

Studentenpreferenties m.b.t. beoordelingsvormen in het algemeen, evaluaties van een „multiple choice-toets” in het bijzonder: een kort verslag¹

Sociologisch Instituut der Rijksuniversiteit te Leiden

L. RADEMAKER

In de periode april 1967 tot september 1969 zijn er door het Sociologisch Instituut te Leiden een 20-tal „multiple choice”-tentamens „Inleiding tot de Sociologie” afgenomen van een 1000-tal studenten merendeels eerstejaars, te onderscheiden in 3 categorieën, nl. a) westers sociologen, b) niet-westers sociologen en cultureel antropologen, en c) psychologen en pedagogen. De door ons onderscheiden categorieën werd een gedeeltelijk identiek tentamen voorgelegd van 30 tot 50 m.c.-vragen met vier antwoordmogelijkheden.

Tegelijk met de toets werd een evaluatie-formulier aangeboden, dat door de docenten was opgesteld om o.a. na te gaan of er wellicht bijzondere weerstanden onder studenten bestonden tegen continuering van de desbetreffende meerkeuzebeoordelingsvorm, alsook om kritisch commentaar te verkrijgen op de afzonderlijke tentamen-vragen, op grond waarvan eventueel de kwaliteit van de items kon worden verhoogd.

Dit evaluatie-formulier kon vrijwillig worden ingevuld en — indien gewenst — anoniem worden ingeleverd. Het grootste gedeelte der studenten vulde het formulier zonder bezwaar in, nl. 895 van de ongeveer 1000.

We zullen hieronder een kort en globaal verslag uitbrengen, en wel per vraag. Aan de genoemde differentiatie in drie categorieën studenten, die ter informatie werd vermeld, gaan we in het volgende over het algemeen voorbij. Volledigheidshalve zullen we ook de gegevens noteren m.b.t. die vragen uit het formulier, welke in het kader van dit verslagje ietwat minder relevant zijn (b.v. vr. 1).

vraag 1: *Hoeveel tijd had u nodig om deze toets te maken?*

Per vraag had men gemiddeld ongeveer 2,5 minuut nodig. De verschillen in benodigde tijd voor de toets als geheel (zoals vermeld variërend van 30 tot 50 vragen) waren opvallend: vele studenten waren binnen 45 minuten reeds klaar, een groot deel echter maakte gebruik van de beschikbare 2 à 3 uur.

vraag 2a: *Wanneer u mocht kiezen tussen een mondeling tentamen, een „gewoon” schriftelijk tentamen met open vragen, en een multiple choice-tentamen, welke beoordelingsvorm zoudt u dan prefereren (prekodering, zie tabel 1).*

De preferentie-verdeling bleek als volgt:

Tabel 1

Preferenties van studenten m.b.t. drie beoordelingsvormen (mondeling, 'gewoon' schriftelijk en multiple choice); abs. en in %.

	abs.	%
mondeling	213	24
gewoon schrift.	213	24
multiple choice	376	42
geen voorkeur	69	8
geen antw./div.	24	3
totaal	895	101

De voorkeur blijkt duidelijk uit te gaan naar de multiple choicebeoordelingsvorm (het verschil met de andere twee beoorde-

lingsvormen, die een gelijke preferentie verkregen, is significant²: chi-kwadraat is ruim 65.

De exact gelijke preferentie voor mondeling en gewoon schriftelijk kan, opmerkelijk genoeg, óók worden waargenomen bij de afzonderlijke categorieën studenten.

Slechts bij de westers sociologen was de hogere preferentie voor multiple choice niet significant (33, tegenover 30 en 30%). Bij de andere categorieën echter wél (bij de pedagogen en psychologen het duidelijkst (50, tegenover 20 en 20%), bij de niet-westers sociologen en cultureel-antropologen eveneens zeer duidelijk (43, tegenover 21 en 21%).

vraag 2b: *En waarom prefereert u persoonlijk het boven omcirkelde?* (open vraag, zie tabel 2).

Door klassifikatie konden de volgende redenen worden onderscheiden: de mogelijkheid om een gegeven antwoord in eigen bewoordingen te beargumenteren of in het algemeen toe te lichten, het persoonlijk contact, de korrektie-mogelijkheid van een in eerste instantie minder gelukkig antwoord, objectiviteit van de beoordeling, efficiënte beantwoording van de vragen, het rustig werken tijdens het tentamen en de relatief makkelijke bestudering van de getoetste stof. De genoemde kode-rubrieken treft men in afgekorte vorm aan in de volgende tabel.

Zoals te verwachten vermeldden degenen die de voorkeur gaven aan een mondeling of gewoon schriftelijk tentamen in het bijzonder de „mogelijkheid van argumentatie enz” als (één van de) reden(en) voor hun preferentie en degenen die multiple choice prefereerden vooral de reden „objectiviteit”.

Opvallend is de relatief geringe betekenis die de factor „objectiviteit” bij de preferentie voor gewoon schriftelijk verkreeg. Saillant is bovendien het feit dat „relatief makkelijke bestudering van de stof” door degenen die multiple choice prefereerden vrij frekwent wordt genoemd, maar t.a.v. de andere preferenties niet of nauwelijks wordt vermeld.

vraag 3: *Hoe vond u dit bepaalde multiple choice-tentamen in het algemeen?*

(drie aspecten: plezierigheid, makkelijk-/moeilijkheid en mate van „goede” toetsing van wat je weet: geprecodeerd, zie tabel 3)

Het desbetreffende tentamen werd aldus door de studenten overwegend aan de onplezierige kant gevonden (en wel door 32% — nl. 25+7-, tegenover 22% — nl. 17+5— aan de andere kant van „neutraal”; dit verschil is significant: chi-kwadraat ruim 12 (de percentages „diversen” en „geen antwoord” laten we, evenals bij tabel 4 en 5, bij onze globale conclusies buiten beschouwing).

Tabel 2

Redenen voor preferenties m.b.t. drie beoordelingsvormen: mondeling, 'gewoon' schriftelijk en multiple choice

mog. argum.	pers. kont.	korr. mog.	obj. t.	effic. beantw.	rustig werken	stof. rel. makk. te best.	div.	g.a.	toaal %	abs.
<i>mondeling</i>										
33	21	12	18	—	—	—	13	3	100	191
<i>gewoon schriftelijk</i>										
50	—	—	11	—	22	—	14	3	100	241
<i>multiple choice</i>										
—	—	—	31	22	7	23	10	7	100	399

(831)³

Tabel 3

Waardering voor het afgelegde multiple choice-tentamen m.b.t. het aspect 'plezierigheid'

plezierig	tamelijke plezierig	neutraal	tamelijk onplezierig	onplezierig	g.a.	totaal %	abs.
5	17	41	25	7	5	100	641 ¹

Tabel 4

Waardering voor het afgelegde multiple choice-tentamen m.b.t. het aspect 'makkelijkheid-moeilijkheid'

makkelijk	tamelijk makkelijk	neutraal	tamelijk moeilijk	moeilijk	g.a.	totaal %	abs.
0.16	4	23	55	14	4	100	633

Een overgrote meerderheid der studenten vond aldus het afgelegde tentamen aan de moeilijke kant (en wel 69% —nl. 55+14—, tegenover ruim 4% —nl. 0.16+4— aan de andere kant van „neutraal”; dit verschil is overduidelijk significant: chi-kwadraat ruim 357).

Tabel 5

Waardering voor het afgelegde multiple choice-tentamen m.b.t. het aspect 'mate van "goede" toetsing wat je weet'

goed	ongeveer	redelijk	matig	slecht	g.a.	totaal %	abs.
11	19	33	24	6	7	100	632

De uitslagen links en rechts van „redelijk” zijn gelijk, en wel 30% (nl. 19 + 11 en 24 + 6). Men was van oordeel dat het tentamen redelijk toeste wat men wist.

Overigens zal de student in deze wel eens een beoordelingsfout hebben gemaakt.

Dat de beantwoording op de gestelde vraag niet representativiteit van de toets aanduidt, behoeft nauwelijks toelichting: de student, die slechts het niet ter zake doende uit de stof heeft bestudeerd, zal m.b.t. een representatieve, op — volgens de docent — relevante thema's gerichte, toets van mening zijn dat het tentamen slecht toetst wat je — i.c. hij — weet.

Wanneer t.a.v. deze vraag overwegend op

de kode „slecht” was geskoord, dan had de docent o.a. kunnen bekijken of a) de toets wel representatief was; b) de studenten wellicht een andere relevantie-bepaling bij de bestudering als uitgangspunt hadden genomen. Ingaande op dit laatste zouden we willen stellen dat een specifieke rele-

vantie-aanwijzing door de docent voor de student attractief is en in het algemeen kan worden aanbevolen.

vraag 4: *Hebt u opmerkingen met betrekking tot bepaalde vragen?*

In dit verband willen we alleen opmerken dat het kritisch commentaar van studenten op enkele, volgens hun inzicht, onjuiste of minder geslaagde vragen, voor de docent bijzonder nuttig kan zijn bij zijn poging tot kwaliteitsverbetering van zijn voorraad vragen te komen. In ons geval werden vaak n.a.v. dit commentaar bepaalde vragen uit de item-kartotheek verwijderd, andere vragen werden „bijgeschaafd”. Een enkele keer werd, eveneens

op basis van de studenten-evaluatie, bij de berekening van de uitslag een vraag aan elke deelnemer goed gerekend, of meer dan één alternatief als goed aangemerkt.

vraag 5: Algemene opmerkingen

Van de mogelijkheid om algemene opmerkingen te maken werd nauwelijks gebruik gemaakt. Bovendien konden vele notities in deze reeds worden afgeleid uit de antwoorden op voorgaande vragen. Tenslotte: de algemene suggesties van de studenten vertoonden een dermate hoge diversiteit dat hieruit geen enkele beleidskonsekwentie kon worden ontleend.

Slotopmerkingen

Gekonstateerd werd dat onder de studenten die wij een tentamen presenteerden de voorkeur uitging naar een multiple choice-beoordelingsvorm. Vastgesteld kon eveneens worden dat zij de desbetreffende toets overwegend als nogal onplezierig en moeilijk beschouwden, terwijl zij niet konden stellen dat het tentamen zonder meer goed toetste wat men wist. Aannemende dat de specifieke ervaring met het afgelegde tentamen van invloed is geweest op de preferentie m.b.t. beoordelingsvormen in het algemeen, zijn we geneigd te stellen dat de voorkeur voor de multiple choice-beoordelingsvorm niet geflatteerd is. Het feit echter dat niet is uitgesloten dat konkrete tentamens met een andere beoordelingsvorm op de genoemde drie punten door de studenten eveneens nogal „ongunstig” worden geëvalueerd, moet echter tot voorzichtigheid bij de konkludering op dit terrein aansporen.

Tenslotte: de vraag in hoeverre men eigenlijk rekening moet houden met studentenpreferenties in deze, zou ik slechts willen opwerpen en vooralsnog niet beantwoorden. Dit vraagstuk dient vanzelfsprekend te worden bezien binnen de gehele interne demokratiseringsproblematiek die thans aktueel is.

Noten

1. De heer P. Weenink, studenten-assistent in de sectie 'Inleiding tot en geschiedenis van de sociologie' van het Sociologisch Instituut te Leiden, wil ik hierbij hartelijk bedanken voor de controle op het cijferwerk, het uitvoeren van enkele significantie-berekeningen en het kritisch commentaar op het conceptverslag.
2. Bij de bepaling van de significantie is steeds een alpha-fout van 5% aangenomen.
3. Dit totaal is minder dan het totaal bij tabel 1 (895), aangezien enkele personen het evaluatie-formulier slechts gedeeltelijk hebben ingevuld.
4. Dit totaal is, evenals bij tabel 4 en 5, aanmerkelijk minder dan 895, niet alleen vanwege de gedeeltelijke invulling door sommige deelnemers, maar bovendien vanwege het feit dat deze vraag in 1967 nog niet in het formulier was opgenomen.

Curriculum vitae

Geboortjaar 1939.

Van 1960 tot 1966 studie Westerse sociologie te Leiden.

Als wetenschappelijk medewerker sinds dien verbonden aan het Sociologisch Instituut van de universiteit te Leiden. Publiceerde o.a. 'Objectieve prestatietoetsen, i.h.b. multiple-choice-toetsen', Wolters-Noordhoff, Groningen, 1969, 68 pag.