

Deze rubriek is een forum voor debat over kwalitatief methodologische kwesties. Wie een idee heeft voor een thema of reactie wil leveren op een eerder verschenen stuk, vragen wij contact op te nemen met Jeanine Evers: j.c.evers@planet.nl.

Theorie toetsen in kwalitatief onderzoek

Tony Hak

Inleiding

Kwalitatieve onderzoekers betreuren het over het algemeen dat kwalitatief onderzoek doorgaans een exploratieve rol krijgt toebedeeld. In dit essay betoog ik dat zij zelf deels schuldig zijn aan deze roltoedeling, omdat zij geen overtuigende strategie ontwikkeld hebben om theorie te toetsen met kwalitatieve methoden. Ik wil laten zien dat en hoe de gangbare methodologische noties over het ontwikkelen van een generaliseerbare theorie in kwalitatief onderzoek tekortschieten. Ik zal ook proberen een alternatief daarvoor te formuleren.

Er zijn twee strategieën binnen de kwalitatieve onderzoekstraditie waarin gepretendeerd wordt dat daarmee generaliseerbare theoretische stellingen worden ontwikkeld, dat wil zeggen stellingen waarvan impliciet of expliciet wordt geclaimd dat zij ook geldig zijn buiten de gevallen waarin zij zijn 'ontdekt'. Dat zijn analytische inductie en constante vergelijking. Analytische inductie is min of meer in onbruik geraakt en in de kwalitatieve onderzoekspraktijk vervangen door constante vergelijking. Dit heeft te maken met het feit dat de meeste kwalitatieve onderzoekers in hun analyse een versie hanteren van de door Strauss en Corbin (1998) gecodificeerde vorm van gefundeerde theorie (GT). Een goed voorbeeld hiervan is het boek *Analyseren in kwalitatief onderzoek* van Hennie Boeije (2005). Ondanks de generieke titel is dit boek een voor Nederlandse lezers toegankelijk gemaakte versie van de GT-benadering. Analytische inductie wordt in dit boek behandeld als een variant op de methode van constante vergelijking. Beide worden beschreven als methoden waarbij 'verwachtingen of tussentijdse hypothesen over de samenhang tussen verschijnselen worden getoetst op hun juistheid bij de nieuwe gevallen bij wie (of waar) de onderzoeker zijn data verzamelt' (Boeije, 2005: 74). Als die toetsen in steeds weer nieuwe gevallen laten zien dat de hypothesen juist blijken te zijn, dan is er 'verzadiging' bereikt en wordt de hypothese als een 'gefundeerde' theoretische stelling beschouwd. Door deze vervolgens te publiceren wordt impliciet of expliciet geclaimd dat de stelling geldig is buiten de gevallen waarin deze is ontwikkeld en 'getoetst'. Waarom zou men anders onderzoeksbevindingen publiceren in nationale en internationale tijdschriften?

Ik zal in dit essay de volgende stellingen toelichten en onderbouwen:

- Omdat 'verzadiging' een toetsingstrategie is, is het een voorbeeld van de manier waarop kwalitatief onderzoek toetsend kan zijn.

- Als een ‘emerging’ relatie of patroon via een strategie van ‘verzadiging’ kan worden getoetst in individuele (nieuwe) gevallen, dan moeten er typen van relaties zijn die in individuele gevallen kunnen worden getoetst.
- Constante vergelijking en het toetsen van theorie in kwalitatief onderzoek kunnen pas serieus worden genomen als wordt gespecificeerd welke typen relaties in individuele gevallen kunnen worden getoetst.

Hierna zal ik eerst de werkwijze van de GT-benadering bespreken en nagaan waar de schoen wringt en vervolgens laten zien hoe met behulp van vergelijkende gevalsanalyse verschillende typen relaties kunnen worden getoetst.

Dataverzameling, meting en analyse in GT

Kwalitatieve analyse in de door Strauss en Corbin (1998) en Boeije (2005) gecodificeerde vorm bestaat uit drie stappen die ieder met het begrip ‘coderen’ worden aangeduid, te weten ‘open coderen’, ‘axiaal coderen’ en ‘selectief coderen’. Open coderen bestaat uit het markeren van stukjes onderzoeksmateriaal (meestal fragmenten van transcripten van interviews) en het labelen van die fragmenten met een naam (code) waaronder ze worden opgeborgen (Boeije, 2005: 85). Axiaal coderen bestaat uit verschillende procedures (zoals het splitsen en samenvoegen van codes, en zo nodig het benoemen van nieuwe codes) die samen ertoe leiden dat iedere code een precieze definitie krijgt, die het mogelijk maakt om tekstfragmenten eenduidig met een code te verbinden. Eigenlijk is dit een methode van operationalisering waarin codes worden gevalideerd. De toegekende codes vertegenwoordigen nu geldig gemeten waarden op eenduidig gedefinieerde variabelen (dimensies). Een voorbeeld van een dergelijke geldig gemeten waarde is de code ‘normalisering’ die aan een fragment van een interview met een patiënt met een chronische ziekte toegekend wordt. Het resultaat van de fase van axiaal coderen is dat alle fragmenten van alle patiënten met deze code in de verdere analyse kunnen worden behandeld als een meting van de waarde ‘ja’ (of ‘aanwezig’) van de variabele ‘normalisering’. Na het axiaal coderen kan de totale groep van patiënten die zijn onderzocht (geïnterviewd) worden verdeeld in twee groepen, de groep met normalisering en de groep zonder normalisering. Axiaal coderen is dus niet alleen een methode om een begrip te funderen maar ook, gezien het resultaat, een meetmethode. De volgende stap in de GT-benadering is ‘selectief coderen’. Eigenlijk is dit geen vorm van coderen, maar een methode van analyse waarin door middel van de methode van ‘constante vergelijking’ wordt nagegaan of en hoe bepaalde codes (dus de gemeten waarden van variabelen) in de eenheden van onderzoek (bijvoorbeeld patiënten) samenhangen.

De GT-benadering bestaat uit een combinatie van drie stappen die ieder als ‘kwalitatief’ kunnen worden gekarakteriseerd: (a) men begint met tekstuele (‘kwalitatieve’) data; (b) er wordt een ‘kwalitatieve’ methode gebruikt om nominale (‘kwalitatieve’) variabelen te ontwikkelen waarop de onderzoekseenheden worden gemeten; en (c) de samenhangen tussen de zo gemeten variabelen worden op een niet-statistische (‘kwalitatieve’) manier geanalyseerd. Een verwarring die ik in de onderzoekspraktijk waarneem, is dat het vaak

als vanzelfsprekend aangenomen wordt dat gegevens die verzameld zijn met behulp van 'kwalitatieve' waarnemingsmethoden (zoals bijvoorbeeld participierend observeren of half-gestructureerd interviewen) alleen 'kwalitatief' (dus nominaal) gecodeerd zouden kunnen worden en dat zij ook alleen op een kwalitatieve manier (zoals bijvoorbeeld door middel van 'constante vergelijking') geanalyseerd kunnen worden. Het gaat hier echter om drie fasen in het onderzoeksproces die (relatief) onafhankelijk zijn en die ieder al of niet (of minder of meer) 'kwalitatief' kunnen zijn. Data kunnen immers in principe altijd met behulp van coderingsrichtlijnen in numerieke codes worden omgezet onafhankelijk van de manier waarop zij zijn verzameld. En die numerieke codes kunnen vervolgens statistisch worden geanalyseerd. Dat geldt ook voor waarden die via axiaal coderen verkregen zijn. Omgekeerd is het ook zeker mogelijk om numerieke scores op te nemen in procedures van 'constante vergelijking'.

In een discussie over de vraag of theorie getoetst kan worden in kwalitatief onderzoek is het essentieel om onderscheid te maken tussen meten en analyseren, en zich te realiseren dat de discussie over het toetsen van theorie enkel betrekking heeft op de methode van analyse (dus op de vraag naar relaties tussen gemeten waarden). In de rest van dit essay gaat het alleen over kwalitatieve toetsing, niet over kwalitatieve data of codes.

Het onderscheid tussen verschillende typen theoretische uitspraken

In de literatuur over kwalitatieve analyse wordt zelden een onderscheid gemaakt tussen de verschillende typen relaties die men als resultaat van de analyse zou kunnen vaststellen. Een voorbeeld van een resultaat van een analyse in een kwalitatief onderzoek is een bevinding dat er verschillende typen of patronen van een verschijnsel bestaan en dat deze samenhangen met een situatie of een aanleiding. Een dergelijk resultaat is bijvoorbeeld de stelling dat er drie typen van omgaan met therapeutische voorschriften bestaan en dat het afhangt van de ernst van de symptomen of van de houding van een mantelzorger welk type feitelijk optreedt. Er wordt dus een oorzakelijke relatie gevonden tussen een conditie (een fysieke toestand of een eigenschap van een verzorger) en een gedrag van een patiënt. Kenmerkend voor zo'n onderzoeksresultaat is dat niet wordt gespecificeerd wat er precies wordt bedoeld. Wordt daarmee bedoeld dat de geanalyseerde gegevens erop wijzen dat er bij een gegeven conditie een grotere kans is op een bepaald gedrag (een kansrelatie)? Of wordt er bedoeld dat er bij aanwezigheid van zo'n conditie onvermijdelijk en dus altijd een bepaald patiëntengedrag volgt ('voldoende voorwaarde')? Of wordt bedoeld dat een vorm van (wenselijk geacht) patiëntengedrag alleen mogelijk is onder bepaalde voorwaarden, bijvoorbeeld als ook de zorgverlener zich op een bepaalde manier opstelt ('noodzakelijke voorwaarde')? Dit zijn niet alleen naar de inhoud drie verschillende typen uitspraken, maar ze verschillen ook naar de wijze waarop ze kunnen worden vastgesteld.

Het is in dit verband essentieel om een onderscheid te maken tussen 'probabilistische' en 'deterministische' uitspraken. Een probabilistische uitspraak formuleert een kans: Als A, dan is er meer kans op B (dan er is zonder A). Een deterministische uitspraak formuleert een zekerheid: Als A, dan zeker B ('voldoende voorwaarde'). Of ook: Als niet-A, dan

ook zeker niet-B. Deze laatste formulering die een voldoende voorwaarde uitdrukt voor de afwezigheid van B, kan ook als een noodzakelijke voorwaarde voor de aanwezigheid van B worden geformuleerd: Alleen B als er ook A is. Kern van mijn betoog is dat het onderscheid tussen probabilistische en deterministische relaties essentieel is zowel voor het ‘vinden’ als voor het toetsen van theorie. Zij moeten op verschillende manieren worden ‘gevonden’ en ook op verschillende manieren worden getoetst.

Ik zal hier eerst bespreken hoe men probabilistische en deterministische verbanden kan ‘vinden’. Daarna bespreek ik het toetsen van deze twee typen verbanden.

Hoe kunnen relaties tussen concepten worden gevonden?

Stel, we hebben gegevens van tien patiënten en we hebben voor ieder van hen (al of niet op een ‘kwalitatieve’ manier gegenereerde) waarden voor (a) ernst van symptomen; (b) de houding van de mantelzorg; en (c) het gedrag van de patiënt. Die waarden plaatsen we in de volgende datamatrix.

	<i>Ernst van de symptomen op schaal 1-10</i>	<i>Houding van mantelzorg</i>	<i>Gedrag van de patiënt</i>
Patiënt 1	10	Ondersteunend	Therapietrouw
Patiënt 2	7	Ondersteunend	Therapietrouw
Patiënt 3	7	Ondersteunend	Therapietrouw
Patiënt 4	6	Ondersteunend	Neutraal
Patiënt 5	10	Neutraal	Neutraal
Patiënt 6	6	Neutraal	Neutraal
Patiënt 7	4	Neutraal	Neutraal
Patiënt 8	6	Neutraal	Totaal geen terapietrouw
Patiënt 9	3	Tegenwerkend	Totaal geen terapietrouw
Patiënt 10	3	Tegenwerkend	Totaal geen terapietrouw

De patiënten zijn in deze datamatrix gerangschikt naar de mate van terapietrouw (van hoog naar laag). Therapietrouw wordt beschouwd als wenselijk. Welke aanwijzingen zijn er in deze datamatrix voor condities die terapietrouw bevorderen of verklaren?

In de tabel zijn de patiënten geordend naar de waarde van het te verklaren gedrag. Meteen ziet men in de kolom van het gedrag van de mantelzorg ook een rangordening van het meer positief geachte gedrag (ondersteunend) naar het als negatief beoordeelde gedrag (tegenwerkend). In de kolom symptoomernst is ook een gradiënt waarneembaar, van een hoge naar een lage symptoomernst, maar de rangorde is niet perfect. We kunnen op basis van deze waarnemingen relaties ‘vinden’ als:

- Bij meer symptoomernst is er meer terapietrouw (of omgekeerd).
- Bij meer ondersteuning van de mantelzorg is er meer terapietrouw (of omgekeerd).

Dit zijn twee probabilistische stellingen waarin het bestaan van een verschil in kans op een bepaald effect wordt uitgedrukt. Dat verschil is gevonden door de trends in de drie kolommen te vergelijken, niet door tussen individuele gevallen te vergelijken. (Een vergelijking tussen patiënten 3 en 5 zou het gevonden probabilistische verband niet zichtbaar maken.) Met andere woorden, het verschil in kansen kan niet waargenomen worden door middel van een 'constante vergelijking'. Er is maar één enkele vergelijking, namelijk tussen de trend in de ene kolom en die in andere kolom. Die gaan al of niet (of in grotere of geringere mate) in dezelfde richting, en verder komt er geen 'vergelijking' aan te pas bij het ontwikkelen van een probabilistische stelling uit een verzameling data.

Dezelfde datamatrix kan ook worden gebruikt voor het formuleren van een ander type stellingen. We zouden bijvoorbeeld op zoek kunnen zijn naar 'kritieke' factoren voor therapietrouw. Om zulke factoren te vinden moeten we groepjes van gevallen maken en daarbinnen kijken of aan een bepaalde voorwaarde wordt voldaan. Een voorbeeld: als we willen vaststellen of de datamatrix aanwijzingen bevat voor een 'kritieke' voorwaarde voor therapietrouw, moeten we een groepje maken van alle patiënten met therapietrouw en dan nagaan of er een factor is die kennelijk altijd (niet meestal) samengaat met dit gedrag. Op deze manier kunnen we tot de volgende stellingen komen:

- Een relatief hoge ernst van symptomen (6 of meer) is een 'kritieke' factor voor therapietrouw.
- Een ondersteunende zorgverlener is een 'kritieke' factor voor therapietrouw.

Men zou ook kunnen zeggen dat beide factoren (hoge symptoomernst en ondersteunende zorgverlener) allebei 'noodzakelijke' voorwaarden lijken te zijn, dus condities waaraan voldaan moet zijn om tot therapietrouw te komen. In deze stellingen is de richting van het verband voorondersteld. Maar het verband zou ook omgekeerd kunnen zijn, zoals in de stellingen:

- Therapietrouw is een 'voldoende' voorwaarde voor ondersteunend gedrag van de zorgverlener.
- Het totaal ontbreken van therapietrouw is een 'noodzakelijke' voorwaarde voor tegenwerkend gedrag van de zorgverlener.

Deze stellingen worden 'gevonden' door steeds een groepje van patiënten te nemen met een zelfde waarde van een variabele (een groepje met therapietrouw, of juist zonder, of een groepje met ondersteunende zorgverlener of juist een tegenwerkende zorgverlener) en dan binnen dat groepje naar een patroon van waarden van de andere variabelen te zoeken. De stelling wordt gevonden door een patroon vast te stellen binnen het groepje als geheel, niet door 'constante vergelijking' tussen afzonderlijke gevallen.

Zowel 'probabilistische' stellingen (die verschillen in kans op een effect uitdrukken) als 'deterministische' stellingen (die uitdrukking geven aan de veronderstelling dat een voorwaarde 'kritiek', 'noodzakelijk' of 'voldoende' is) worden dus *niet* gevonden door een 'constante vergelijking' tussen gevallen maar door een trend of een patroon in een groep

van gevallen waar te nemen. Het is moeilijk te verklaren dat in de literatuur over kwalitatieve analyse alsmat beweerd wordt dat er in kwalitatief onderzoek maar één type analyse bestaat, namelijk 'constante vergelijking'. Het is aannemelijk dat dit vooral te maken heeft met de verwarring tussen meten en analyseren. Het is immers wel goed voorstelbaar dat men aan een patiënt pas een waarde van een bepaalde variabele (zoals therapietrouw) toekent na vergelijking van alle relevante tekstfragmenten van alle patiënten. Een dergelijke vergelijking zou men 'constante vergelijking' kunnen noemen. Meten (het bepalen van de waarde van een variabele in een geval) is echter heel iets anders dan analyseren (het vaststellen van een relatie tussen waarden van variabelen).

Nu is vastgesteld dat verschillende typen relaties op verschillende manieren 'gevonden' moeten worden in een datamatrix (en dat 'constante vergelijking' niet tot die manieren behoort), kunnen we nagaan op welke manier die verschillende typen relaties getoetst kunnen of moeten worden. Moeten verschillende typen relaties ook op een verschillende manier getoetst worden? En, kan 'constante vergelijking' hierbij een rol spelen?

Verzadiging en generalisatie

Stel, dat de tien patiënten van wie de gegevens zijn opgenomen in de datamatrix deel uitmaken van een groep van dertig patiënten die op een 'theoretische', 'doelgerichte' of 'purposive' (dus in ieder geval niet-probabilistische) wijze waren geselecteerd uit het ledenbestand van een landelijke vereniging van patiënten met een bepaalde ziekte. In een dergelijk geval is doorgaans de claim van het onderzoek dat de resultaten in die dertig gevallen wel zullen gelden voor een grotere groep patiënten. Dat is toch ten minste een impliciete claim als het onderzoek in het ledenblad wordt gepubliceerd en er praktische implicaties (voor de leden) worden besproken. Dat is ook een impliciete claim als het onderzoek wordt gepubliceerd in *Social Science & Medicine* als een bijdrage aan onze kennis over de determinanten van therapietrouw. Hoe kunnen zulke claims worden onderbouwd?

De standaardmanier om een dergelijke claim te rechtvaardigen is die van 'verzadiging'. Er wordt gesteld dat er nieuwe patiënten voor het onderzoek waren geselecteerd nadat de relatie tussen de variabelen in een eerste groep patiënten was gevonden, en dat in die nieuwe groep patiënten geen bevindingen zijn gedaan die strijdig waren met de eerder gevonden conclusies. Bij deze rechtvaardiging kunnen de volgende twee kanttekeningen worden gemaakt:

- 'Kijken of een bevinding ook geldt in nieuwe gevallen' is een vorm van toetsen, die niet kan worden onderscheiden van toetsen in theorietoetsend onderzoek. Er is een (voorlopige) hypothese geformuleerd en die wordt getoetst, want anders zou niet de conclusie kunnen worden getrokken dat 'verzadiging' is bereikt. Impliciet wordt dus geclaimd dat men weet hoe een probabilistische of een deterministische stelling kan worden getoetst in een klein aantal gevallen.

- Als de nieuwe gevallen opnieuw uit het ledenbestand van de vereniging worden geselecteerd – wat de meest voorkomende strategie is – dan kan alleen maar naar de populatie van de leden van de vereniging worden ‘gegeneraliseerd’.

De laatstgenoemde beperking – dat we na verzadiging iets weten over de leden van de vereniging, maar niets over patiënten daarbuiten – wordt in het meeste kwalitatieve onderzoek genegeerd met verwijzing naar het begrip ‘analytische generalisatie’ (vaak met Yin als bronvermelding). Echter, ‘analytische generalisatie’ is bij Yin geen claim maar een ‘aim’. Yin geeft geen vrijbrief om (analytische) generaliseerbaarheid te claimen, maar legt uit hoe je deze generalisatie kunt bereiken, namelijk door replicatie. Replicatie vereist dat nieuwe gevallen moeten worden geselecteerd uit andere hoeken van het theoretische domein – de verzameling van alle instanties van het theoretische object – en juist niet uit dezelfde pot als waaruit eerdere gevallen waren geselecteerd. Het domein in het voorbeeld bestaat uit de verzameling van alle chronische patiënten – waar ook ter wereld – die te maken hebben met mantelzorgers en therapietrouw. Aannemelijk maken dat een bevinding (die altijd lokaal is en een beperkte reikwijdte heeft) geldt voor het hele domein doet men door te repliceren in zeer diverse gevallen uit dat domein. Verzaadiging zoals doorgaans gepraktiseerd is een methode om naar een specifieke populatie (hier: het ledenbestand van de vereniging) te generaliseren, niet naar het theoretische domein.

Theorietoetsend kwalitatief onderzoek

Zowel bij ‘verzaadiging’ als bij ‘analytische generalisatie’ wordt een voorlopige stelling getoetst op nieuwe gevallen. Hoe kan een dergelijke toetsing plaatsvinden? Veronderstel dat de analyse van de eerste tien gevallen van leden van de vereniging (zie de datamatrix) de volgende stelling heeft opgeleverd: *‘Symptoomernst 6 of meer is een noodzakelijke voorwaarde voor therapietrouw.’* Om deze stelling te toetsen moeten we om te beginnen (andere) patiënten vinden met therapietrouw. Patiënten zonder therapietrouw zijn niet relevant voor de toets. Als we een patiënt met therapietrouw hebben gevonden, dan kijken we naar de ernst van de symptomen. Een therapietrouwe patiënt met een lagere symptoomernst is een aanwijzing dat de stelling niet klopt. Iedere therapietrouwe patiënt met een hoge symptoomernst is een aanwijzing dat de stelling klopt. Als we er maar niet in slagen, zelfs na lang zoeken, om een therapietrouwe patiënt te vinden met een lage symptoomernst (‘verzaadiging’), dan kunnen we in de stelling gaan geloven. Dat wil zeggen: dan is de stelling waarschijnlijk waar voor de leden van de vereniging. Als we willen generaliseren naar vergelijkbare patiënten buiten de vereniging, dan zullen we dezelfde toets vele keren moeten herhalen bij patiënten uit allerlei hoeken van het domein, dus met andere ziektes, uit andere landen, met of zonder lidmaatschappen, enzovoort. Die herhaalde toetsing wordt replicatie genoemd, en *‘verzaadiging’ is eigenlijk een vorm van replicatie die zich om praktische redenen beperkt tot maar één populatie uit het domein.* Het voorbeeld laat zien dat toetsing van een ‘noodzakelijke’ voorwaarde (herhaaldelijk) plaatsvindt op een enkel geval (single case). In dat ene geval wordt een ‘vergelijking’ gemaakt tussen de verwachte waarde van de variabele (hier: symptoomernst 6 of meer) en de geobserveerde waarde (‘pattern matching’). Buiten deze eenmalige matching van een

verwachte en een geobserveerde waarde van de relevante variabele vindt er in een toets geen andere vergelijking plaats. Er is dus geen sprake van ‘constante’ vergelijking.

Veronderstel nu dat uit de eerste tien gevallen een probabilistische stelling naar boven is gekomen, zoals: ‘*Bij patiënten met een grotere symptoomernst is er gemiddeld genomen meer therapietrouw*’. Om deze stelling te toetsen moet eerst worden vastgesteld op welke populatie van patiënten deze stelling betrekking heeft. Deze stelling kan immers niet slaan op een groepje gevallen dat op een ‘purposive’ manier bij elkaar is gezocht, omdat dan het gevaar van selectie bias te groot is. Om deze stelling te toetsen moet een *random* (of in ieder geval probabilistische) steekproef worden getrokken uit een afgebakende populatie. Het is daarom op voorhand duidelijk dat een dergelijke ‘emerging’ stelling niet hard(er) kan worden gemaakt in een strategie van ‘verzadiging’. En van ‘constante vergelijking’ kan hier al helemaal geen sprake zijn.

Omdat in GT geen systematisch onderscheid wordt gemaakt tussen deterministische en probabilistische stellingen die het resultaat zouden zijn van ‘selectief coderen’ worden in de praktijk vaak probabilistische stellingen geformuleerd (zoals: Patiënten met A hebben meer B) die helemaal niet kunnen volgen uit de voorgeschreven technieken (‘constante vergelijking’ en ‘verzadiging’) omdat ze (a) alleen kunnen worden ontdekt door de observatie van een trend in een groepje gevallen (zoals in de datamatrix) en (b) alleen gegeneraliseerd kunnen worden naar een populatie als dat groepje op een probabilistische manier is geselecteerd.

Het is echter wel mogelijk gebleken om deterministische stellingen (zoals een stelling over een ‘noodzakelijke’ voorwaarde) te ontdekken in een datamatrix (dus bijvoorbeeld in een groepje van tien patiënten die op een *purposive* manier uit een ledenbestand zijn gehaald) en vervolgens die stelling te toetsen op individuele gevallen (single cases) in een strategie van replicatie. Publicatie in een internationaal tijdschrift van een deterministische stelling die is gevonden in kwalitatief onderzoek (met een impliciete of een expliciete claim dat die stelling generaliseerbaar is naar een theoretisch domein) is gerechtvaardigd als de replicaties zich hebben uitgestrekt tot gevallen uit verschillende hoeken van het theoretische domein (en dus niet alleen tot gevallen uit de lokale setting waarin men de stelling heeft ontdekt).

Het kwalitatieve onderzoek zou zich een veel sterkere positie verwerven als (a) daarin de aard van de gevonden samenhangen veel preciezer werd omschreven (namelijk met een specificatie van het deterministische dan wel probabilistische karakter daarvan) en (b) zichtbaar werd gemaakt hoe de gehanteerde analytische technieken (zowel technieken van het ‘ontdekken’ van verbanden als technieken van ‘toetsing’ daarvan) passen bij de aard van de geformuleerde samenhangen. De gangbare manier om generaliseerbaarheid van bevindingen van kwalitatief onderzoek te rechtvaardigen – met gebruik van termen als ‘analytische generaliseerbaarheid’, ‘theoretische selectie’ en ‘verzadiging’ – is in dit perspectief nogal gebrekkig.

De pretentie van het kwalitatieve onderzoek kan beter als volgt worden geherformuleerd:

In kwalitatief onderzoek kunnen verbanden worden ontdekt door vergelijkende gevalsanalyse ('cross case analysis'). Deterministische stellingen kunnen worden gegeneraliseerd naar een theoretisch domein door ze te toetsen in individuele gevallen ('single cases') die in een strategie van replicatie uit dat domein zijn geselecteerd.

Literatuur

Boeije, H. (2005). *Analyseren in kwalitatief onderzoek*. Amsterdam: Boom.

Strauss, A.L. & Corbin, J. (1998). *Basics of qualitative research: techniques and procedures for developing grounded theory*. Thousand Oaks: Sage.

Yin, R.K. (2003). *Case study research. Design and methods*. (third edition). Beverly Hills: Sage.

Wat kun je met de gefundeerde theoriebenadering?

Commentaar op Tony Hak

Hennie Boeije

Het is goed dat onderzoekers nadenken en discussiëren over hun vakgebied. Ik heb Tony's bijdrage opgevat als een poging om stil te staan bij enkele lastige termen en moeilijkheden met het doel te komen tot opheldering. Met zijn essay wil hij denk ik laten zien dat sommige kwalitatieve benaderingen, en dan met name de gefundeerde theorie, niet geschikt zijn om theorie te toetsen en dat andere typen (kwalitatief) onderzoek daar wel geschikt voor zijn. Hij beschrijft wat hij onder toetsing verstaat en hoe daarbij te werk dient te worden gegaan.

Ik denk dat de redenering klopt voor wat in de wetenschap onder toetsen wordt verstaan en uitgaande van de door Tony gehanteerde uitgangspunten. Het voorbeeld gaat in op de omstandigheden waaronder patiënten met een zekere aandoening therapietrouw zijn. Wanneer toetsing van het verband tussen de ernst van de ziekte, de aard van de mantelzorg en therapietrouw het doel van het onderzoek is, dan ben ik het eens met zijn redenering. Een experimenteel design, een enquête of een case-study kunnen hier goede diensten bewijzen.

Tony's verzet lijkt zich te concentreren op de gefundeerde theoriebenadering en de onmogelijkheid daarmee te toetsen. Laat ik voorop stellen dat je met de gefundeerde theoriebenadering helemaal niet hoeft te toetsen. Met deze benadering kan je hele mooie andere doeleinden realiseren. Ik denk dat zeer weinigen de gefundeerde theoriebenadering associëren met toetsing zoals Tony zich dat voorstelt en dat weinigen op deze manier te werk zouden gaan. Met name de procedure waarin kwalitatieve gegevens worden omgezet in