

PROEFWERKEN MET VIERKEUZE-VRAGEN: EEN AFRONDING.

E. WARRIES

Tot mijn spijt heb ik gekonstateerd dat om onverklaarbare redenen in mijn artikel in Paed. Stud. aflevering 1 ruim een pagina tekst is weggevallen. De beknopte handleiding voor het schrijven van schooltoets-items waarmee het stuk eindigde is zo wat àl te beknopt geworden. In overleg met de redactiesecretaris werd een „afronding” geschreven die hieronder volgt. De reeds beschreven eerste vier stappen zijn hier nog eens kort samengevat. Bij de samenstelling van het daarna volgende heb ik tevens gebruik gemaakt van enkele opmerkingen van drs. Sj. Sandbergen die een hoofdstuk schreef in het boek „Studietoetsen”¹

Beknopte handleiding: de eerste vier stappen

In het Research Instituut voor de Toegepaste Psychologie is in de loop der jaren op grond van eigen ervaring en die van anderen² een bruikbare procedure voor toetsconstructie ontwikkeld. Voor een vak of studie-onderdeel wordt een itemschrijfgroep van 5 of meer personen (waaronder een toetsconstructeur van professie) samengesteld. Men begint in die groep circa één jaar vóór de geprojecteerde toetsdag met het doorlopen van de hieronder beschreven weg, waarvan het begin dus zoals gezegd uitvoeriger reeds is beschreven in aflevering 1, 1969, van dit blad.

Eerste stap: In eerste bespreking(en) wordt vastgelegd over welke leerstof gevraagd zal worden en welke soort vragen in concreto bij de stof horen. De gewenste aantallen opgaven worden vastgesteld. Een voorlopige werkverdeling regelt de verantwoordelijkheden van de teamleden.

Tweede stap: Uitgaande van de eisen die men aan goede items stelt maken de itemschrijvers elk afzonderlijk hun huiswerk. Zij sturen dit vervolgens naar de centrale administratie.

Derde stap: De administratie maakt het ingezonden werk leesbaar, nummert de items, vermenigvuldigt de stukken en zendt deze rond aan alle leden van het betreffende team.

Vierde stap: De teamleden bestuderen grondig elkaars werk. Zij vergelijken niveau en inhoud van de items met de gemaakte afspraken over de leerstof. Ze toetsen de vormgeving aan de bekende eisen. Ze keuren

items af, vullen onafgemaakte items aan en amenderen waar nodig. Tegenvoorstellen worden zo mogelijk reeds ingezonden aan de administratie zodat ze in de eerste bespreking kunnen worden rondgedeeld.

5. *Bespreking over de items*

Het allereerste doel van de itembespreking is te komen tot een voorraad van items waarover alle aanwezige teamleden een gunstig oordeel hebben uitgesproken. Vervolgens kan als tweede (eind-)doelstelling nog gelden het doen van een uiteindelijke, voor de stof representatieve „evenwichtige” keuze uit het beschikbare materiaal: het aanwijzen van de opgaven die in de af te nemen toets worden opgenomen. Ten derde kan elke bespreking meebrengen dat een aantal nog te verrichten verdere werkzaamheden wordt verdeeld. Hieronder valt: het thuis verbeteren van voorlopig verworpen items, het systematisch aanvullen van de voorraad en het opnieuw bestuderen van andermans items.

Het is onmogelijk een algemeen geldig voorschrift te verschaffen voor de gang van zaken tijdens een itembespreking. De volgende agenda is dan ook een voorbeeld van hoe het kan:

1. De ingezonden items worden gescheiden in drie stapels: De nieuwgemaakte die volgens de auteur zelf nog niet af zijn of inderdaad nog incompleet zijn. Verder de nieuwgemaakte die compleet zijn. Ten derde de items waar kwantitatieve gegevens van bekend zijn. (We bepalen ons hier tot de nieuwgemaakte complete items).
2. De complete items worden stuk voor stuk aan de orde gesteld, waarbij men begint met zich af te vragen of aan de objectiviteits-eis is voldaan. Als er discussie ontstaat over het juiste antwoord, moet het item worden verworpen; eventueel krijgt de maker of iemand anders de verbetering van het item als huiswerk mee.
3. Nadat voor een bepaald item de eenstemmigheid over het juiste antwoord is vastgesteld, wordt zo mogelijk critiek geleverd na toetsing aan de eisen uit het rijtje van zes. Tegenvoorstellen worden met het oorspronkelijke item vergeleken, voorstellen om het item te verwerpen worden overwogen. Goedgekeurde items worden toegevoegd aan de voorraad.
4. Aan het einde van de bespreking wordt de inventaris opgemaakt en vastgesteld welke soort opgaven nog bijgemaakt zullen moeten worden om de eerder overeengekomen aantallen te bereiken.

Het is aanbevelenswaard een agenda als de bovenstaande — of een andere — in het hoofd te hebben alvorens men een itembespreking begint want anders bestaat het gevaar dat men in interessante discus-

sies verzeild raakt die géén items opleveren. Tijdens de bespreking — of besprekingen als er veel materiaal is — moet door de voorzitter een tamelijk straf beleid worden gevoerd, omdat een itembespreking gemakkelijk kan terugvallen naar een doeleindenbespreking — die al eerder gevoerd moet zijn. Een ander gevaar tijdens de bespreking is, dat men al te langdurig bij één itemvoorstel blijft stilstaan. In het algemeen geldt dat een discussie van vijf tot tien minuten over de vormgeving van een enkel item bewijst, dat hier geen goed item aan het „groeien” is. Op zeker moment moet de knoop dan worden doorgesneden door het besluit dit item te verwerpen — eventueel met de afspraak het thuis nog eens te bekijken. Tijdens de itembesprekingen mag geen zuinigheidsbeleid worden gevoerd; het weggooien van items moet niet als een verspilling worden gezien, maar moet worden beschouwd als een noodzakelijke selectie-procedure. De schrijver van het item dient zich er dan ook tevoren op in te stellen, dat minstens de helft van zijn werk zal worden „afgekraakt”. Tijdens de besprekingen moet — zeker bij beginnende itemschrijvers — de „itemmortaliteit” hoog zijn. Als deze lager is dan 50%, is men waarschijnlijk niet kritisch genoeg.

6. *Herhaalde itembesprekingen*

Hierboven werd aangenomen, dat men al een jaar vóór de toetsdag met de constructie van een „massaal” af te nemen toets moet beginnen. Deze bewering is niet alleen gegrondvest op de zienswijze dat het doelstellingengesprek waarmee het eerste overleg aanvangt bijzonder tijdrovend kan zijn, maar ook op de ervaring dat men in één itembespreking op geen stukken na klaarkomt met het samenstellen van de gehele toets. De ervaring heeft geleerd dat voor het samenstellen van een vrij grote toets — als b.v. de Amsterdamse Schooltoets — herhaalde besprekingen en schrijfpogingen nodig zijn. Het aantal items is te groot om in één bijeenkomst af te handelen. Daar komt nog bij dat de technische en democratische spelregels eisen dat items die door één teamlid gereviseerd zijn opnieuw in de voltallige vergadering bekeken moeten worden. Dat kost tijd — goed bestede tijd overigens. Verder kunnen meerdere besprekingen nodig zijn omdat de schrijfgroepen tot taak hebben een voorstel te doen over een afgeronde subtoets — eventueel met een parallelvorm. Ten slotte is het ook nog mogelijk dat de groepsaandacht moet besteden aan de kwantitatieve gegevens van een proefafname die gebruikt moeten worden om items te veranderen of de samenstelling van de toets te wijzigen.

Herhaalde besprekingen zullen een herhaling betekenen van het

schrijven, het vermenigvuldigen, het bestuderen en amenderen van elkaars werk en van de gedachtenwisseling daarna in de eigenlijke bespreking. In feite betekenen herhaalde itembesprekingen dus het opnieuw — en even gedisciplineerd — doorlopen van de voorafgaande stadia. In sommige gevallen zelfs zal men besluiten om terug te gaan naar het aanvangsstadium: de bespreking over wat, hoe, hoeveel en wanneer.

7. *De eindredactie*

Het is altijd goed om de geschreven producten waar de groepsleden een tijdlang mee hebben omgegaan en waarop ze misschien wel enigszins uitgekeken zijn, voor te leggen aan mensen van buiten de groep. Een kritische buitenstaander die goed geïnformeerd is (over onderwijs, leerstof en toetsconstructie) ziet de dingen waarschijnlijk heel anders dan de samenstellers zelf. Er kan in dit late stadium nog fundamentele critiek op details en het geheel worden geleverd. Een andere reden om buitenstaanders aan te trekken is de omstandigheid dat de leden van itemschrijfgroepen nu eenmaal niet zijn aangezocht om hun redactionele kwaliteiten. Itemschrijvers zijn zelden goede redacteuren. De juistheid van deze bewering is gemakkelijk aan te tonen: goede redacteuren zijn in het algemeen zeldzaam en er is geen reden om aan te nemen, dat itemschrijvers op dit punt uitblinkers zijn. Eerder integendeel. Omdat nochtans een eindbeslissing moet worden getroffen over de uiteindelijke vormgeving van de schooltoets, is een redactie-commissie het aangewezen middel.

Het werk van de eindredacteuren kan zeer eenvoudig gestructureerd zijn: „Ja” of „nee” zeggen tegen het eindproduct dat eventueel compleet met instructies voor proefleider en leerlingen wordt afgeleverd door de toetsconstructeur. Hun werk kan echter ook gecompliceerder zijn en bijvoorbeeld omvatten: het doen van voorstellen over de volgorde van de subtoetsen of van de items in een subtoets, het afwijzen van sommige items of het vragen om een andere vormgeving. De eindredacteuren kunnen ook tevoren vragen om een ruimere keus aan items of een reservevoorraad aan items, zodat zij onmiddellijk beslissingen kunnen nemen over de definitieve vorm.

Welke taak de redacteuren ook hebben, altijd is een duidelijke omschrijving daarvan noodzaak. Een andere noodzakelijke voorwaarde is: tijd. Redacteuren, die de semi-definitieve vorm van een Schooltoets in mogen zien, maar niet meer dan een week de tijd hebben om die — niet anders dan telefonisch — met andere redacteuren te bespre-

ken, verrichten geen doorwrocht werk. Een maand is wel nodig.

8. Proefafname

Het uitproberen van een Schooltoets of van een parallelvorm van een Schooltoets is noodzakelijk om een aantal in de handboeken beschreven redenen³

De proefafname kan weinig of veel pretenderen en naar verhouding arbeidsuren vragen van de test-specialisten, wier taak het is de afname te organiseren en de analyse uit te voeren. In het eenvoudigste geval verzamelt men in een klein aantal klassen wat gegevens van de onderwijzers of proefleiders over „hoe 't ging”. In het meest uitgewerkte geval, begint men met een landelijke steekproef van leerlingen te trekken en eindigt men met een „normering vooraf”, waardoor de uitslagen van de Schooltoets snel — op de dag van afname (of eigenlijk van scoring) in gestandaardiseerde scores, bijv. percentielscores, kunnen worden omgezet.

Indien men tijd genoeg heeft, kan men in de besprekingsgroepen aan de uitkomsten van een proefafname de nodige aandacht besteden. Op grond van bijv. de antwoordpercentages van de vier alternatieven bij één item, kunnen items verbeterd worden. Dit behoort evenwel tot de soort rekenkundige bewerkingen die wij hiervóór hebben aangekondigd niet te zullen bespreken.

Literatuur.

1. GROOT, A. D. DE EN NAERSSSEN, R. F. VAN. (red.) *Studietoetsen*. Groningen: Wolters Noordhoff, nog niet verschenen.
2. EBEL, R. L. *Measuring Educational Achievement*. New Jersey, Prentice Hall, 1965.
3. LINDQUIST, E. F. (ed) *Educational Measurement*, Washington: American Council on Education, 1951.