

**Ontwikkeling van een vaardigheidstest voor het anamnestic interview in de geestelijke gezondheidszorg; betrouwbaarheid en eerste validatie**

S.M. Bögels, H.G. Schmidt en C.P.M. van der Vleuten

**1 Inleiding**

In de geestelijke gezondheidszorg (GGZ) is het anamnestic interview het belangrijkste instrument om een diagnose en behandelplan op te baseren. Onder het anamnestic interview wordt verstaan het afnemen van een ziektegeschiedenis. Het anamnestic interview heeft twee doelen; het creëren van een vertrouwensrelatie met de patiënt en het verkrijgen van relevante informatie om een diagnose op te baseren (Schouten, 1982).

In de opleiding van (toekomstig) hulpverleners in de GGZ, zoals psychiatrisch verpleegkundigen, psychiaters, klinisch psychologen en geestelijk gezondheidkundigen, dient training in anamnestic interviewvaardigheden dan ook een belangrijke plaats in te nemen. Hoewel er door hulpverleners veel geschreven is over hoe men een psychiatrisch interview dient af te nemen (zie o.m. Hersen & Turner, 1985, Shawn, 1988), is er weinig systematisch onderzoek gedaan naar psychiatrische interviewvaardigheden (Rutter & Cox, 1981). Onderzoek naar het meten en trainen van vaardigheden heeft zich vooral gericht op het gebied van counselling (Ford, 1979, Kurtz, Marshall & Banspach, 1985, Marmar, 1990), behavioral assessment (Iwata, Wong, Riordan, Dorsey & Lau, 1982, Veltum & Miltenberger, 1989) en medische vaardigheden (Carrol & Monroe, 1980, Van der Vleuten & Swanson, 1990).

Bij verschillende medische faculteiten in binnen- en buitenland, is inmiddels ruime ervaring opgedaan met het meten van medische vaardigheden. Een vaardigheidstoets bestaat doorgaans uit een aantal "stations", waar de student verschillende klinische taken dient uit te voeren. Hierbij wordt gebruik gemaakt van simulatiepatiënten. Vaardigheidstoetsen worden gebruikt voor het afnemen van een anamnese, lichamelijk onderzoek en communicatievaardigheden. Uit onderzoek naar medische vaardigheden blijkt steeds weer dat de prestatie van de student bij één patiënt een slechte voorspeller is voor de prestatie bij een andere patiënt. Dit verschijnsel wordt wel de "inhoudsspecificiteit van klinische competentie" genoemd. Verwacht kan worden dat ook bij het anamnestic interview van patiënten met geestelijke gezondheidsproblemen, inhoudsspecificiteit een rol speelt.

Het doel van dit onderzoek was een vaardigheidstest te ontwikkelen voor het

anamnestisch interview in de GGZ. Aan het anamnestic interview in de GGZ kunnen twee vaardigheidsaspecten onderscheiden worden (Mumford, Schlesinger, Cuerdo & Scully, 1987). Ten eerste de communicatieve vaardigheden van de interviewer, ook wel *gespreksvaardigheden* genoemd. Ten tweede het vermogen om bepaalde informatie te verzamelen, de inhoudelijke agenda van de interviewer. We noemen dit aspect de feitelijke *anamnesevaardigheden*. Op basis van deze overwegingen werd een lijst ontwikkeld voor het beoordelen van het anamnestic interview in de GGZ. De lijst bevat twee schalen; gespreksvaardigheden (11 items) en anamnesevaardigheden (14 items). De gespreksvaardigheden-schaal meet de kwaliteit en functionaliteit van de communicatieve vaardigheden van de interviewer (Bögels & Kreutzkamp, 1990). De anamnese-schaal meet in hoeverre de interviewer bepaalde onderwerpen navraagt en uitdiept. Alle items werden in gedragstermen gedefiniëerd en heldere criteria voor slecht en goed gedrag werden omschreven. De items werden gescoord op een vijfpuntsschaal van 1 (slecht) tot 5 (goed).

Verschillende vragen werden onderzocht:

- 1) Kunnen de vaardigheden betrouwbaar worden beoordeeld?
- 2) Hoeveel interviews zijn nodig om een betrouwbaar beeld te krijgen van de vaardigheden van de interviewer?
- 3) Zijn de vaardigheden valide, d.w.z. zijn de vaardigheden relevant voor het bewerkstelligen van een effectieve communicatie en het verkrijgen van belangrijke informatie over de ziektegeschiedenis van de patiënt?

## 2 Methode

### 2.1 Proefpersonen

Twintig gevorderde studenten Gezondheidswetenschappen, afstudeerrichting Geestelijke Gezondheidskunde, namen deel aan de studie. Tijdens hun studie werden zij getraind in gespreksvaardigheden en anamnestic interviewvaardigheden. Simulatiepatiënt-contacten maakte deel uit van hun opleiding. Zij waren derhalve bekend met het interviewen van simulatiepatiënten.

### 2.2 Casus

Op basis van gevallen uit de praktijk van de GGZ werden vier zeer verschillende patiëntrollen geschreven; het betrof een jongeman met een psychotisch beeld, een jonge vrouw met verschillende angstklachten, een man van middelbare leeftijd met sociale angsten en een oudere vrouw met depressieve klachten. De rollen werden vertolkt door ervaren simulatiepatiënten.

### 2.3 Instrumenten

*Simulatiepatiëntstisfactie*. De tevredenheid van de simulatiepatiënt met het gesprek werd gemeten met behulp van een op de GGZ aangepaste versie van de

"Patiënt Satisfactie Checklist" (Kraan & Crijnen, 1987). De aangepaste lijst bestond uit 18 beweringen. Op een vijfpuntsschaal vulde de patiënt in, in hoeverre hij het eens was met de bewering.

*Relevante informatie.* De hoeveelheid verzamelde relevante informatie werd als volgt gemeten. Per casus werd een lijst opgesteld met relevante informatie nodig voor de diagnose, differentiaaldiagnose, etiologie en behandelplan (ong. 30 items per casus). Bij elk item diende de beoordelaar aan te geven of de betreffende informatie wel of niet genoemd werd door de patiënt.

*Probleemoplossen.* De analyses van de interviewers van de verschillende casus werden gemeten met behulp van een semi-gestructureerde vragenlijst die in eerder onderzoek naar anamnestiche interviewvaardigheden werd gebruikt (Kraan & Crijnen, 1987). De interviewer diende een aantal open vragen te beantwoorden over diagnose, differentiaaldiagnose, etiologie en behandelplan. Dit instrument geeft een beeld van de cognitieve vaardigheden van de interviewer; hoe interpreteert hij de door hem verzamelde informatie?

#### 2.4 Procedure

Proefpersonen werden gevraagd de vier simulatiepatiënten te interviewen. Voor elk interview hadden zij 30 minuten de tijd. Ze kregen de instructie om 1) een eerste vertrouwensrelatie met de patiënt te vestigen en 2) informatie te verzamelen om een (differentiaal)diagnose, etiologische formulering en behandelingsplan op te baseren. De interviews werden opgenomen op video. Direct na elk interview vulden de proefpersonen de probleemoplossingslijst in. De simulatiepatiënten vulden tegelijkertijd de patiëntsatisfactie vragenlijst in.

Voor het beoordelen van de videobanden werd gekozen voor inhoudsdeskundige beoordelaars; drie klinisch psychologen en een geestelijk gezondheidkundige. Alle beoordelaars waren ervaren in het interviewen van patiënten met geestelijke gezondheidsproblemen en tevens bekend met het trainen van studenten in deze vaardigheden. De beoordelaars werden gezamenlijk getraind in het hanteren van het scoringssysteem totdat een bevredigende betrouwbaarheid ( $r=.80$ ) werd bereikt. Alle videobanden werden vervolgens door twee beoordelaars gescoord.

De hoeveelheid verzamelde relevante informatie werd beoordeeld door twee studenten Geestelijke Gezondheidskunde. Zij werden vooraf getraind in het hanteren van de lijst. Omdat de interbeoordelaarsovereenstemming zeer hoog was (.98) werd elk gesprek slechts door één beoordelaar gescoord. De ene beoordelaar scoorde van alle proefpersonen casus 1 en 2, de andere beoordelaar scoorde van alle proefpersonen casus 3 en 4.

Voor het beoordelen van de probleemoplossing werd eerst een antwoordsleutel ontwikkeld. De vier casus werden voorgelegd aan 8 professionele hulpverleners (andere hulpverleners dan degenen die de videobanden beoordeelden); psychiaters, arts-assistenten aan het eind van hun opleiding tot psychiater en psychologen. Zij hadden ruime ervaring in de diagnostiek van mentale stoornissen. Hen werd gevraagd van elke casus de probleemoplossingslijst in te vullen. Op basis van

deze antwoorden werd een antwoordsleutel met juiste, half-juiste en foute antwoorden gemaakt. Twee geestelijk gezondheidkundigen beoordeelden onafhankelijk van elkaar de probleemoplossingen van de proefpersonen, op basis van de antwoordsleutel.

### 2.5 Statistische analyse

Voor de schatting van de betrouwbaarheid werd gebruik gemaakt van generaliseerbaarheidsanalyse (Cronbach et al., 1972). De interbeoordelaarsbetrouwbaarheid werd geschat door de scores van alle beoordelaarsparen te onderwerpen aan een random beoordelaars-genest-binnen-personen ANOVA ("i:p" design in generaliseerbaarheidstheorie). Variantiecomponenten werden geschat en generaliseerbaarheidscoëfficiënten berekend (voor één beoordeling). Deze betrouwbaarheidsindex kan worden opgevat als een overeenstemmingsmaat tussen beoordelaars, waarbij zowel rangordeverschillen als absolute verschillen tussen beoordelaars zijn verdisconteerd. Voor de totale testbetrouwbaarheid werd een random beoordelaars-genest-binnen-casus x personen ("j:ixp) ANOVA uitgevoerd en variantiecomponenten geschat. Generaliseerbaarheidscoëfficiënten werden vervolgens berekend voor een variabel aantal casus en beoordelaars (volgens een norm-georiënteerd perspectief). Deze betrouwbaarheidsindices zijn op te vatten als de verwachte correlaties tussen deze test en een andere test bestaande uit een willekeurige steekproef van andere beoordelaars, casus en simulatiepatiënten. Voor de validiteitsanalyse werden correlatieve analyses toegepast.

## 3 Resultaten

### 3.1 Variantie-analyse

De verdeling van de scores legde geen beperkingen op aan een ANOVA-analyse. Tabel 1 bevat de variantiecomponenten van de gespreks- en anamnesevaardigheden. De persoonscomponent vertegenwoordigt de ware variantie, het geeft aan hoeveel proefpersonen van elkaar verschillen. De casuscomponent kan opgevat worden als een indicatie voor de moeilijkheidsgraad van de casus. De beoordelaar-binnen-casus component geeft aan in hoeverre beoordelaars (binnen casus) qua strengheid van elkaar verschillen. De persoon-maal-casus component reflecteert in hoeverre personen anders gerangordend worden tussen casus, ook wel het casus-specificiteitsprobleem genoemd. De laatste component betreft de (residuele) error.

Bij de gespreks- en anamnesevaardigheden waren de pure persoonsvariantie en de persoon-maal-casusvariantie verreweg de grootste variantiebronnen. Bij de anamnesevaardigheden lag dit anders; hier waren de persoon-maal-casus en de pure casusvariantie de grootste variantiebronnen. Zowel bij de gespreks- als anamnesevaardigheden was de algemene error term een aanzienlijke variantiebron. De variantie die het gevolg is van verschillen in strengheid tussen beoordelaars was zeer laag.

### **3.2 Interbeoordelaarsbetrouwbaarheid**

In tabel 2 worden de interbeoordelaarsbetrouwbaarheden gegeven voor de gespreks- en anamnesevaardigheden en probleemoplossen per casus en voor de totale testafname. Er werd een hoge interbeoordelaarsbetrouwbaarheid gevonden voor de gespreksvaardigheden, voor de anamnesevaardigheden was de betrouwbaarheid iets lager. De interbeoordelaarsbetrouwbaarheid over de probleemoplossing was eveneens hoog.

### **3.3 Totale toetsbetrouwbaarheid**

In tabel 3 zijn de generaliseerbaarheidscoëfficiënten weergegeven als functie van het aantal casus en het aantal beoordelaars. Uitgaande van 4 casus en 2 beoordelaars, het design van het huidige experiment, werd wat betreft de gespreksvaardigheden een niet onredelijke betrouwbaarheid bereikt (.74). Voor de anamnesevaardigheden zou een veel groter aantal casus nodig zijn (10) om tot een aanvaardbare betrouwbaarheid te komen.

### **3.4 Validiteit**

De validiteit van de gemeten vaardigheden wordt onder meer bepaald door hun samenhang met de doelen van een anamnestic interview; het vestigen van een vertrouwensrelatie met de patiënt en het verkrijgen van relevante informatie om diagnose en behandelplan op te baseren. In tabel 4 worden de correlaties tussen de verschillende variabelen gegeven.

De correlaties gingen in de verwachte richting. Gespreksvaardigheden vertoonden een sterk verband met de satisfactie van de simulatiepatiënt met het gesprek, en enig verband met de hoeveelheid verzamelde relevante informatie. De anamnesevaardigheden vertoonden een sterke samenhang met hoeveelheid verzamelde relevante informatie, en enige samenhang met de kwaliteit van de probleemoplossing en patiëntsatisfactie.

## **4 Conclusies en discussie**

In dit onderzoek werd een instrument ontwikkeld en onderzocht om anamnestiche vaardigheden voor het domein van de GGZ te meten. Interbeoordelaarsbetrouwbaarheden waren bevredigend, en vergelijkbaar met de waarden zoals die in onderzoek naar medische interviewvaardigheden doorgaans gevonden worden (Van der Vleuten & Swanson, 1990). Een bedreiging voor de totale testbetrouwbaarheid werd gevormd door de zg. casusspecificiteit van de vaardigheden; iets wat in bijna alle studies naar medische vaardigheden wordt gevonden. Bovendien bleken de feitelijke anamnesevaardigheden gevoelig voor verschillen in moeilijkheidsgraad van de casus. Als gevolg van deze bevindingen generaliseerden de gespreks- of communicatievaardigheden beduidend beter over casus heen dan de anamnesevaardigheden. Dit valt als volgt te verklaren. Het afnemen van een anamnese of ziektegeschiedenis bij een patiënt vereist specifieke kennis van bepaalde diagnos-

tische beelden. Hoewel alle proefpersonen in dit onderzoek de verschillende diagnostische beelden tijdens het verplichte curriculum hadden behandeld, kan verwacht worden dat zij op basis van bijvoorbeeld "vrije" studie-onderdelen of eigen interesse, verschilden in hun kennis van bepaalde beelden.

De validiteit van de test werd ondersteund door de correlatieve analyses; gespreksvaardigheden vertoonden een sterke samenhang met patiëntstatisfactie en enige samenhang met verzamelde relevante informatie, terwijl anamnesevaardigheden omgekeerd een sterke samenhang met verzamelde relevante informatie en enige samenhang met patiëntstatisfactie vertoonden. Dit patroon kan als volgt geïnterpreteerd worden. Gespreksvaardigheid is vooral belangrijk voor een goed contact met de patiënt; patiëntstatisfactie is hiervoor een indicatie. Vaardigheid in het afnemen van een anamnese is vooral van belang om relevante diagnostische informatie te verzamelen. Er werd tot slot slechts enig verband gevonden tussen anamnesevaardigheid en de kwaliteit van de probleemoplossing. Dit kan worden verklaard uit het feit, dat of het verzamelen van relevante diagnostische informatie ook werkelijk leidt tot een adequate probleemoplossing, vooral afhankelijk is van cognitieve vaardigheden van de interviewer.

De hier ontwikkelde test lijkt bruikbaar voor het meten van anamnestiche interviewvaardigheden bij geestelijke gezondheidsproblemen. Voor toetsdoeleinden in onderwijssituaties is het echter nodig een groter aantal interviews af te nemen dan in dit onderzoek het geval was om een betrouwbaar beeld te krijgen van de *anamnesevaardigheden*. Voor de gespreksvaardigheden kan worden volstaan met een viertal interviews.

Zowel bij de variantie-analyses als bij de correlatieve analyses werden belangrijke en goed interpreteerbare verschillen gevonden tussen de gespreksvaardighedenschaal en de anamneseschaal. Het onderscheiden van deze twee aspecten van anamnestiche interviewen lijkt daarom zinvol. Met behulp van factor-analytisch onderzoek dient nog te worden nagegaan of deze twee aspecten van anamnestiche interviewen ook werkelijk onderscheidbare factoren zijn.

Verder onderzoek naar dit instrument dient bovendien gericht te worden op een andere validiteitsvraag, namelijk of het instrument gevoelig is voor voortgang in vaardigheden. Op het gebied van onderwijs in anamnestiche vaardigheden voor de GGZ is het belangrijk het effect van training in het algemeen en van verschillende trainingsmethoden in het bijzonder te onderzoeken.

## Literatuur

- Bögels, S.M. en Kreutzkamp, R. (1990) Effecten van een training in basale gespreksvaardigheden. *Tijdschrift voor Onderwijsresearch*, 15, 4, 201-214.
- Carrol, J.G. & Monroe, J. (1980) Teaching clinical interviewing in the health professions. *Evaluation of the Health Professions*, 3, 21-45.
- Crick, J. & Brennan, R. (1983) *Manual for GENOVA: A Generalized Analysis of Variance System*. Iowa: American College Testing Program.
- Cronbach, L.J., Goldine, C.G., Nanda, H. & Rajaratnam, N. (1972) *The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles*. New York: Wiley & Sons.
- Ford, J.D. (1979) Research on training counselors and clinicians. *Review of Educational Research*, 49, 87-130.
- Hersen, M. & Turner, S.M. (1985) *Diagnostic interviewing*. New York, Plenum Press.
- Iwata, B.A., Wong, S.E., Riordan, M.M., Dorsey, M.F. & Merle, M.L. (1982) Assessment and training of clinical interviewing skills. *Journal of Applied Behavior Analysis*, 191-203.
- Kraan, H.F. and Crijnen, A.A.M. (1987) *The Maastricht History-taking and Advice Checklist: Studies of instrumental utility*. Proefschrift, Amsterdam: Velden van den Hazelaar.
- Marmar, C.R. (1990) Psychotherapy process research: progress, dilemmas and future directions. *Journal of Consulting and Clinical Psychology*, 3, 265-272.
- Mumford, E., Schlesinger, H., Cuedon, T. & Scully, J. (1987) Ratings of videotaped simulated patients and four other methods of evaluating psychiatric clerkship. *American Journal of Psychiatry*, 3, 316-322.
- Rutter M. & Cox, A. (1981) Psychiatric interviewing techniques: I Methods and measures. *British Journal of Psychiatry*, 138, 273-282.
- Schouten, J. (1982) *Anamnese en Advies*. Alphen aan den Rijn. Brussel: Stafleu.
- Shawn, S.C. (1988) *Psychiatric Interviewing: The Art of Understanding*. Philadelphia: W.B. Saunders.
- Veltum, G.L. & Miltenberger, R.G. (1989) Evaluation of a self-instructional package for training initial assessment interviewing skills. *Behavioral Assessment*, 11, 165-177.
- Vleuten, C.P.M. van der en Swanson D.B. (1990) Assessment of clinical skills with standardised patients: state of art. *Teaching and Learning in Medicine*, 2: 58-76.

Tabel 1  
 Variantiecomponenten van de gespreks- en anamnesevaardigheden (standard error  
 tussen haakjes)

Bron <sup>1</sup>	Variantiecomp. gespreksvh.	Perc. v.d. tot. var.	Variantiecomp. anamnesevh.	Perc. v.d. tot. var.
P	19.84 (8.33)	34%	7.08 (3.88)	15%
C	4.61 (4.10)	8%	13.56 (9.84)	28%
B:C	0.92 (0.87)	2%	1.85 (1.41)	4%
PC	21.59 (5.12)	37%	14.48 (3.87)	30%
PB:C	11.51 (1.84)	20%	11.84 (1.90)	24%

Tabel 2  
 Interbeoordelaarsbetrouwbaarheid voor gespreks- en anamnesevaardigheden en  
 probleemoplossing per casus en voor de totale testafname

	gespreksvh.	anamnesevh.	probleemopl.
Psychose	.85	.78	.77
Agorafobie	.73	.71	.78
Sociale angst	.83	.48	.81
Depressie	.72	.70	.86
Totaal	.79	.72	.84

<sup>1</sup> P is personen, C is casus, B is beoordelaar



Tabel 3

Generaliseerbaarheidscoëfficiënten voor de gespreks- en anamnesevaardigheden, als functie van het aantal casus en het aantal beoordelaars

	Aantal casus	Eén beoordelaar	Twee beoordelaars
Gespreks- vaardigheden	1	.38	.42
	2	.55	.59
	4	.71	.74
	10	.86	.88
Anamnese- vaardigheden	1	.21	.26
	2	.35	.41
	4	.52	.58
	10	.73	.78

Tabel 4

Correlatiematrix van de gespreks- en anamnesevaardigheden, hoeveelheid verzamelde relevante informatie, probleemoplossing en simulatiepatiënt-satisfactie<sup>2</sup>

	gespreksvh.	anamnesevh.	rel. info	probleemopl.	satisfactie
gespreksvh.	-	.71***	.37	.07	.68***
anamnesevh.		-	.57**	.45*	.45*
rel. info			-	.40*	.43*
probleemopl.				-	.40*

<sup>2</sup> \* = p < .05  
 \*\* = p < .005  
 \*\*\* = p < .001