

Don Mellenbergh - De laatste der generalisten

□ Jackelien ter Burg en Henk van Berkel

We spraken met een ontspannen Don Mellenbergh in de stationsrestauratie in Utrecht in de herfst van het vorig jaar. Hij is inmiddels 72 jaar, maar nog steeds actief. Hij legt de laatste hand aan het schrijven van een inleidend boek voor onderzoeksmasters getiteld *A Conceptual Introduction to Psychometrics* en hij geeft cursussen aan onderzoeksmasters en promovendi in de Methodenleer. Daarnaast is hij nog medebegeleider van promovendi, waarbij hij zich vooral richt op de methodologische kant van het promotieonderzoek.

Hoe en wanneer bent u geraakt door het onderwerp examens?

Mellenbergh is begin jaren zestig tijdens zijn studie psychologie begonnen als stagiair bij prof.dr. A.D. de Groot. De Groot zal een belangrijke plaats gaan innemen in het werkzame leven van Mellenbergh. In het souterrain van het huis van De Groot, dichtbij het Vondelpark, werkte hij aan de kwaliteit van testen en toetsen. De Groot woonde zelf op de bovenste verdiepingen en kwam geregeld rond 11 uur's ochtends in zijn kamerjas binnen om te kijken of zijn handgeschreven teksten, die hij de avond daarvoor rond drie uur's nachts op het bureau van zijn secretaresse had gelegd, goed waren uitgetypt.

Dit was nog de tijd van de ponskaarten en de ponsbanden. Als gebruiker moest je daar gegevens intikken zodat een computer het kon verwerken. De vaak meterslange ponsbanden leverde je bij het mathematisch centrum in en, als je geluk had, dat wil zeggen als je



Prof.dr. Gideon Jan (Don) Mellenbergh (1938) studeerde psychologie aan de Universiteit van Amsterdam en verkreeg in 1965 het diploma. In 1971 promoveerde hij bij prof.dr. A.D. de Groot op het proefschrift *Studies in studietoetsen*. Met uitzondering van een kwartaal in de Verenigde Staten en vier jaar aan de Universiteit Utrecht heeft Mellenbergh carrière gemaakt aan de Universiteit van Amsterdam. Op dit moment is Mellenbergh als emeritus-hoogleraar nog steeds verbonden aan de Programmagroep Psychologische Methodenleer.

Mellenbergh heeft onderzoek gedaan naar de constructie en analyse van psychologische tests en studietoetsen, onzuiverheid van vragen (*differential item functioning*, afgekort DIF), cesuurbepaling, psychometrische begrippen en de relatie tussen verschillende psychometrische modellen. Op zijn naam staan meer dan 150 publicaties, waaronder

IJking van de AVL, de Nederlandse bewerking van de *Study of Values* (1970), *Testleer en testconstructie* (1989), *Beslissen met tests en studietoetsen* (1993) en *Modellen op de divan: Psychometrische, besliskundige en statistische modellen belicht* (1999). Zie voor een compleet overzicht: <http://home.medewerker.uva.nl/g.j.mellenbergh/page2.html>. Daarnaast heeft hij 88 promovendi begeleid bij het schrijven van hun dissertaties, adviseert hij studenten en onderzoekers op het gebied van methodologie op verschillende terreinen als psychologie, onderwijskunde, tandheelkunde en psychiatrie.



Destijds waren er ook al twijfels over het niveau van leerlingen

geen invoerfouten had gemaakt, kreeg je de volgende dag de uitkomst van de berekeningen. Kom daar maar eens mee in dit tijdperk, waarin de computers honderden keren sneller zijn dan toen en tevens een honderdvoud kleiner. Bovendien kom je er tegenwoordig meteen achter wanneer je een tikfout maakt.

In het souterrain werkte Mellenbergh aan de vertaling en validatie van de Nederlandse versie van de *Study of Values* (S.O.V.). Dit resulteerde in zijn eerste publicatie in 1966 in het *Nederlands Tijdschrift voor de Psychologie*. Daarnaast deed hij onderzoek naar de kwaliteit van toetsen, in die tijd nog uitsluitend gedefinieerd in termen van KR20, Rit, Rat en p-waarden. Kortom, klassieke testtheorie. De Groot had in die tijd, onder invloed van een lang verblijf bij ETS, net de Amsterdamse schooltoets ontwikkeld. Dat was een toets bestaande uit vierkeuze-vragen, destijds een unicum in Nederland, die tot doel had de schoolprestaties van leerlingen van Amsterdamse lagere scholen in kaart te brengen. De getoetste vakken waren rekenen, taal, aardrijkskunde en geschiedenis. Destijds bestonden er ook al twijfels over het niveau van leerlingen. De Groot en zijn medewerkers kregen opdracht dat veronderstelde lage niveau in kaart te brengen. Deze toets kan worden gezien als de voorloper van de huidige Cito-toets. Mellenbergh hielp mee aan de verwerking en deed onderzoek naar de passendheid van het Rasch-model op de data; dat model paste redelijk.

Na zijn afstuderen stelde De Groot hem aan bij de Afdeling Examen Techniek (AET) bij het Psychologisch Laboratorium van de UvA. Hier is hij met de nodige pressie door De Groot, want "er moest wel productie gedraaid worden," snel gepromoveerd op het proefschrift *Studies in studietoetsen*. Het is een proefschrift waarmee je heden ten dage niet voor de dag kunt komen. Niet vanwege de inhoud, de kwaliteit staat niet ter discussie, maar vanwege de uiterlijke vorm: A4 formaat en de bladzijden zijn gestencild in smoezelig grijs/zwart. In het proefschrift werd voor het eerst het Rasch-model gebruikt bij het analyseren van toetsen. Naast het schrijven aan het proefschrift had Mellenbergh regelmatig overleg met een leesgroepje wiskundigen, gaf hij werkcolleges psychometrie en programmeerde hij in ALGOL-60. Omdat hij in het onderwijs

werkzaam was, hoefde hij niet in militaire dienst. Mellenbergh werd in 1971 benoemd tot lector bij de Universiteit Utrecht. Studenten roerden hun staartje en eisten democratisering, *one man one vote*. Studenten vonden vooral de medewerkers die zich bezig hielden met testen reactionair. Testen waren volgens hun ideologie immers elitair, omdat die onderscheid maken tussen mensen waarbij de ene mens als beter voor de dag komt dan de ander. Mellenbergh kan smeulige verhalen vertellen uit die tijd. Maar het eind van het liedje was wel dat hij na enkele jaren terugkeerde naar de UvA, waar hij in 1982 de hoogleraarplaats overnam van professor F.N. Kerlinger. De Groot was toen al enkele jaren eerder vertrokken naar Groningen. Hij had geen zin in de continue discussie met studenten.

Wie zijn uw leermeesters geweest?

Over het antwoord hoeft Mellenbergh niet lang na te denken: De Groot als algemene methodoloog en Bob van Naerssen¹ die meer een specialist in de psychometrie was. Mellenbergh vindt zichzelf een generalist. Hij heeft veel verschillende dingen gedaan en is daardoor geen echte specialist geworden.

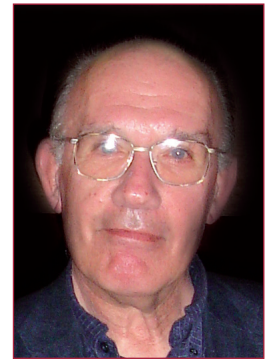
Wie beschouwt u als vooraanstaand in de examenwereld?

Ook hier is de eerste reactie De Groot en Van Naerssen. Beiden zijn overleden. In latere jaren is daar De Gruijter² bijgekomen. Op internationaal gebied noemt hij namen als Lindquist³, de eerste *editor* van *Educational Measurement* (1951), Cronbach⁴, Lord⁵ (die zei nooit wat!) en Novick⁶. Maar dan praat hij over het verleden. In het heden zijn volgens hem Paul Holland⁷, Wim van der Linden⁸ en Norman Verhelst⁹ de koplopers. Maar Mellenbergh noemt naast namen van personen ook instituten waarvan hij vindt dat ze een vooraanstaande positie innemen in de examenwereld: het Amerikaanse ETS (Educational Testing Service) en natuurlijk het Nederlandse Cito. Hij vindt Cito een prachtig instituut, waar veel data beschikbaar zijn. Het is jammer, zo gaat hij verder, dat het bestuur van Cito niet in meerdere mate bevordert dat er wordt gepubliceerd.

Wat drijft u?

"Ik houd van publiceren." Er staan dan ook inmiddels meer dan 150 publicaties op zijn naam. Mellenbergh is

De overheid moet een veel belangrijker rol gaan spelen



zeer gefascineerd door begrippen die in verschillende vakgebieden worden gebruikt, maar die hetzelfde betekenen. Of begrippen die in verschillende vakgebieden juist iets anders betekenen, zoals het begrip precisie. Precisie betekent in de klassieke testtheorie betrouwbaar, maar in de itemresponsetheorie betekent het informatie. Hij probeert de begrippen met elkaar te verbinden door ze te analyseren en terug te brengen tot hun essentie.

Daarnaast vindt Mellenbergh het begeleiden van promovendi en onderzoekers erg leuk. "Je kunt de mensen zien groeien. Ze komen vaak wat onzeker binnen en de meesten schrikken van het vele commentaar op hun eerste artikel. Maar aan het eind zijn ze veel zelfverzekerder en kunnen ze veel beter schrijven." Mellenbergh geeft promovendi en onderzoekers vaak de tip om na een review het commentaar een week in de la te laten liggen en er dan eens serieus naar te kijken. Hij heeft inmiddels 88 promoties begeleid, hetzij als eerste, hetzij als tweede promotor. Van die promovendi zijn er negentien hoogleraar geworden. En op zo'n getal kan iedere promotor trots zijn.

Ondanks dit grote aantal leerlingen (studenten en promovendi) vindt Mellenbergh niet dat hij school heeft gemaakt. Hij heeft geen leerlingen opgeleid die op hem zijn gaan lijken. Het vakgebied, zegt hij, is veel meer gespecialiseerd geworden en daardoor minder overzichtelijk. Je moet je als beginnend onderzoeker snel gaan specialiseren, anders onderscheid je je niet van anderen. Mellenbergh beschouwt zichzelf meer als een generalist die het vakgebied psychometrie heeft helpen ontwikkelen. Dat het zich dan in latere jaren meer gaat verspreiden over een aantal subdisciplines is niet meer dan logisch.

Wat voor visie heeft u op examens en examineren? Wat is het doel?

Het eindexamen heeft drie hoofdfuncties: afsluiten of certificeren, toelating tot het vervolgonderwijs en richting geven aan het onderwijs (de *feedback* en *feedforward* functie). "De overheid moet een veel belangrijker rol gaan spelen. Ze laten het nu te veel gaan. Ze moeten niet zo veel verantwoordelijkheid bij de scholen leggen, maar het meer in eigen hand houden en zelf de standaarden stellen."

Mellenbergh heeft met veel plezier deelgenomen aan

de profielcommissie die ontwerpadviezen heeft opgesteld voor onder andere het VO-eindexamen. Uit onderzoek bleek dat er gemiddeld één punt verschil was tussen het schoolonderzoek en het centraal eindexamen. Vooral bij levenbeschouwelijke scholen en privéscholen was het verschil groot. De commissie heeft voorgesteld om het schoolonderzoek en het centraal eindexamen uit elkaar te halen. Een ander voorstel was dat leerlingen voor zowel het centraal eindexamen als het schoolonderzoek een voldoende moeten halen om het eindexamen te behalen. Dan pas wordt het schoolonderzoek een echt examen. Scholen hebben in die opzet zelf de vrijheid om het schoolonderzoek in te richten. Als een school zich wil profileren in bijvoorbeeld kunst of sport, hebben ze met dit voorstel de mogelijkheid dat te doen binnen het schoolonderzoek. Het advies was om alle cognitieve vakken in het centraal eindexamen op te nemen, door de overheid geregeld waardoor het niveau van het onderwijs veel beter te bewaken zou zijn. Het voorstel is opgenomen in het rapport van de profielcommissie, maar kreeg tegenwerking van het ministerie. "Er zaten goede dingen in, maar misschien was het niet politiek genoeg," zegt Mellenbergh berustend. Dat de huidige examens met een gecombineerd eindcijfer van schoolexamen en centraal examen als toelatingsexamen functioneren tot de universiteit, vindt Mellenbergh een idiote situatie. Het eindexamen moet doen wat het moet doen. Namelijk een oordeel geven over de prestaties in de afzonderlijke vakken die nodig zijn voor het onderwijs dat volgt. Bijvoorbeeld voor een studie psychologie moeten de eerstejaars studenten Engels kunnen lezen, Nederlands kunnen schrijven en een basiskennis hebben van wiskunde en biologie. Je moet een onvoldoende voor wiskunde op het centraal examen niet kunnen compenseren met een hoog cijfer op een schoolonderzoek. Het is goed dat de overheid hier ingrijpt.

Innovatieve examens geven volgens Mellenbergh ook een nieuwe impuls voor nieuwe psychometrie. Psychometrici staan nu voor een nieuwe uitdaging: hoe maak je goede examens met nieuwe technologie?

Welke missers heeft u gemaakt?

Mellenbergh vindt dat hij de moderne ontwikkelingen op het gebied van computertechnologie niet goed heeft bijgehouden en noemt zichzelf een computernul.



Er is nu zoveel mogelijk met behulp van de computer

“Misschien heb ik te veel aandacht voor de begrippen gehad. Er is nu zoveel mogelijk met behulp van de computer, zoals computer adaptieve toetsen en simulaties. Dit geeft zulke prachtige testsituaties.” Zelf zal hij geen voorloper in deze ontwikkelingen zijn. Hij noemt zich zelf niet behoudend, maar hij bekent wel dat hij nog steeds houdt van de oude programmeertaal ALGOL-60.

Geef eens een boodschap aan de lezers van het blad.

“Houd de moderne ontwikkeling in de gaten. Laat je niet te veel afschrikken door bureaucratie. Examens zijn een sterk onderdeel voor de onderwijswereld. De overheid moet wel de randvoorwaarden creëren. Een examen moet gewoon goed zijn. Als dat alles wordt verworpt, zal Nederland zijn vooraanstaande positie op het examengebied behouden.” Mellenbergh doet ook nog een constatering. Hij ziet met lede ogen gebeuren dat de toelatingsfunctie van het VO-eindexamen in het gedrang komt. Dat blijkt, bijvoorbeeld bij psychologie,

uit de vele klachten over het Nederlands schrijven en de gebrekkige wiskundekennis van eerstejaarsstudenten. De kwaliteit van het examen is goed, maar het sluit niet goed aan bij de eisen van het hoger onderwijs, aldus Mellenbergh.

Literatuur

Bruning, L., Meijs, L., Veldhuis, J.G.F. (red.). (2006). *Bruggen tussen Natuur en Maatschappij*. Ontwerpadvies. Profielcommissie Natuur en Techniek / Natuur en Gezondheid en Profielcommissie Economie en Maatschappij / Cultuur en Maatschappij. Den Haag.
Mellenbergh, G.J. (2009). Robert F. van Naerssen: Een eminent examengeleerde van het eerste uur. *Examens*, 6(3), 26-29.

□ Mw. drs. J. ter Burg en de heer dr. H.J.M. van Berkel zijn redacteur respectievelijk hoofdredacteur van *EXAMENS*. E-mail: jterburg@xs4all.nl.

¹Zie *Examens* 2009, 6(3), 26-29.

²Zie *Examens* 2010, 7(4), 24-29.

³Everett Franklin Lindquist (1901-1978) was professor of education aan de University of Iowa.

⁴Lee J. Cronbach (Fresno, Californië 1916-2001), Amerikaanse psycholoog, bekend door zijn belangrijke bijdragen aan de ontwikkeling van psychologische methoden en technieken (Cronbach's alpha).

⁵Frederic M. Lord (1912-2000) was psychometricus bij ETS. Op zijn naam staan twee belangrijke boeken: *Statistical Theories of Mental Test Scores* (1968), geschreven met Melvin Novick en Allen Birnbaum, en *Applications of Item Response Theory to Practical Testing Problems* (1980).

⁶Melvin R. Novick (1932-1986) werkte bij ETS (1963-1970) voordat hij aan de University of Iowa werd benoemd als 'professor of education and statistics' (1970-1986). Van 1970-1974 was hij bovendien werkzaam voor het American College Testing Program (ACT).

⁷Paul Holland Ph.D. is directeur van het 'Center for Statistical Theory and Practice' van ETS.

⁸Zie *Examens* 2010, 7(2), 25-28.

⁹Zie *Examens* 2010, 7(3), 25-29.

In het volgende nummer van **EXAMENS** komen onder andere de volgende onderwerpen aan bod:

- Ik spiek, jij spiekt, wij spieken
- Kwaliteit van toetsing in het wetenschappelijk onderwijs
- Op een diploma moet je kunnen vertrouwen
- Kopstukken uit de examenwereld: Jules Peschar

EXAMENS 2, 2011 verschijnt in juni