concretere, wat minder hoog scorende, criteria. Het item 'angstig zijn' was in het onderzoek van Rutter en Graham (1968) vrij betrouwbaar te scoren maar differentieerde minder tussen een doorsnee- en een kinderpsychiatrische populatie dan aanverwante items, met een iets lagere betrouwbaarheidsgraad, als 'een angstige gelaatsuitdrukking' en 'gepreoccupeerd zijn met beangstigende thema's'. Blijkbaar hebben laatstgenoemde gedragsitems een veel grotere klinische relevantie.

Meerdere factoren kunnen een rol spelen bij de nogal uiteenlopende onderzoeksresultaten die we in de literatuur aantreffen betreffende de betrouwbaarheid van de diverse criteria. Zo worden weinig frequent optredende gedragsitems in een in tijd beperkte onderzoekssituatie uiteraard vaak gemist, maar ook onbekendheid van de onderzoeker met een dergelijk gedrag kan er oorzaak van zijn dat een, eventueel, sporadisch optreden niet wordt opgemerkt. Persoonlijk ervaren we steeds weer in onderwijssituaties dat studenten bij het bekijken van videobanden over gedragsstoornissen bij kinderen meer significante gedragsaspecten opmerken wanneer ze vooraf een theoretische uiteenzetting krijgen, dan wanneer dit slechts nadien gebeurt. Trainings- en ervaringsachtergronden van de diverse onderzoekers zullen dus wel degelijk een rol spelen. Illustratief in dit verband zijn onder meer de bevindingen van Kendell (1973) dat, althans in zijn onderzoek, beoordelaars nauwelijks gebruik maakten van non-verbale informatie. Of men nu bij de classificatieopdracht een gevalsbeschrijving zonder ver-

wijzingen naar concrete gedragingen hanteerde of een videotape was nauwelijks van invloed op de tussenbeoordelaarsbetrouwbaarheidsscore. Het hanteren van non-verbale criteria zal bij beoordelaars die voornamelijk verbaal zijn ingesteld zeker de betrouwbaarheidsscore drukken.

Daar een classificatiesysteem als de DSM-III ook gebruik maakt van door ouders verstrekte gegevens is het belangrijk dat men ook van die informatiebron een betrouwbaarheidsgraad bepaalt. Globaal is dit van een vrij aanvaardbaar niveau (Verhulst e.a. 1985) doch ook hier is het belangrijk om naar de individuele items te kijken. Concrete gedragingen als bedplassen en nagelbijten werden in het onderzoek van Graham en Rutter (1968) door ouders veelal accurater gesignaleerd dan items met een meer subjectieve inbreng. Zo werd er door hen nogal inconsistent geantwoord op de vraag of hun kind in een bepaalde periode een depressieve indruk maakte.

Overigens moeten we opmerken dat doordat de diverse auteurs andere betrouwbaarheidsmaten hanteren, een onderlinge vergelijking erg bemoeilijkt wordt en dan ook de nodige voorzichtigheid bij de interpretatie geboden is.

Hoe bereikt men dat elk criterium daadwerkelijk gescoord wordt? Wil men tot een zo accuraat mogelijke inventarisatie van gedragsitems komen, dan is een gestandaardiseerde onderzoekstechniek wenselijk. Een dergelijke benadering verhoogt de kans op volledigheid (Saghir 1971; Weitzel 1973; Greist 1973; Meikle 1970) en vermindert de 'bias' (Welner 1974; Kendell 1968). Horowitz (1982) toont in een aantal proefopstellingen aan dat het risico van 'bias' erg hoog is wanneer men prototypeconcepten, zoals de DSM-III-categorieën, hanteert. Het aan diverse beoordelaars laten lezen van gevalsbeschrijvingen waarin meerdere maar net niet voldoende criteria vervat zijn om aan een welbepaalde classificatiecategorie te beantwoorden heeft tot gevolg dat een aantal onder hen terecht concluderen dat aan de gestelde voorwaarden niet voldaan is, maar een relatief grote groep acht ze wel aanwezig.

Een meer gestructureerde scoring kan het risico op dergelijke discrepanties verminderen en alzo de informatievariantie beperken (Beck 1962). Deze is namelijk naast de reeds benoemde criteriumvariantie een belangrijke ontstaansfactor in de gebrekkige betrouwbaarheid van bestaande classificatiemodellen.

Een belangrijke bijdrage kan daarbij geleverd worden door *gestructureerde kinderpsychiatrische onderzoeksmethoden*. In dit verband kennen we momenteel onder meer de KIDDIE-SADS (Chambers en Puig-Antich 1974), de ISC (Kovacs 1977) en de CAS (Hodges 1981). In de literatuur komt duidelijk naar voren dat de extra-informatie die een dergelijke gestructureerde aanpak oplevert, opweegt tegen de beperkingen (Rutter en Graham 1968; Climent 1975; Herjanic 1975; Helzer 1978). Eén der nadelen van een dergelijke benadering is dat bepaalde

items zoals antisociaal gedrag minder aan bod komen of minder betrouwbaar door het kind worden gerapporteerd (Rutter 1967; Verhulst 1985) doch dit kan enigszins opgevangen worden door een wat vrijere onderzoekssituatie in de procedure op te nemen (Carpenter 1976).

Het belangrijkste voordeel dat het gebruik van een gestandaardiseerde onderzoeksmethode biedt is dat er doelgericht alle gedragsitems in kunnen worden opgenomen die onder meer nodig zijn om tot classificatie volgens de DSM-III-criteria over te kunnen gaan.

Validering van een classificatiecategorie

Streven naar een hoge tussenbeoordelaarsbetrouwbaarheid heeft uiteindelijk slechts zin wanneer de desbetreffende categorie ook betekenisvol is, dat wil zeggen: enige validiteit bezit. Men kan daarbij diverse vormen, niveaus, onderscheiden. Een eerste vorm is *validiteit opzicht*, face validity. Men kan dit slechts ten dele als een validiteitsconcept beschouwen daar het slechts uitdrukt dat bij voorbeeld een categorie als MDD waardevol lijkt te zijn op basis van beschrijvingen door ervaren klinici. Belangrijker is het al of niet bestaan van *inhoudsvaliditeit*, content validity. Dit betekent dat de desbetreffende categorie ook een waardige representant is van de psychische problematiek die ze wenst te omschrijven. In dit artikel hebben we gepoogd aan te geven dat het blijven aansluiten met de klinische praktijk niet zo'n eenvoudige zaak is.

Een volgende stap is het aantonen van *gelijktijdige validiteit*, concurrent validity. Dit betekent dat er een relatie kan aangetoond worden met een aan het oorspronkelijke concept externe variabele. Voorbeelden hiervan zijn onderzoeken naar het verband tussen het effect van de dexamethasone-suppressietest (Puig-Antich 1979a; Poznanski 1981; Geller 1983) of afwijkingen in de monoamine-stofwisseling (Puig-Antich 1979b; Cytryn en McKnew 1982) en MDD. Klinisch belangrijk is het bestaan van *voorspelbaarheidsvaliditeit*, predictive validity. Een concept als MDD is nog te recent ontwikkeld om op dit vlak al enige uitspraak te kunnen doen. De meest omvattende validiteitsvorm is de *begripsvaliditeit*, construct validity. Deze is gericht op een meer abstracte vorm van gedragsbeschrijving. Het vereist het geleidelijk opstapelen van informatie uit verschillende bronnen: klinisch, demografisch, familiaal, factoranalytisch onderzoek enz. Een concreet voorbeeld hiervan is research naar het voorkomen van problematiek beantwoordend aan de MDD-criteria bij specifieke doelgroepen. Zo verrichtten Cytryn en McKnew (1982) een exploratief onderzoek waarbij nagegaan werd of de MDD-classificatie vaker van toepassing was bij kinderen van manisch-depressieve ouders dan bij een controlegroep.

Conclusie

Het consequent gebruiken van 'diagnostische' criteria is zonder meer belangrijk bij de uitbouw van een betrouwbaar en valide classificatiesysteem. In dit artikel hebben we pogen aan te tonen dat zowel keuze van criteria als van scoringsmethode nog verder onderzoek vereist wil het uiteindelijke doel bereikt worden.

Het feit dat de aansluiting met de klinische praktijk nog maar ten dele geslaagd is bij een ontwerp als de DSM-III, zoals onder meer door Smith en Kraft (1983) onderstreept wordt, mag ons inziens geen reden zijn om niet op de ingeslagen weg verder te gaan. Een voortdurende terugkoppeling naar de klinische praktijk, een vergelijken met bevindingen uit empirisch onderzoek, gericht onderzoek naar de betrouwbaarheidsscores van diverse potentiële gedragscriteria e.d. kan daarbij een belangrijke bijdrage leveren.

Een aspect dat men echter niet uit het oog mag verliezen is dat classificeren niet identiek is met diagnostiseren. Vele critici van de DSM-III maken dit onderscheid onvoldoende. *Classificeren*, althans wanneer we dit begrip hanteren uitgaande van wat Stengel (1954) noemt 'publieke classificatie', beoogt het bevorderen van communicatie over psychopathologie tussen deskundigen van diverse theoretische vorming, zodat klinische ervaringsgegevens en onderzoeksresultaten een zo breed mogelijk bereik kunnen hebben. *Diagnostiseren* daarentegen is een voortdurend voortschrijdend proces waarin gepoogd wordt het psychisch (dys)functioneren van een individu, binnen zijn biologische en psychosociale ontwikkelingsgeschiedenis en huidige context, zo veelomvattend mogelijk te omschrijven.

Een classificatiecategorie kan bijdragen aan de diagnostiek in zoverze specifieke informatie toevoegt, bij voorbeeld door te verwijzen naar een door onderzoek gefundeerd waarschijnlijk verloop of vermoedelijke etiologie. Ze is echter ontoereikend als 'diagnose', wat treffend geïllustreerd en bekritiseerd wordt in een aflevering van Medisch Contact (1984).

Men kan in dit verband terecht opmerken dat het verwarrend is om bij een classificatiesysteem te spreken van 'diagnostische' criteria. Consequenter en juister zou zijn te spreken van classificatiecriteria.

Literatuur

- Achenbach, T.M. (1978), The child behavior profile: I-Boys aged 6-11. *Journal of Consulting and Clinical Psychology*, 46, 478-488.
- Achenbach, T.M., en C.S. Edelbrock (1979), The child behavior profile: II-Boys aged 12-16 and girls aged 6-11 and 12-16. *Journal of Consulting and Clinical Psychology*, 47, 223-233.
- American Psychiatric Association (1980), *Diagnostic and statistical manual of mental disorders*, 3e dr., Washington DC.
- Arnold, L.E., en D.J. Smeltzer (1974), Behavior checklist factor analysis for

- children and adolescents. *Archives of General Psychiatry*, 30, 799-804.
- Beck, A.T. (1962), Reliability of psychiatric diagnosis. 1. A critique of systematic studies, *American Journal of Psychiatry*, 119, 210-216.
- Berden, G.F.M.G. (1982), Het gebruik van diagnostische criteria in de kinderpsychiatrie: een exploratief onderzoek (interne publikatie).
- Carlson, G.A., en D.P. Cantwell (1982), Diagnosis of childhood depression: A comparison of the Weinberg and DSM-III criteria. *Journal of the American Academy of Child Psychiatry*, 21, 247-250.
- Carpenter, W.T., M.H. Sacks, J.S. Strauss, J.J. Bartko en J. Rayner (1976), Evaluating signs and symptoms: Comparison of structured interview and clinical approaches. *British Journal of Psychiatry*, 128, 397-403.
- Chambers, W., J. Puig-Antich en M.A. Tabrizi (1978), *The ongoing development of the KIDDIE-SADS*. Paper presented at the meeting of the American Academy of Child Psychiatry, San Diego.
- Climent, C.E., R. Plutchik, H. Estrada, L.F. Gowiria en W. Arévalo (1975), A comparison of traditional and symptom-checklist-based histories. *American Journal of Psychiatry*, 132, 450-453.
- Cytryn, L., D.H. McKnew, J.J. Bartko, M. Lamour en J. Hamovit (1982), Offspring of patients with affective disorders. *Journal of the American Academy of Child Psychiatry*, 21, 389-391.
- Feigner, J.P., E. Robins, J.B. Guze, R.A. Woodruff, G. Winokur en R. Munor (1972), Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry*, 26, 57-63.
- Frommer, E.A. (1968), Depressive illness in childhood, In: Coppen en Walk, *Recent developments in affective disorders: a symposium*, Headley, London.
- Geller, B., A.D. Rogol en E.F. Knitter (1983), Preliminary data on the Dexamethasone Suppression Test in children with major depressive disorder. *American Journal of Psychiatry*, 140, 620-622.
- Glaser, K. (1967), Masked depression in children and adolescents. *American Journal of Psychotherapy*, 21, 565-574.
- Greist, J.H., L.H. van Cura en N.P. Kneppreth (1973), A computer interview for emergency room patients. *Computer Biomedical Res.* 6, 257-265.
- Graham, P., en M. Rutter (1968), The reliability and validity of the psychiatric assessment of the child: II. Interview with the parent. *British Journal of Psychiatry*, 114, 581-592.
- Helzer, J.E., P.J. Clayton, R. Lambakion en R.A. Woodruff (1978), Concurrent diagnostic validity of a structured psychiatric interview. *Archives of General Psychiatry*, 35, 849-853.
- Herjanic, B., M. Herjanic, F. Brown en T. Wheatt (1975), Are children reliable reporters? *Journal of Abnormal Child Psychology*, 3, 41-48.
- Hodges, K., J. Kline, L. Stern, L. Cytryn en D. McKnew (1982), The development of a child assessment interview for research and clinical use. *Journal of Abnormal Child Psychology*, 10, 173-189.
- Horowitz, L.M., J.C. Wright, E. Lowenstein en H.W. Parad (1981), The prototype as a construct in abnormal psychology: 1. A method for deriving prototypes. *Journal of Abnormal Psychology*, 90, 568-574.
- Kanner, L. (1969), The children haven't read those books: reflections on differential diagnosis. *Acta Paedopsychiatrica* 36, 2-11.
- Kashani, J.H., G.J. Barbero en F.D. Bolander (1981), Depression in hospitalized pediatric patients. *Journal of the American Academy of Child Psychiatry*, 20, 123-134.

- Kashani, J.H., Z. Lababidi en R.S. Jones (1982), Depression in children and adolescents with cardiovascular symptomatology: the significance of chest pain. *Journ. of the Am. Acad. of Child Psychiatry*, 21, 187-189.
- Kendell, R.E. (1968), *The classification of depressive illnesses*. Oxford University Press, Londen.
- Kendell, R.E., J.E. Cooper, A.J. Gourlay, J.R.M. Copeland, L. Sharpe en B.J. Gurland (1971), Diagnostic criteria of American and British psychiatrists. *Archives of General Psychiatry*, 25, 123-130.
- Kendell, R.E. (1973), Psychiatric diagnosis: study of how they are made. *British Journal of Psychiatry*, 122, 437-445.
- Kendell, R.E. (1975), *The role of diagnosis in psychiatry*, Blackwell Scientific Publication, Oxford.
- Kovacs, M., en A.T. Beck (1977), An empirical-clinical approach toward a definition of childhood depression. In: J.G. Schulterbrandt en A. Raskin (eds.), *Depression in childhood*, Raven Press, New York, p. 1-25.
- Kovacs, M. (1978), Children's Depression Inventory (CDI). University of Pittsburgh, ongepubl. manuscript.
- Kreitman, N. (1961), The reliability of psychiatric diagnosis. *Journal of Mental Science*, 107, 876-886.
- Ling, W., G. Oftedal en W. Weinberg (1970), Depressive illness in childhood presenting as severe headache. *Am. Journ. of Diseases in Childhood*, p. 120.
- Luria, R.E., en R. Berry (1979), Reliability and descriptive validity of PSE-syndromes. *Archives of General Psychiatry*, 36, 1187-1195.
- Medisch Contact* (1984), 15, 470.
- Meikle, S., en R. Gerritse (1970), A comparison of psychiatric symptom frequency under narrative and check list conditions. *American Journal of Psychiatry*, 127, 379-382.
- Meyer, A. (1951), *The collected papers of Adolph Meyer*. Winters E.E. (ed.). Baltimore, John Hopkins Press, 3, p. 224.
- Poznanski, E.O., en J.P. Zrull (1970), Childhood depression, clinical characteristics of overtly depressed children. *Arch. of General Psychiatry*, 23, 8-15.
- Praag, H.M. van (1982), Een transatlantische visie op diagnostiek van depressies volgens de DSM-III. *Tijdschrift voor Psychiatrie*, 24, 487-496.
- Puig-Antich, J., J.M. Perel, W. Lupatkin, W.J. Chambers, C. Shea, M.-A. Tabrizi en R.L. Stiller (1979a), Plasma levels of imipramine and desmethylimipramine and clinical response in prepubertal major depressive disorder. *Journal of the American Academy of Child Psychiatry*, 616-627.
- Puig-Antich, J., W. Chambers en F. Halpern (1979b), Cortisol hypersecretion in prepubertal depressive illness: a preliminary report. *Psychoneuroendocrinology*, 4, 191-197.
- Puig-Antich, J. (1982), The use of RDC-criteria for major depressive disorder in children and adolescents. *Journ. of the Am. Acad. of Child Psych.*, 291-293.
- Rie, H.E. (1966), Depression in childhood. A survey of some pertinent contributions. *Journal of the American Academy of Child Psychiatry*, 5, 653-685.
- Rochlin, G. (1959), The loss complex. *Psycho-analysis*, Ass. 7, 299-316.
- Rosch, E., C.B. Mervis, W.D. Gray, D.M. Johnson en P. Boyes-Braem (1976), Basic objects in natural categories. *Cognitive Psychology*, 8, 382-439.
- Rutter, M. (1967), A children's behaviour questionnaire for completion by teachers: preliminary findings. *Journal of Child Psychology and Psychia-*

- try, 8, 1-11.
- Rutter, M., P. Graham (1968), The reliability and validity of the psychiatric assessment of the child: I. Interview with the child. *British Journal of Psychiatry*, 114, 563-579.
- Rutter, M. (1978), Diagnostic validity in child psychiatry. *Advances in Biological Psychiatry*, 2, 2-22.
- Saghir, M.T. (1971), A comparison of some aspects of structured and unstructured psychiatric interviews. *American Journal of Psychiatry*, 128, 180-184.
- Smith, D., en W.A. Kraft (1983), DSM-III, Do psychologists really want an alternative? *American Psychologist*, 777-785.
- Spitzer, L.R., en J.L. Fleiss (1974), A re-analysis of the reliability of psychiatric diagnosis. *British Journal of Psychiatry*, 125, 341-347.
- Spitzer, L.R., J. Endicott en E. Robins (1975), Clinical criteria for psychiatric diagnosis and DSM-III. *American Journal of Psychiatry*, 132, 1187-1192.
- Spitzer, L.R., J. Endicott en E. Robins (1978), Research diagnostic criteria. *Archives of General Psychiatry*, 35, 773-782.
- Stengel, E. (1959), *Classification of mental disorders*. Bulletin of WHO, 21, 601-663.
- Toolan, J.M. (1962), Depression in children and adolescents. *American Journal of Orthopsychiatry*, 32, 404-415.
- Verhulst, F.C. (1983), DSM-III in de kinderpsychiatrie. *Tijdschrift voor Psychiatrie*, 25, 127-141.
- Verhulst, F.C., G.W. Akkerhuis en M. Althaus (1985), Mental health in Dutch children (1): a cross-cultural comparison. *Acta Psychiatrica Scandinavica*, suppl. 323 vol. 72.
- Verhulst, F.C., G. Berdenen J. Sanders-Woudstra (1985), Mental health in Dutch children (2): the prevalence of psychiatric disorder and relationship between measures. *Acta Psychiatrica Scandinavica*, suppl. 324 vol. 72.
- Ward, C.H., A.T. Beck en M. Mendelson (1962), The psychiatric nomenclature. *Archives of General Psychiatry*, 7, 198-205.
- Weinberg, W.A., J. Rutman, L. Sullivan, E. Pennick en S.G. Dietz (1973), Depression in children referred to an educational diagnostic center. *Journal of Paediatrics*, 83, 1065-1072.
- Weitzel, W.D., D.W. Morgan, T.E. Guyden en J.A. Robinson (1973), Toward a more efficient mental status examination. *Archives of General Psychiatry*, 28, 215-218.
- Welner, A., J.L. Liss en E. Robins (1974), A systematic approach for making a psychiatric diagnosis. *Archives of General Psychiatry*, 31, 193-196.
- Wing, J.K., J.L.T. Birley, J.E. Cooper, P. Graham en A.D. Isaacs (1967), Reliability of a procedure for measuring and classifying 'present state examination'. *British Journal of Psychiatry*, 113, 499-515.
- Wing, J.K., J.E. Cooper en N. Sartorius (1974), *The measurement and classification of psychiatric symptoms*. Cambridge University Press, Londen.
- Zubin, J. (1967), Classification on the behavior disorders. *Annual Review of Psychology*, 18, 373-406.

Schrijver is kinder- en jeugdpsychiater, werkzaam op de polikliniek Kinder- en Jeugdpsychiatrie van het Sophia Kinderziekenhuis/Academisch Ziekenhuis Rotterdam (hoofd: prof. dr. J.A.R. Sanders-Woudstra).