

Misverstanden over functionele analfabeten.

Een reactie op J.H.L. Oud:
'Analfabetisme in Nederland?'

In een bijdrage in het februarinummer van *Pedagogische Studiën* gaat Oud (1986) in op de Voorstudie Periodieke Peiling van het Onderwijsniveau. Deze Voorstudie is van 1983 tot 1985 onder leiding van dr. H. Wesdorp verricht door de Stichting Centrum voor Onderwijsonderzoek van de Universiteit van Amsterdam (S.C.O.). Oud zet terecht enkele kritische kanttekeningen bij bepaalde aspecten van de Voorstudie. Zijn commentaar heeft betrekking op veel verschillende aspecten van de Voorstudie. In het volgende zullen wij enkele zaken die Oud aanvoert verduidelijken.

1 De voorstudie PPO

Het doel van de Voorstudie was tweeledig. Ten eerste moest vastgesteld worden of een periodieke peiling van het onderwijsniveau gerealiseerd kon worden. Deze vraag had betrekking op een aantal 'technische' aspecten als instrumentatie, steekproeftrekking, dataverzameling, scoring en beoordeling, analyse en rapportage. Ten tweede diende nagegaan te worden in hoeverre een peiling geaccepteerd zou worden door de scholen. Wat is de houding van leerkrachten, ouders en onderwijsverzorgingsinstellingen ten aanzien van een peilingsonderzoek? In overleg met de opdrachtgever, het ministerie van Onderwijs en Wetenschappen, is besloten een proefpeiling te verrichten voor het vak taal in het laatste leerjaar van het basisonderwijs. Bewust is voor het vak taal gekozen, omdat het een aantal lastig operationaliseerbare doelstellingen kent. Zou een proefpeiling voor taal slagen, dan zal ook een peiling voor bijvoor-

beeld rekenen of wereldoriëntatie realiseerbaar zijn.

In de Voorstudie is gebruik gemaakt van een aantal functionele taalopdrachten. Dit zijn taken, ontleend aan taalgebruikssituaties, waarmee in principe elke 11 à 12-jarige leerling in aanraking kan komen. Sommige taken lijken erg op situaties die zich buiten school voordoen. Andere taken zijn ontleend aan situaties die typerend voor de school zijn. De leerlingen hebben dus opdrachten uitgevoerd als: het opzoeken van programma's in een radio- en t.v.gids, het mondeling uitleggen van een spel aan een klasgenoot, het schrijven van een bestelbriefje en het schrijven van een fantasie-opstel.

Het projectteam heeft gekozen voor het meten van functionele taalvaardigheid, omdat de school leerlingen taal voor het leven leert. De school bereidt de leerlingen voor op zulke situaties, allerlei schoolse taalvaardigheden als het vormen van verkleinwoorden, het toepassen van spreekwoorden etc. zijn als intermediaire doelstellingen te beschouwen. Ten tweede heeft de keuze voor realistische taken tot voordeel dat ook niet direct bij het onderwijs betrokkenen zich een oordeel over de wenselijkheid van het beheersen van zulke vaardigheden kunnen vormen. Bovendien wordt met de keuze voor functionele taalvaardigheid aangesloten bij de keuze zoals die o.a. door buitenlandse peilingsinstituten is gemaakt (zie o.a. *Assessment of Performance Unit* van de National Foundation for Educational Research in Engeland, of de *National Assessment of Educational Progress* in de V.S.).

Door de keuze voor functionele taalvaardigheden ontstond de noodzaak tot het houden van een doelstellingenenquête. Immers, welke functionele doelstellingen worden algemeen wenselijk geacht voor zesdeklassers? Hiertoe is een flink aantal leerkrachten in het basisonderwijs, ouders en moedertaalspecialisten aangeschreven. Allen gaven een oordeel over de wenselijkheid van de in de enquête genoemde doelstellingen. Echter, het zich kunnen redden in een bepaalde doelstelling kan wel wenselijk geacht worden, doch (nog) ver buiten het vermogen van veel zesdeklassers liggen. Vandaar dat in de enquête aan de leerkrachten een vraag naar de 'haalbaarheid' van de situaties was toegevoegd: kan naar uw mening het merendeel der zes-

deklassers zich in deze situatie redden? Deze vraag naar de haalbaarheid van de doelstellingen is dus een poging tot het vermijden van de operationalisering van doelstellingen die (nog) buiten het bereik van de meerderheid der zesdeklassers ligt. Er is dus géén sprake van een a priori normering, zoals Oud abusievelijk meldt.

In de Voorstudie zijn ook moeilijk meetbare vaardigheden als de stel- en spreekvaardigheid gemeten. Dit is juist één van de belangwekkende aspecten van de Voorstudie. Het is al jaren bekend dat meerkeuze- en andere geprecodeerde toetsen op grote schaal afneembaar zijn. Hiermee bezit het CITO een grote ervaring. Echter, is het ook mogelijk op grote schaal schrijf- en spreektaken af te nemen en betrouwbaar te beoordelen? In de Voorstudie is aangetoond dat dit kan. Door nieuwe beoordelingsmethoden te hanteren konden in korte tijd veel stel- en spreekprodukten, door jury's van twee of drie beoordelaars, betrouwbaar beoordeeld worden. Afhankelijk van de taak varieerde de betrouwbaarheid van de beoordeling van .71 tot .98 bij spreken, en van .79 tot .99 bij het stellen. Deze coëfficiënten zijn het resultaat van uitgebreide pretesting van de beoordelingsprocedures. Met name bij het spreken, waar veel minder ervaring is opgedaan met de beoordelingsproblematiek, bleek de interbeoordelaarsovereenstemming in het begin lager. Abusievelijk meldt Oud in zijn commentaar de coëfficiënten uit deze pretestfase.

2 Functionele analfabeten

Een belangrijk onderdeel van het commentaar van Oud betreft de zogenaamde functionele analfabeten. Oud heeft er bezwaar tegen dat wij de resultaten van de Voorstudie niet alleen in een wetenschappelijk rapport vastgelegd hebben, maar ook vertaald hebben voor een breder publiek, van ouders, leerkrachten, schoolbesturen en andere overheden. Wij achten dit wel belangrijk, daar wetenschappelijke resultaten zonder een dergelijke vertaling geen maatschappelijke implicaties kunnen hebben. Bij deze vertaling hebben wij onder meer de nadruk gelegd op de bevinding dat een groep zesdeklassers zoveel moeite met de taken heeft, dat men zich ern-

stig af moet vragen hoe zij zich in alledaagse situaties kunnen redden. Deze groep leerlingen heeft in de pers gefigureerd als functionele analfabeten. De omvang van deze groep is bepaald door na te gaan hoeveel leerlingen op maximaal twee van de gemaakte taken (5 of 6) boven een bepaalde grensscore presteerden. Bij de keuze van deze grensscores zijn wij zeer voorzichtig te werk gegaan, waardoor de schatting van 7% functioneel analfabeten voor lezen (en 9% voor schrijven) conservatief is, d.w.z. aan de lage kant. Oud merkt terecht op dat een dergelijke dichotomie wetenschappelijk gezien niet interessant is. Peilingonderzoek heeft echter naast het wetenschappelijk ook een maatschappelijk belang. Het is maatschappelijk van groot belang om te weten dat er een groep 11 à 12-jarigen is die over een zeer slechte lees- of schrijfvaardigheid beschikt. Er is een groep leerlingen in het basisonderwijs die extra aandacht nodig heeft. Immers, deze leerlingen zijn zowel in vele maatschappelijke (taal-)situaties, als bij hun verdere schoolcarrière, onthand door hun gebrekkige lees- of schrijfvaardigheid.

Oud verwijt ons dat wij de criteria voor de bepaling van de omvang van de groep functionele analfabeten onvoldoende beargumenteerd hebben. Hij lijkt ons zelfs te verwijten dat er überhaupt subjectieve elementen voorkomen in onze procedure. Kennelijk acht hij alleen een volstrekt objectieve procedure waarin geen enkele overweging zichtbaar is acceptabel. Zulke procedures bestaan in de onderwijsresearch niet. Argumenten zullen altijd noodzakelijk blijven. Onze argumenten staan open voor discussie. We willen ze ook inruilen voor betere argumenten (met de bijbehorende beslissingsregels). Hoe zijn wij te werk gegaan? De minimale lezers – respectievelijk schrijvers – zijn achterhaald door per gemaakte lees- en schrijftaak grensscores vast te stellen. Natuurlijk is een leerling die op één van zijn 5 of 6 leestaken minimaal presteert géén minimale lezer. Hij kwam pas voor het etiket minimale lezer of schrijver in aanmerking, wanneer hij op meerdere taken minimaal presteerde. De keuzen voor de grensscores zijn door het projectteam gemaakt, per taak zijn grensscores bepaald die betrekking hadden op de communicatieve functie van de taak: kan de communicatie slagen op grond van de door de leerling ver-

strekke informatie? Zo moest in een bestelbriefje aan Greenpeace minimaal één artikel besteld worden, de datum waarvoor het artikel ontvangen moest worden, en een adres waarheen het bestelde artikel gestuurd moest worden. Echter, niet bij alle taken bleek een dergelijke inhoudelijke motivatie mogelijk. Hoe bijvoorbeeld te bepalen of een leerling een leestekst al dan niet begrepen heeft? Voor deze taken zijn zeer voorzichtige grensscores gekozen, waarbij zowel met de aard van de taak, als het type vragen rekening gehouden is. Zo is voor een taak met tweekeuze vragen een hogere grensscore vastgesteld, dan bij een taak met open vragen. Ondanks deze voorzichtigheid waarmee de grensscores gekozen zijn, is een normeringsonderzoek géén overbodige luxe. Het lijkt noodzakelijk om de taken aan een panel van leerkrachten, ouders en vakspecialisten voor te leggen, met de vraag wat zij een minimaal voldoende prestatie vinden. Kortom, er is onderzoek nodig naar de minimeindtermen van het basisonderwijs.

3 Ontbrekende waarnemingen

Ten tweede roert Oud het probleem van de ontbrekende waarnemingen aan. Het bleek namelijk dat bij de vier opdrachtboekjes een flink aantal leerlingen één of meer taken overgeslagen had. Bij taakboekje A bleek dat in totaal 12% der leerlingen één of meer taken overgeslagen had, bij de boekjes B en C was dit 4% en bij boekje D, dat enkel aan allochtone leerlingen afgenomen was, bleek dat per taak gemiddeld 7,5% der leerlingen deze taak overgeslagen had. Het percentage allochtone leerlingen dat een taak overgeslagen had, achten wij zo hoog dat het niet zinvol leek om ook voor deze groep leerlingen minimale lezers en schrijvers te onderscheiden. Bovendien bleek het overslaan van een taak sterk gecorreleerd met de gemiddelde moeilijkheid van die taak ($r_s = .93$). Voor de taakboekjes A, B en C zijn wel minimale lezers en schrijvers berekend. Het percentage leerlingen dat een taak overgeslagen had, was doorgaans per taak laag (gemiddeld 1%). Het is dus zeer onwaarschijnlijk dat veel leerlingen als minimale lezer of schrijver geclassificeerd zijn, enkel en alleen omdat zij veel taken hebben overgeslagen. Zeker wan-

neer ook in aanmerking genomen wordt dat, ondanks de verschillen in de percentages leerlingen die één of meerdere taken in een toetsboekje overgeslagen hebben, en toch bij alle drie de toetsboekjes ongeveer hetzelfde percentage leerlingen als minimale lezer dan wel minimale schrijver geclassificeerd is.

4 Acceptabiliteit

Een ander punt in het commentaar van Oud is de medewerking van de scholen, die wij groot genoemd hebben (52% van de aangeschreven scholen bleek bereid tot medewerking aan de Voorstudie). Dit gaat zeker op wanneer ook de tijd die aan de Voorstudie besteed diende te worden in ogenschouw genomen wordt. De klasseleerkracht moest behalve de taalbeheersingsinstrumenten, ook een vragenlijst aan de leerlingen afnemen, en zelf nog een (lange) vragenlijst invullen. Ook het hoofd der school diende een (lange) vragenlijst in te vullen. Daarnaast werden nog 50 van de 159 aan het onderzoek deelnemende scholen door een lid van het projectteam bezocht, om een spreek- en luistertoets af te nemen. In veel onderwijskundig onderzoek bleek de bereidwilligheid van de scholen tot medewerking gering.

Ook uit een gecombineerd SVO-Intomart onderzoek zijn gegevens over de bereidheid tot het verlenen van medewerking aan peilingsonderzoek bekend (Scheerens, 1985). In dit onderzoek werden de meningen van leerkrachten en ouders geïnventariseerd. Deze inventarisatie vond zowel plaats, voordat de Voorstudie haar beslag kreeg, als erna. Voor dat de proefpeiling uitgevoerd was, bleek 64% der geënquêteerde schoolhoofden positief ten opzichte van peilingsonderzoek te staan, en 14% bleek een positieve houding te hebben mits aan enkele voorwaarden werd voldaan. In 1985 zijn deze percentages gestegen tot respectievelijk 82% en 15%.

In dit SVO-Intomart onderzoek is ook aan de hoofden van de scholen gevraagd naar de bereidheid van hun school om medewerking te verlenen aan een peilingsonderzoek. In 1984 bleek 64% hiertoe bereid en nog eens 21% bleek bereid mits aan enkele voorwaarden werd voldaan. Voor 1985, nadat de proefpeiling had plaatsgevonden, bleek 76% bereid en 15% onder bepaalde voorwaarde

bereid. Kortrom, er kan van een grote bereidheid van de scholen gesproken worden.

een methode, die door het CITO gehanteerd wordt, met anker items/taken de voorkeur.

H. van den Bergh (S.C.O.)

5 Onopgeloste problemen

Natuurlijk zijn er in de Voorstudie veel onopgeloste problemen. Met name de cross-sectionele (en longitudinale) vergelijking die, wanneer er meerdere peilingen achter de rug zijn, nog de nodige hoofdbrekens zal kosten. Hoe een dergelijke vergelijking precies gemaakt kan worden is vooralsnog onduidelijk. Het kan zijn dat een homogene Rasch-itempool een goede mogelijkheid is. Mits vaardigheden unidimensioneel zijn. Dit is echter nog de vraag. Waarschijnlijk verdient

Literatuur

Oud, J. H. L., Analfabetisme in Nederland? Commentaar op de Voorstudie Periodieke Peiling van het Onderwijsniveau (PPON). *Pedagogische Studiën*, 1986, 63, 85-91.

Scheerens, J., *Opinies over een periodieke peiling van het onderwijsbeleid*. SVO/Intomart, Amsterdam: 1985.

Manuscript aanvaard 30-1-'86