

# De kwaliteit van kwalitatief onderzoek

M. H. VAN IJZENDOORN EN  
S. MIEDEMA

*Vakgroep Wijsgerige en Empirische  
Pedagogiek  
Rijksuniversiteit te Leiden*

## Samenvatting

*In deze bijdrage bespreken we criteria voor de kwaliteit van kwalitatief onderzoek. Betrouwbaarheid en validiteit beschouwen we als verbijzonderingen van de omvattende eis 'intersubjectieve navolgbaarheid', waaraan ook kwalitatief onderzoek moet voldoen.*

*In de fase van data-verzameling zal doorgaans met 'argumentatieve' betrouwbaarheid moeten worden volstaan; in de fase van data-analyse moet ook aandacht aan de 'technische' betrouwbaarheid worden besteed. Validatie van kwalitatief onderzoek kan langs drie wegen plaatsvinden: 'communicatieve', 'ideologie-kritische', en 'empirische' validatie. De vanzelfsprekende geldigheid, c.q. 'levensechtheid' van kwalitatief onderzoek wordt hierdoor gerelativeerd.*

## 1 Inleiding

Als wetenschapsfilosofische reflectie op onderzoek één ding heeft duidelijk gemaakt, dan is het wel de argumentatieve grondslag van kennisverwerving. Wetenschappelijke kennis is het – steeds voorlopige – resultaat van discussies onder vakgenoten over de merites van het onderzoek waaruit de kennisclaims zijn voortgevloeid. De rapportage van het onderzoek moet daarom voor discussie vatbaar zijn. Dit betekent dat aan ieder onderzoek de eis van intersubjectieve navolgbaarheid mag worden gesteld. Hiermee wordt bedoeld dat de onderzoeker de plicht heeft zijn activiteiten en keuzes zo helder te beschrijven dat zijn collega's het onderzoek in principe stap voor stap kunnen reconstrueren c.q. repliceren. Ongeacht de 'wetenschappelijke traditie' waarin men zich be-

vindt (Hetebrij & Wardekker, 1986), en ongeacht verschillen in objectvisies en epistemologische criteria is de gemeenschappelijke noemer voor iedere wetenschappelijke activiteit haar intersubjectieve navolgbaarheid door collega's en/of consumenten van het onderzoek.

'Navolgbaarheid' is om verschillende redenen kenmerkend voor wetenschapsbeoefening. Allereerst bestaat de wetenschappelijke discussie bij gratie van de navolgbaarheid van onderzoek. Het formuleren van kritiek en het voortbouwen op prestaties uit het verleden is alleen mogelijk als het onderzoeksverslag meer voorstelt dan de allerindividueelste expressie van de allerindividueelste emotie. Esoterische onderzoeksproducten, ontsproten aan de rijke maar onnavolgbare fantasie van een wetenschapper en interpretaties op intuïtie gebaseerd, immuniseren tegen kritiek en kunnen geen bouwsteen voor vervolgonderzoek vormen. Wetenschapsbeoefening is bij uitstek een 'onpersoonlijke', collectieve en publieke aangelegenheid (Popper, 1972).

Ten tweede staat 'navolgbaarheid' garant voor de mogelijkheid (on-)bewuste vertekening van onderzoeksresultaten en bedrog op het spoor te komen. Vooral in de natuur- en medische wetenschappen ontkomt geen enkel belangrijk experiment aan replicatie en het is juist in die wetenschappen dat op gezette tijden zelfs vooraanstaande onderzoekers en onderzoeksinstituten op bedrog worden betrapt (Broad & Wade, 1982), 'Navolgbaarheid' vervult hier de functie van waarschuwing tegen al te lichtzinnige interpretatie, c.q. vertekening van onderzoeksuitkomsten. Maar ook in de gedrags- en sociale wetenschappen zijn voorbeelden van bedrog te vinden die door 'navolging' van het onderzoeksproces zijn opgespoord. Zo is nogal wat 'onderzoek' van Cyril Burt door de mand gevallen, omdat herberekening van zijn resultaten een al te perfecte pasvorm tussen theorie en data liet zien (cfr. Medawar, 1977).

Hoewel intersubjectieve navolgbaarheid uitgangspunt voor iedere vorm van wetenschapsbeoefening moet zijn, kan niet altijd

op dezelfde wijze aan dit fundamentele criterium worden voldaan. Afhankelijk van onderzoeksmethode en -object kunnen meer of minder sterke garanties voor navolgbaarheid gegeven worden. De sterkste garantie voor navolgbaarheid, verbijzonderd als betrouwbaarheid (zie verder) is natuurlijk die, waarbij het gehele onderzoeksproces door een tweede onderzoeker of onderzoeksgroep volledig gerepliceerd kan worden. Wellicht kan hier alleen bij laboratoriumonderzoek sprake van zijn. Zo is er bijvoorbeeld in de discussies rondom de Mead-Freeman polemiek op gewezen dat er weliswaar sprake is van vertekening bij de resultaten die Mead presenteerde, maar dat de historische ontwikkelingen tussen 1925 en 1948 (het begin van hun respectievelijke studies), en de verschillende belangstelling en theoretische perspectieven (cfr. Nardi, 1984; Scheper-Hughes, 1984) van de onderzoekers het onmogelijk maken bij Freeman van een zuivere replica te spreken. Freemans resultaten kunnen dan ook beter als argumenten tegen Meads claim op navolgbaarheid van haar onderzoek gezien worden.

Hoe het ook zijn, tekortkomingen in de gevolgde methode van onderzoek kunnen alleen aangetoond worden als deze methode zo nauwkeurig mogelijk beschreven, en dus minstens in gedachte navolgbaar is gemaakt. Met dit laatste is de minimale waarborg geformuleerd voor de intersubjectieve navolgbaarheid van onderzoek. Tussen de strikte en de minimale vorm van navolgbaarheid ligt een vorm van navolgbaarheid waarbij voor gedeelten van het onderzoeksproces de replicerbaarheid kwantitatief wordt aangegevoerd, terwijl voor andere aspecten argumenten worden aangevoerd dat het onderzoeksproces niet volstrekt willekeurig is verlopen en niet louter afhankelijk is geweest van de persoon van de onderzoeker.

We zullen in onderstaande de fundamentele eis van intersubjectieve navolgbaarheid toepassen op kwalitatief onderzoek. Daarbij zullen twee kwaliteitscriteria centraal staan, te weten het bekende koppel 'betrouwbaarheid' en 'validiteit'. Deze criteria worden als verbijzonderingen opgevat van het omvattende criterium 'intersubjectieve navolgbaarheid'. Bij betrouwbaarheid gaat het om (virtuele) herhaalbaarheid van het onderzoeksproces langs ongeveer dezelfde weg als oorspronke-

lijk werd gevolgd. Bij validiteit ofwel geldigheid gaat het om de reproductie van het onderzoeksproduct langs verschillende wegen. Het zal duidelijk zijn dat aan deze verbijzonderingen van 'intersubjectieve navolgbaarheid' een epistemologische visie ten grondslag ligt, waarin geen ruimte is voor beslissende confrontaties tussen kennisclaims en een daarvan onafhankelijke werkelijkheid. We kunnen in dit korte bestek deze visie helaas niet uitwerken. Overeenkomsten en verschillen tussen kwalitatief en kwantitatief onderzoek kunnen tegen deze achtergrond worden gespecificeerd.

## 2 De betrouwbaarheid van kwalitatief onderzoek

We kunnen een onderscheid maken tussen 'technische' en 'argumentatieve betrouwbaarheid'. *Technische betrouwbaarheid* duidt op de exacte vaststelling van de mate van herhaalbaarheid van een bepaald facet van het onderzoek. Twee of meer personen leggen onafhankelijk van elkaar eenzelfde traject van de onderzoeksweg af – bijvoorbeeld het categoriseren van interviewprotocollen of het coderen van gedrag, vastgelegd op video. De mate van overeenstemming is dan een indicatie voor de intersubjectiviteit van de gevolgde werkwijze. Het is gebruikelijk 'technische' betrouwbaarheid kwantitatief te bepalen. *Argumentatieve betrouwbaarheid* berust op het gegeven dat niet in ieder onderzoek de technische betrouwbaarheid exact kan worden bepaald, terwijl er toch niet-kwantitatieve indicaties gegeven kunnen worden voor de herhaalbaarheid van het onderzoeksproces. De argumentatie heeft dan bijvoorbeeld betrekking op waarborgen voor herhaalbaarheid die de onderzoeker tijdens het verzamelen van de gegevens heeft ingebouwd. Te denken valt hierbij aan explicatie van het begrippenkader waarmee de observaties zijn verricht, of aan de weergave van de interviewleidraad aan de hand waarvan halfgestructureerde interviews zijn gehouden. Ook de letterlijke weergave van gedeelten van de interviews of gedetailleerde beschrijvingen van situaties en interacties die werden geobserveerd kan het vertrouwen in de herhaalbaarheid van het onderzoek verhogen (zie Goetz & LeCompte,

1984; Janssens, 1985). Het is duidelijk dat argumentatieve betrouwbaarheid nooit tot sluitende garanties voor de feitelijke herhaalbaarheid kan leiden en zeker niet gelijkgesteld mag worden aan het resultaat van feitelijke herhaling. Wel is het mogelijk een discussie te voeren over de houdbaarheid van de argumenten die ten gunste van de herhaalbaarheid naar voren zijn gebracht en wellicht kan de uitkomst van deze discussie ook verschillend gewicht voor verschillende onderdelen van het onderzoeksverslag impliceren. Buitenstaanders zijn in staat sterke en zwakke kanten van het onderzoek naar waarde te schatten.

Betrouwbaarheid heeft betrekking op verschillende fasen in het onderzoeksproces. Niet alleen het verzamelen van de gegevens maar ook het analyseren van de data kan aan vertekende invloeden onderhevig zijn en voor beide fasen geldt de eis van (virtuele) herhaalbaarheid. Afhankelijk van de opzet van de studie is het gewenst de vraag naar de technische of argumentatieve betrouwbaarheid te stellen ten aanzien van de fase van het verzamelen van gegevens of de analysefase. Er zijn dan ideaaltypisch vier alternatieven:

Schema 1 *Kwaliteitscriteria voor verschillende onderzoeksfasen*

Fase	Betrouwbaarheid	
	Technisch	argumentatief
verzamelen	1	3
analyseren	2	4

Als gegevens op gestandaardiseerde wijze zijn verzameld dan is er geen enkele reden om het vaststellen van de technische betrouwbaarheid na te laten. Doorgaans wordt dit de interobserverbetrouwbaarheid genoemd. Als de gegevens via gestructureerde interviews zijn verzameld dan laat zich de vraag naar de technische betrouwbaarheid interpreteren als de vraag naar de invloed van de interviewer(s) op de data. In de kwantitatieve traditie is het gebruikelijk de technische betrouwbaarheid voor de fase van verzamelen van de data te bepalen, hoewel hierop ook heel belangrijke uitzonderingen zijn (zie Lamb, Thompson, Charnov & Gardner, 1985).

Kwalitatief onderzoek verschilt van kwantitatief onderzoek juist op het punt van standaardisatie van de gegevensverzameling: kenmerkend voor kwalitatief onderzoek is de 'open' vorm van verzamelen van gegevens die de onderzoeker gelegenheid geeft ook onverwachte gebeurtenissen en opinies te registreren die niet zouden hebben gepast in een vooraf geprepareerd keurslijf van observatie-categorieën (Van IJzendoorn & Van Vliet-Visser, 1985). Dit heeft echter tot gevolg dat het vrij zinloos is te streven naar technische betrouwbaarheid voor deze fase van kwalitatief onderzoek. Percentages overeenstemming of andere kwantitatieve betrouwbaarheidsmaten zijn gebaseerd op gelijke wijze van gegevensverzameling. Hiermee strookt de open, semi-gestructureerde kwalitatieve benadering niet. Vandaar dat juist hier de argumentatieve betrouwbaarheid van belang is. De onmogelijkheid technische betrouwbaarheid te bepalen mag geen vrijbrief vormen voor het negeren van de betrouwbaarheidsvraag. Ook de kwalitatief onderzoeker zal zich moeten verantwoorden tegenover het forum voor de wijze waarop hij zijn gegevens heeft bijeengesprokkeld. Zoals gezegd zijn hier meerdere typen argumenten mogelijk: al sinds jaar en dag wordt in de kwalitatieve traditie de explicatie van het 'voorbegrip' als startpunt van een hermeneutische cirkel noodzakelijk geacht: wat triviaal maar niet minder belangrijk is de beschrijving van de observatie- of interviewleidraad. Daarnaast wordt tegenwoordig steeds meer gepleit voor de inzet van meerdere onderzoekers van wie de produkten met elkaar kunnen worden vergeleken. Ook een onderzoeksoepzet met ingebouwde 'replicaties' zoals een  $n=5$  studie (Van IJzendoorn & Van Vliet-Visser, 1985) kan een waarborg tegen willekeur en kapitalisatie op onnavolgbare waarnemingen zijn. Kortom: voor kwalitatief onderzoek dat door open verzameling van gegevens wordt gekenmerkt, is argumentatieve betrouwbaarheid (3) paradigmatisch, voor kwantificerend onderzoek is dat de technische betrouwbaarheid (1).

Is de fase van verzameling van gegevens achter de rug dan moet het ruwe materiaal nog geanalyseerd worden. De vraag is dan of tijdens deze fase voor alternatief 2 of 4 uit Schema 1 gekozen moet worden. Als interviews gestandaardiseerd zijn afgenomen of

als de observaties gestructureerd hebben plaatsgevonden dan doet zich in de analyse-fase de vraag voor hoe dit materiaal zo bewerkt kan worden dat een spaarzame beschrijving mogelijk wordt. Dat betekent doorgaans het zoeken naar een onderliggende structuur van waaruit de meeste ruwe gegevens begrijpelijk, c.q. 'verklaarbaar' zijn. Deze speurtocht die zowel in de kwantitatieve, als ook in de kwalitatieve traditie (zie Spradley, 1980) plaatsvindt, kan aan subjectieve vertekening in de slechte zin van het woord onderhevig zijn. Categorieën-systemen kunnen slordig worden toegepast, de ontwikkeling van concepten die de onderliggende structuur weerspiegelen kan louter op projectie van de onderzoeker berusten in plaats van te resulteren uit een navolgbare interactie tussen het voorbegrip van de onderzoeker en zijn materiaal. Ook in deze fase moet daarom de betrouwbaarheidsvraag gesteld worden.

In de kwantitatieve traditie wordt deze vraag doorgaans beantwoord door de inter-codeurbetrouwbaarheid te bepalen, afgezien van geavanceerdere methoden om de invloed van codeurs op de gegevens te bepalen (generaliseerbaarheidsstudies, zie Tavecchio, 1977). Het ligt voor de hand in de analyse-fase de technische betrouwbaarheid vast te stellen, als datzelfde ook in de daaraan voorafgaande fase is gebeurd (1→2). Men verliest onnodig precisie, ook in de fase van rapportage als hier louter argumentatieve betrouwbaarheid wordt bepaald. In het kwalitatieve onderzoek daarentegen doen zich wel degelijk twee alternatieven voor. Hoewel doorgaans ook in de analyse-fase slechts de argumentatieve betrouwbaarheid wordt bepaald (zie Spradley, 1980) is het in veel gevallen mogelijk juist in deze fase minstens voor enkele facetten van het analyseren de technische betrouwbaarheid vast te stellen en geen concessie te doen aan de precisie. Dit is bijvoorbeeld in onderzoek van Komter (1985) en Van IJzendoorn & Van Vliet-Visser (1985) gebeurd. Als het verzamelde materiaal op enigerlei wijze is geobjectiveerd, bijvoorbeeld integraal is uitgeschreven, dan is er geen enkele reden af te zien van een demonstratie van de herhaalbaarheid van het analyse-proces of althans onderdelen daarvan. Het indelen van interview- of observatieprotocollen in abstracte categorieën waar-

van men vermoedt dat daarmee de onderliggende structuur zichtbaar wordt gemaakt, kan heel wel op een technisch betrouwbare wijze gebeuren. In feite gaat men dan de weg van 3→2.

De vraag is of hierdoor de precisie en dus navolgbaarheid van het onderzoek als geheel wordt verhoogd: wat is de zin van een 'objectieve' analyse op 'subjectief' verzamelde gegevens? Het antwoord moet luiden dat vertekening en willekeur in verschillende onderzoeksfasen cumuleren: de navolgbaarheid erodeert steeds verder naarmate niet alleen in de fase van de gegevensverzameling maar ook in de analyse-fase met zwakkere argumenten voor betrouwbaarheid moet worden gewerkt. De erosie wordt een halt toegevoerd door in de analyse-fase wel degelijk technische betrouwbaarheid vast te stellen. Het gevolg hiervan is dat de rapportage op een hoger niveau van precisie kan plaatsvinden, waardoor het onderzoek beter navolgbaar zal zijn. De fout die echter onder geen beding gemaakt mag worden, is menen dat bij een bevredigende technische betrouwbaarheid van de analyse-fase de open, kwalitatief verzamelde gegevens nu ook langs statistische weg bewerkt moeten worden. Illustratie hiervoor is de dissertatie van Komter (1985), die zich ook zelf kennelijk niet aan de macht van het getal heeft kunnen onttrekken (cfr. Komter, 1983, 210).

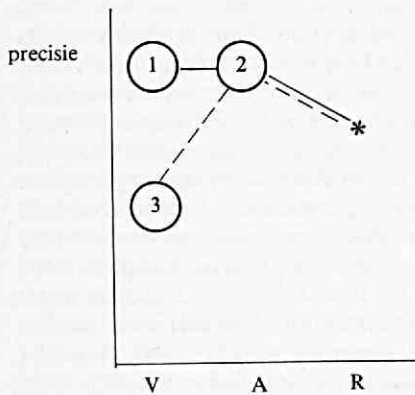
Het voorafgaande kan nog wat verhelderd worden met behulp van grafieken (zie Figuur 1). Ook het onderscheid tussen kwalitatief en kwantitatief onderzoek kan hierdoor worden verduidelijkt. In grafiek a is kwantitatief onderzoek weergegeven, in grafiek b kwalitatief onderzoek.

Gangbaar kwantitatief onderzoek is in grafiek a weergegeven als een ononderbroken lijn: de precisie van de rapportage moet hier enigszins onder het niveau van precisie van de daaraan voorafgaande fasen liggen omdat 'onzuiverheid', hoe gering ook, cumuleert. De stippellijn duidt een suboptimaal onderzoekstype aan dat vanwege onduidelijkheid over het verschil tussen interobservator- en intercodeurbetrouwbaarheid nogal eens te zien valt. Lamb et al. (1985) hebben aangetoond dat de fameuze Baltimore-studie van Ainsworth aan dit tekort laboreert. Het is in zo'n geval onjuist de rapportage op een al te hoog niveau van pre-

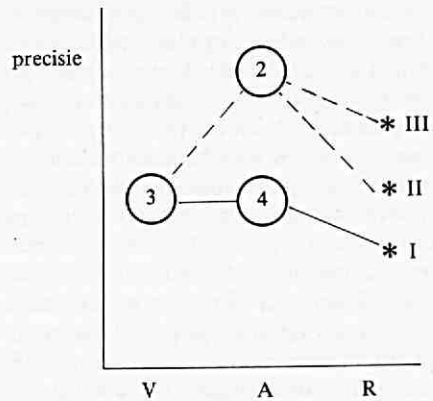


Figuur 1 *Betrouwbaarheid in verschillende fasen van kwantitatief en kwalitatief onderzoek.*

a: *kwantitatief onderzoek*



b: *kwalitatief onderzoek*



legenda:

- V = verzameling van gegevens
- A = analyse
- R = rapportage

- 1, 2: technische betrouwbaarheid
- 3, 4: argumentatieve betrouwbaarheid

cisie te situeren en aldus de schijn van navolgbaarheid te wekken. Ook hier geldt dat de rapportage niet een hogere mate van precisie, en dus navolgbaarheid, mag suggereren dan op grond van de betrouwbaarheid van de verzameling en analyse van de gegevens toelaatbaar is.

Gangbaar kwalitatief onderzoek is in grafiek b door een ononderbroken lijn aangegeven. Het is duidelijk dat de precisie van de rapportage minder groot is dan bij kwantitatief onderzoek (I). Daar staan allerlei voordelen tegenover die we hier echter niet zullen bespreken (zie Spradley, 1980). Wel is het mogelijk de precisie van het onderzoek nog wat te verhogen door bijvoorbeeld in de analyse-fase een intercodeurbetrouwbaarheid te bepalen (zie Van IJzendoorn & Van Vliet-Visser, 1985). In dat geval cumuleert het tekort aan precisie niet en kan de rapportage op een wat hoger niveau van nauwkeurigheid plaatsvinden (II). Het gevaar bestaat echter dat onderzoekers zich door de hoge mate van navolgbaarheid in de analyse-fase laten verleiden tot een rapportage op een hoog niveau van precisie, zonder rekening te houden met de wijze waarop de gegevens zijn verzameld (III). Een voorbeeld hiervan is het al genoemde onderzoek van Komter (1985), die weliswaar op zorgvuldige wijze de be-

trouwbaarheid van het coderen tijdens de analyse-fase bepaalde maar dit ten onrechte aanleiding vond op de bekende kwantitatieve wijze te rapporteren met behulp van frequentie-tabellen en significantie-toetsen. In dat geval vindt een onjuiste mixture van 'open', dus kwalitatieve gegevensverzameling, en kwantitatieve rapportage plaats. Deze hybride vorm van onderzoek doet geen recht aan de beperkingen en mogelijkheden van kwalitatief onderzoek.

### 3 De validiteit van kwalitatief onderzoek

Niet zelden wordt de vraag naar de validiteit van de resultaten uitsluitend aan kwantitatief onderzoekers gesteld. Kwantitatief onderzoek mag dan misschien wel heel precies zijn (grote betrouwbaarheid), aan de geldigheid wordt met name door kwalitatieve onderzoekers nogal eens getwijfeld. Kwalitatief onderzoek zou daarentegen dicht bij de praktijk blijven, mede door het open karakter van het dataverzamelen, en daardoor bijna per definitie een hoge mate aan validiteit kunnen claimen. Toch is ook de kwalitatief onderzoeker niet van de verplichting ontslagen, argumenten aan te dragen voor de geldigheid van zijn onderzoeksresultaten. De

kwalitatieve methode van onderzoek garandeert uiteraard niet zonder meer een onvertekkende kijk op 'de authentieke werkelijkheid'. In de kwalitatieve onderzoekstraditie kunnen verschillende validatie-strategieën (cfr. Terhart, 1981; 1982) worden onderscheiden, waarbij tevens een bepaald antwoord wordt gegeven op de vraag naar de relatie tussen kwalitatief en kwantitatief onderzoek.

Een eerste strategie is de *communicatieve validatie*. De onderzoeker tracht de geldigheid van zijn interpretatie met betrekking tot de betekenis-kaders van de onderzochten aan te tonen, door met hen naar overeenstemming over het realiteitsgehalte van zijn interpretatie te streven. De discussie tussen onderzoeker en onderzochten moet tot consensus over de adequaatheid van zijn interpretatie van een interview of participerende observatie leiden. In feite beslissen de onderzochten zelf over de communicatieve validiteit van het onderzoek. Hun betekenisverlening van het eigen handelen is de toetssteen. De – veelal impliciete – aanname over het object van onderzoek (de onderzochte), luidt dat diens handelen voor hemzelf volstrekt transparant is. Zijn betekenisverlening is een geldige representatie van de sociale werkelijkheid. De consequentie hiervan is dat slechts die uitspraken valide zijn, waarmee de betrokkene zelf instemt – hetzij spontaan of met behulp van de onderzoeker. De interpretaties van de onderzoeker zijn daarmee ook voor de 'proefpersonen' navolgbaar geworden: Het onderzoeksprodukt is langs verschillende wegen gereproduceerd – er is sprake van convergentie tussen de visies van onderzoekers en onderzochten, waarbij uiteraard niet is uitgesloten dat beide partijen een onjuiste kijk op de zaak zouden kunnen hebben.

Deze vorm van communicatieve validatie kan aangevuld, bijgesteld of verworpen worden met behulp van een ideologiekritische of empirische objectivering. Na een geslaagde communicatieve validatie kunnen mogelijke misinterpretaties of inconsistenties in de uitspraken van de onderzochten worden geanalyseerd. In de ideologiekritische variant wordt een maatschappij theorie gebruikt om te trachten de voorwaarden waaronder misinterpretaties mogelijkerwijze zijn ontstaan, bloot te leggen. Hier is dan sprake van *ideologie-kritische validatie*. De resultaten

van deze validatie kunnen overigens weer worden voorgelegd aan de betrokkenen en wellicht een leerproces bij hen bewerkstelligen. Dit laat onverlet dat ook door sommige kwalitatieve onderzoekers alles op de kaart van de objectieve, ideologiekritische validatie gezet wordt. De door Oevermann (cfr. Oevermann et al., 1983) uitgewerkte objectieve hermeneutiek is hiervan een voorbeeld.

De communicatief gevalideerde inzichten kunnen ook het uitgangspunt vormen voor empirisch onderzoek in verklarende en voorspellende zin. We hebben hier met een empirische objectiveringsvariant te maken. Kwalitatief en kwantitatief onderzoek worden gekoppeld. In plaats van op indirecte wijze – via waarneembaar gedrag – de hand te leggen op achterliggende processen, regelmatigheden of naïeve theorieën, wordt 'direct' toegang tot de sociale werkelijkheid gezocht. En het communicatief gevalideerde resultaat hiervan wordt vervolgens empirisch getoetst, doordat nagegaan wordt of de 'naïeve' theorieën ook de facto het handelen van de betrokkenen sturen, c.q. correct beschrijven. In deze opzet is dan sprake van *empirische validatie* die vooral is bedoeld om het wetenschappelijk forum te overtuigen van de geldigheid van het resultaat. Overigens dreigt in dit geval de communicatieve validatie ondergeschikt te worden aan de empirische validatie, tenzij ook de resultaten van de empirische validatie weer systematisch naar de betrokkenen worden teruggekoppeld. *Traditioneel* kwantitatief onderzoek wordt gekenmerkt door het ontbreken van communicatieve validatie of door de afwezigheid van bovengenoemde terugkoppeling van communicatief en empirisch gevalideerde kennis. Vindt deze terugkoppeling wel plaats dan kan dat aanleiding geven tot communicatieve validatie van de tweede orde: onderzoekers en onderzochten buigen zich nogmaals over de communicatief en empirisch gevalideerde onderzoekresultaten, om te kijken of ze tot consensus over de adequaatheid ervan kunnen geraken. Lukt dit, dan mag hierin een extra garantie voor de geldigheid van de verworven kennis gezien worden.

De parallellie met het betrouwbaarheidsvraagstuk zal duidelijk zijn. Naarmate er meer garanties kunnen worden gegeven voor de communicatieve, empirische en ideologiekritische validiteit van de resultaten, is de

geldigheidskracht groter en zullen ze een belangrijker rol spelen in toekomstige wetenschappelijke discussies. Dat geldt zeker voor onderzoek naar verschijnselen die zowel micro-, meso-, als macro-sociale aspecten hebben.

#### 4 Besluit

In het voorgaande zijn we precies daar begonnen waar Hetebrij en Wardekker (1986) eindigden. We zijn van mening dat de door hen geëntameerde discussies al volop gevoerd worden (cfr. Miedema, 1986), en hebben een uitwerking gegeven van de kwaliteitscriteria voor kwalitatief onderzoek; een noodzakelijke stap op weg naar de integratie van kwalitatief onderzoek in het sociaal-wetenschappelijk methoden-arsenaal. We hebben enkele begrippen geïntroduceerd die naar onze ervaring de discussie over de kwaliteit van kwalitatief onderzoek vergemakkelijken: technische en argumentatieve betrouwbaarheid, communicatieve, empirische en ideologiekritische validatie; dit alles onder de noemer van intersubjectieve navolgbaarheid. Het lukte ons overigens niet – uiteraard conform onze opvatting – om in het betoog de relatie met het kwantitatieve onderzoek buiten schot te houden. Onzes inziens is zulks op dit moment ook slechts mogelijk bij gratie van onvruchtbaar ideaaltypische en verzuilde beschrijvingen van de traditionele wetenschapsvisies. De wetenschaptheoretische ‘verzuiling’ wordt echter intussen door de onderzoekspraktijk zelf al achterhaald.

#### Literatuur

- Broad, W. & N. Wade, *Betrayers of truth. Fraud and deceit in the halls of science*. New York: 1982.
- Freeman, D., *Margaret Mead and Samoa. The making and unmaking of an anthropological myth*. Harvard: 1983.
- Goetz, J.P. and M.D. Le Compte, *Ethnography and qualitative design in educational research*. Orlando: 1984.
- Hetebrij, M., *De ontwikkeling van researchprogramma's als methodisch probleem*. Amsterdam: 1983.
- Hetebrij, M. en W.L. Wardekker, De kwaliteit van discussies over kwalitatief onderzoek. *Pedagogische Studiën*, 1986, 63, 485-497.

- IJzendoorn, M.H. van en S. van Vliet-Visser, Gezinsuitbreiding. Een kwalitatieve voorstudie naar veranderingen in de relatie tussen eerstgeboren kind en zijn ouders. *Pedagogische Studiën*, 1985, 62, 313-328.
- Janssens, F.J.G., Betrouwbaarheid en validiteit in interpretatief onderzoek. *Pedagogisch Tijdschrift*, 1985, 10, 149-161.
- Komter, A., De onmacht van het getal. Argumenten voor een intensieve methode, toegelicht aan onderzoek naar macht in huwelijksrelaties. *Kennis en Methode*, 1983, 7, 202-226.
- Komter, A., *De macht van de vanzelfsprekendheid in relaties tussen vrouwen en mannen*. Leiden: 1985 (diss.)
- Lamb, M.E., R.A. Thompson, W. Gardner & E.L. Charnov, *Infant-mother attachment*. Hillsdale, N.J.: 1985.
- Mead, M., *Coming of Age in Samoa*. New York: 1961.
- Medawar, P.B., Unnatural Science. *The New York Review of Books*, 1977, 14, no. 1, 13-18.
- Miedema, S., Relevantie en precisie. Over de relatie tussen kwalitatieve en kwantificerende methoden van empirisch-pedagogisch onderzoek. *Pedagogische Verhandelingen*, 1986, 9, 3-30.
- Nardi, B.A., The height of her powers: Margaret Mead's Samoa. *Feminist Studies*, 1984, 10, 323-337.
- Oevermann, U., T. Allert, E. Konan en J. Krambeck, Die Methodologie einer 'objectiven Hermeneutik'. In: P. Zedler en H. Moser (Hrsg.), *Aspekte qualitativer Sozialforschung. Studien zu Aktionsforschung, empirischer Hermeneutik und reflexiver Sozialtechnologie*. Opladen: 1983, 95-123.
- Popper, K.R., *Objective knowledge. An evolutionary approach*. Oxford: 1972.
- Scheper-Hughes, N., The Margaret Mead controversy: culture, biology and anthropological Inquiry. *Human Organization*, 1984, 43, 85-93.
- Spradley, J.P., *The ethnographic interview*. New York: 1979.
- Spradley, J.P., *Participant observation*. New York: 1980.
- Tavecchio, L.W.C., *Quantification of teaching behavior in physical education. A methodological study*. Groningen: 1977.
- Terhart, E., Intuition – Interpretation – Argumentation. Zum Problem der Geltungsbegründung von Interpretationen. *Zeitschrift für Pädagogik*, 1981, 27, 769-793.
- Terhart, E., Interpretative approaches in educational research: a consideration of some theoretical issues – with particular reference to recent developments in West Germany. *Cambridge Journal of Education*, 1982, 12, 141-160.

## *Curricula vitae*

*M.H. van IJzendoorn* (1952) is als hoogleraar theoretische en historische pedagogiek verbonden aan de vakgroep Wijsgerige en Empirische Pedagogiek van de Rijksuniversiteit te Leiden.

*S. Miedema* (1949) studeerde in 1979 af in de wijsgerige en historische pedagogiek en in 1983 in de wetenschapsfilosofie te Groningen; promoveerde in 1986 te Leiden op het proefschrift 'Kennen en handelen. Bijdragen aan het theorie-praktijk-

debat in de opvoedingswetenschap'. Sedert 1981 is hij verbonden aan de vakgroep Wijsgerige en Empirische Pedagogiek van de Rijksuniversiteit te Leiden.

*Adres:* Vakgroep Wijsgerige en Empirische Pedagogiek, Rijksuniversiteit te Leiden, Postbus 9507, 2300 RA Leiden.

*Manuscript aanvaard 5-9-'86*

## **Summary**

IJzendoorn, M.H. van & S. Miedema. 'The quality of qualitative research.' *Pedagogische Studiën*, 1986, 63, 498-505.

In this paper we discuss criteria for measuring the quality of qualitative research. We consider reliability and validity as operationalizations of the fundamental criterion 'intersubjectivity'. During data collection only 'argumentative' reliability can be determined; during data analysis attention should also be given to 'technical' reliability. Validation of qualitative research has to take place in three different ways: 'communicative', 'empirical', and 'critical' validation.