

AMSTERDAM, 21 november 1975

Beste mensen,

COWO-ers

Hierbij een concept voor een bibliografie m.b.t. alles wat met beoordeling en selectie te maken rond de propedeuse in het w.o. (binnen een paar dagen volgt de literatuurlijst ook nog, het stuk dat je nu in handen hebt bevat een heel summiere aanduiding van inhoud en eventueel belang van de onderscheiden categorieën).

De bibliografie, die te gelegener tijd van aantekeningenperitum voorzien zal worden, is niet bedoeld een volledig overzicht van beschikbare literatuur te geven, maar slechts de meest representatieve illustratieve of belangrijke bronnen te noemen.

Mocht je de indruk hebben dat bepaalde lievelingsliteratuur van jou ten onrechte niet in de lijst voorkomt, dan houd ik me voor die extra informatie van harte aanbevolen.

Ook commentaar op de korte toelichtende stukjes zou ik graag willen hebben, omdat ik zo snel mogelijk de bibliografie in deze zeer voorlopige vorm de deur uit zou willen doen. Dus graag opmerkingen in de trant van : kan dit niet wat duidelijker, waarom heb je dit soort bronnen niet opgenomen, is die categorie nou wel zo relevant, kan het niet wat uitgebreider of korter, zoals het zaakje nu in elkaar zit kan ik er niet mee werken maar misschien wel als je dit en dat...., etc. wat je zelf nog kunt bedenken.

Hartelijk dank,

Ben.

Selectief literatuur overzicht m.b.t. beoordeling en selectie, i.h.b. bij de propedeutische studiefase.

Ben Wilbrink.

C.O.W.O. Spui 21 Amsterdam oktober 1975.

Beoordeling en selectie in de meest ruime zin, althans zo ruim dat geen relevante aspecten buitengesloten blijven, dat is de opzet.

Belangrijkste uitgangspunt is dat ieder beoordelings- of selectieprobleem een beslissingsprobleem is, en dan ook langs besliskundige lijnen aangepakt dient te worden. De moeilijkheid met dit uitgangspunt is dat het, ondanks de publicatie van Cronbach en Gleser in 1957 van Psychological Tests and Personnel Decisions, nog steeds een grote uitzondering is beoordelings- en selectieproblemen besliskundig aangepakt te zien. Bestaande literatuur wordt daarmee nog niet als irrelevant gekenmerkt, maar blijkt wel moeilijk te plaatsen binnen het hier te hanteren indelings systeem voor ons bronnen materiaal.

Grofweg zal een ieder die voor een bepaalde beslissing staat

1) zijn alternatieven met alle bijbehorende mogelijke uitkomsten moeten beschrijven;

2) de utiliteit en waarschijnlijkheid van iedere mogelijke uitkomst moeten schatten; en

3) de verbanden tussen de hem ter beschikking staande informatie en de waarschijnlijkheid van bepaalde uitkomsten moeten schatten, waarna de 'beste' beslissing uitgerekend kan worden. Technieken die direct op 1), 2) of 3) van toepassing zijn, zullen dan ook onder resp. 1), 2) of 3) in deze bibliografie opgenomen worden. Naast beschrijving van technieken zoeken we naar theoretische stukken, verslagen van empirische bevindingen, e.d.

Het bijzondere karakter van deze drie rubrieken hebben we beklemtoond door ze te nummeren, waar de overige rubrieken met letters aangeduid zullen worden.

In het algemeen is het zo, dat voor verschillende 'beslissers' de alternatieven niet gelijk zijn, de mogelijke uitkomsten verschillen, utiliteiten ook van eventueel dezelfde uitkomsten nog sterk uiteen kunnen lopen, en de schatting van verbanden tussen gegeven informatie en mogelijke uitkomsten niet gelijk hoeft te zijn. Of korter gezegd: het beslissingsmodel en de invulling ervan hangt af van wie, of welke instantie, de beslissing neemt. Daaruit volgt direct dat we voor onze propedeuse met alles daar direct omheen vier verschillende groepen van beslissingsmodellen kunnen onderscheiden, namelijk modellen die passen bij de (individuele) student, de docent, de (sub)faculteit, en politieke lichamen (Uraad, kabinet, parlement e.d.). Het onderscheid is niet alleen maar van academisch belang: het is eenvoudig aantoonbaar dat in een bepaalde beslissingssituatie de belangen van genoemde vier groepen niet op één lijn liggen, en dat daarmee ook de 'beste' beslissing voor de één, dat niet is voor de anderen.

Dit zijn de categorieën A, B, C, en D resp.

Het is duidelijk dat we er met deze logische indeling nog niet zijn. Vandaar dat we een nogal fors aantal categorieën toevoegen,

die naast of voorafgaand aan besliskundige uitwerking van bepaalde problemen van belang zijn. Daaronder vallen beschikbare technieken en beschrijving van methodologische problemen, ethische aspecten, onderwijskundige principes, e.d. Ook hier geldt dat de indeling prematuur zal blijken te zijn, en later ongetwijfeld aangepast moet worden. De volledige lijst wordt op de volgende blz. nog eens gegeven.

Ondanks de hier gegeven toelichting, zal vaak niet duidelijk zijn wat nu precies met de verschillende categorie aanduidingen bedoeld wordt. Hetzelfde geldt eigenlijk ook voor de titels van opgegeven literatuur, die in veel gevallen niet voor zichzelf spreken. Pas wanneer in een later stadium de aantekeningen bij de diverse items gereed zijn, kan een ook in dit opzicht meer bevredigende bibliografie gegeven worden. Voorlopig bedoelen we niet meer dan aan te geven welke bronnen mogelijk relevant zijn (ook in de zin eventueel van liever niet na te volgen voorbeelden). De bibliografie beoogt ook niet meer dan een aantal belangrijke publicaties, en mogelijk een aantal minder relevante, te geven, waarbij volledigheid niet nagestreefd is. In het algemeen zal men op grond van de gegeven bronnen wel in staat zijn om voor zichzelf een meer volledig literatuuroverzicht te maken.

1. beschrijving van beslissingsalternatieven en mogelijke uitkomsten. 2. kwantificering van utiliteit en waarschijnlijkheid van uitkomsten. 3. verband van beschikbare informatie met waarsch. van uitkomsten.

- A. een en ander m.b.t. studenten als beslissingnemers.
- B. idem docenten
- C. idem (sub)faculteit
- D. idem beleidsinstanties (Uraad, kabinet, parlement)
- E. normatieve en ethische aspecten bij beoordeling en selectie.
- F. rechtspositie regelingen voor onderwijs deelnemers.
- G. gelijke kansen problematiek.
- H. methodologische problemen.
- J. beschikbare (mogelijk bruikbare) technieken en modellen.
- K. onderwijskundige uitgangspunten.
- L. feedback functie van beoordelingen (incl. motivationele aspecten).

- N. onderwijsfilosofie (globale doelstellingen, maatschappelijke opvattingen
- O. beoordelingsgewoonten (oude en nieuwe, beschrijving en onderzoek).
- P. evaluatie van onderwijs, curriculum beschrijving.
- Q. (de betekenis van) persoonlijke verschillen in het onderwijs.

Voor iedere categorie zal eerst een korte beschrijving gegeven worden, en daarna telkens de bijbehorende literatuur. Zoals gezegd, zal in een volgende editie ieder literatuur item van een korte kritische samenvatting en bespreking worden voorzien.

Categorie 1.

Beschrijving van beslissingsalternatieven en mogelijke uitkomsten.

Een systematische aanpak van welk beslissingsprobleem dan ook vereist dat alle beslissingsalternatieven zorgvuldig geïnventariseerd worden. Daarvoor zijn geen speciale technieken nodig, slechts wat vindingrijkheid om niet al bij voorbaat uitsluitend traditionele paden te betreden. Hetzelfde geldt voor de inventarisatie, bij ieder alternatief, van de mogelijke uitkomsten waartoe dat alternatief zou kunnen leiden.

Een beschrijving van deze in wezen common sense activiteit is te vinden in teksten als van Raiffa 1968, en Lindley 1971, beide aan de orde komend onder categorie 2 (voortaan tussen haakjes vermeld als volgt: Raiffa 1968 (2).)

Dit is een bijzondere categorie, in deze zin dat er geen literatuur onder dit hoofd op te geven valt. Desondanks wordt de indeling gehandhaafd. om uit te laten komen dat hier toch wel van een leemte in onderwijskundig onderzoek sprake is. Er is bijv. behoefte aan met zorg samengestelde beschrijvingen van alternatieven en mogelijke uitkomsten voor een aantal in het onderwijs vaak voorkomende beslissingsituaties, beschrijvingen die het de diverse betrokkenen makkelijker kunnen maken hun beslissingsstrategie te bepalen, en die daarnaast ook voor (beschrijvend) onderzoek als uitgangspunt zouden kunnen dienen.

Ook kan gedacht worden aan cursuspakketten, zoals bijv. in de Verenigde Staten door de College Entrance Examination Board ontwikkeld om high school leerlingen wat beter beslagen op het beroeps en studiekeuze ijs te zetten: Gelatt e.a. 1972 (A), en Gelatt e.a. 1973 (A).

Het belang van een goede inventarisatie van beleidsalternatieven, en de resultaten die verschillende mogelijke keuzen kunnen opleveren, voor een verantwoorde beslissings strategie lijkt evident. Toch zien we in de praktijk maar al te dikwijls dat voor de hand liggende alternatieven niet bij de keuze betrokken worden. Een voorbeeld daarvan zijn de vele pogingen ondernomen om de zak slaag problematiek bij tentamens op te lossen door de 'objectiviteit' van de cesuurbepaling te vergroten, maar waarbij geen aandacht wordt besteed aan de vraag of bijv. studenten met een onvoldoende niet beter in staat gesteld zouden moeten worden om deze onvoldoende te compenseren door een (ev. extra) ander studieonderdeel te doen in de tijd die hij anders aan herhaling van de oude stof zou moeten besteden (zie Wilbrink dec. 1974 (J)).

Categorie 2.

Kwantificering van utiliteit en waarschijnlijkheid van uitkomsten

Het is hier niet de plaats om een uitleg te geven over het begrip utiliteit. Een heel zorgeloze omschrijving is, dat utiliteiten verbonden aan verschillende mogelijke uitkomsten van verschillende alternatieven die te kiezen zijn, de relatieve en subjectieve

waarde aanduiden die de uitkomsten ten opzichte van elkaar hebben. Hoe een en ander ook zij, een heel korte uiteenzetting is te vinden in Wilbrink sept. 1975, en een vollediger behandeling in Raiffa 1968 of Lindley 1971

Soms zijn uitkomsten te evalueren in guldens, dat is dan makkelijk, maar het is geenszins noodzakelijk om eerst iedere mogelijke uitkomst een (fictieve) geldelijke waarde toe te kennen. Voorbeelden van monetaire waarderingen zijn te vinden in Roche (afgedrukt in Cronbach en Gleser 1965), in Witmer 1970, en verwante literatuur.

Voor beleidsinstanties zijn indicators voor effectiviteit van het gegeven onderwijs van belang om mogelijke uitkomsten te kunnen waarderen. Als voorbeeld daarvan (schatting van effectiviteit van onderwijs) zie Marco 1974, en de literatuur vermeld in categorie P: evaluatie van onderwijs.

De waarschijnlijkheid van een bepaalde uitkomst is nooit precies bekend. Met name wanneer het gaat om eenmalige beslissingen, moet daar nogal een slag naar geslagen worden. Mogelijk kan in de literatuur over subjectieve waarschijnlijk een enkel aanknopingspunt gevonden worden om technieken te ontwikkelen die de beslissingnemer bij kunnen staan in het maken van reële schattingen. Zie bijv. Van Naerssen 1962, Vlek 1973

Bijzondere problemen kunnen zich voordoen wanneer meerdere personen met elkaar tot overeenstemming moeten komen over de utiliteit van verschillende mogelijke uitkomsten. Voor zover er al technieken ontwikkeld zijn om dit probleem benaderbaar te maken, verkeren die nog in een vroeg experimenteel stadium. Zie bijv. Davis 1973.

In de meeste gevallen zal de waarschijnlijkheid van bepaalde uitkomsten geschat moeten worden op basis van in experimenteel onderzoek verkregen empirisch materiaal. De aanpak van dergelijk valideringsonderzoek is in principe voldoende bekend, zie bijv. Cronbach en Gleser 1965 (C). Maar dit is een onderwerp waarvoor categorie 3 ingeruimd is.

Categorie 3.

Verband van beschikbare informatie met waarschijnlijkheid van uitkomsten.

De waarschijnlijkheid van uitkomsten moet geschat worden. Dergelijke schattingen kunnen nauwkeuriger zijn, wanneer meer informatie beschikbaar is: zie voor een algemene bespreking Raiffa 1968 (2) of Lindley 1968 (2), en voor een behandeling die wat meer afgestemd is op de problematiek waar we ons bij deze bibliografie op richten Novick en Jackson 1974 (J). Er is in de loop der tijden nogal wat onderzoek gedaan naar de samenhang tussen allerlei variabelen en later studiesucces. We noemen Astin 1964, Hackman e.a. 1974. Harris 1940, Mehrabian 1969, en Warburton 1972.

Veel onderzoek op dit gebied gedaan is van twijfelachtige waarde door methodologische tekortkomingen (zie literatuur in categorie 11). Veel informatie ook kan niet in de praktijk gebruikt worden omdat ethische motieven belangrijker zijn (categorie E) of omdat dat in strijd zou zijn met onderwijskundige uitgangspunten (categorie K) of de onderwijsfilosofie die men hanteert (categorie N). Meer alledaags blijkt veel informatie te kostbaar te zijn om te kunnen verzamelen, onderzoeken en gebruiken: een telkens terugkerend thema in Cronbach en Gleser 1965 (c).

Problematisch is vaak hoe je verschillende informatieve gegevens moet combineren om tot de beste resultaten te komen: Wiggins en Kohn 1971, Kahneman en Tversky 1973, en Dawes en Corrigan 1974. We zitten hier in zeer netelige methodologische problemen (zie categorie H), die nog groter worden wanneer ook de te voorspellen uitkomsten op meerdere criteria gewaardeerd moeten worden (technieken hiervoor: zie categorie J).

Zijn alle drie soorten gegevens bekend (alternatieven en mogelijke uitkomsten; utiliteiten en waarschijnlijkheden van mogelijke uitkomsten; verband tussen informatie die men kan en wil gebruiken en de waarschijnlijkheid van mogelijke uitkomsten), dan kan in principe de beste beslissing uitgerekend worden. Daarvoor: Raiffa 1968 (2) en Lindley (2), en vooral ook weer Cronbach en Gleser 1965 (G).

Categorie A,

De student als beslissingnemer.

De student wordt in de onderwijs situatie vaak voor beslissingen geplaatst die zeer belangrijke gevolgen voor hem kunnen hebben: studiekeuze, keuze van specialisaties, studiestrategie, al dan niet voortzetten van de studie, beroepskeuze.

Hoewel hij bij het nemen van die beslissingen zich vaak kan laten begeleiden zal de verantwoordelijkheid voor de te nemen beslissingen toch uitsluitend bij hem of haar liggen. En hoe neem je eigenlijk dergelijke voor je verdere leven toch zo belangrijke beslissingen? In de Verenigde Staten is een programma gestart om tenminste een begin er mee te maken de high school leerling op dergelijke belangrijke beslissingen beter voor te bereiden, waarbij met name de nadruk valt op een zorgvuldige inventarisatie van beslissingsalternatieven en de mogelijke uitkomsten die de keuze voor ieder alternatief zou kunnen opleveren: Gelatt e.a. 1972. en Gelatt e.a. 1973. In veel landen is ook de keuze van college of universiteit belangrijk: Astin 1971 (O), Watts 1972.

Beschrijvend onderzoek naar wie welke keuzen maakt, en welke motieven daar een rol bij spelen, kan belangrijke gegevens opleveren die weer door studenten of begeleiders uitgebuit kunnen worden als extra informatie bij het nemen van hun beslissingen. Voorbeelden van dergelijk beschrijvend onderzoek: Abe en Holland 1965, Baird 1967, Saenger-Ceha 1970, Tinto 1975, en Wish en Hasazi 1973. Ook Wilbrink 1968 (j) is daar een voorbeeld van.

In overleg tussen studenten en andere belangengroepen over de inrichting van beoordeling en eventueel selectie, zal de student zich ervan bewust moeten zijn dat voorgestelde maatregelen vaak berusten op modellen die afgestemd zijn op de belangen van bijv. de institutionele beslissingnemer, belangen die althans voor een deel strijdig zullen zijn met de belangen van de individuele student. Ook daarom is een explicitering van de te volgen beslissingsstrategie voor de student wenselijk, om aan dergelijk

overleg meer inhoud te geven. Aan de uitwerking van een en ander, ook voor de groep van studenten, is nog vrijwel niets gedaan (maar zie Wilbrink april 1975 (J)).

Categorie B.

De docent als beslissingnemer.

De belangrijkste beslissing waar de individuele docent telkens voor geplaagd wordt is het bepalen van de zakslaag grens voor zijn tentamen. Het is de mening van de schrijver dat dit probleem alleen bevredigend aangepakt kan worden wanneer het langs de bekende besiskundige lijnen gaat, en wanneer de overige aspecten van de onderwijssituatie er expliciet bij betrokken worden. Onmiddellijke implicatie hiervan is dat de zakslaag beslissing niet meer een beslissing kan zijn van de individuele docent, maar in belangrijke mate door de (sub)faculteit in haar geheel, althans de te volgen procedure moet door de (sub)faculteit tot haar verantwoordelijkheid gerekend worden. De individuele docent dient ontlast te worden van beslissingen die een selectief karakter hebben, zodat hij zijn handen weer vrij krijgt om zijn beoordeling z6 uit te voeren dat ze weer een functie binnen zijn onderwijs, en ten dienste van zijn onderwijs kan krijgen (zie categorie L).

En binnen een meer onderwijskundige opzet kunnen zich dan nog problemen voordoen of iemand aan bepaalde criteria voor beheersing van de stof heeft voldaan of niet, maar daarvoor wordt verwezen naar de langzamerhand behoorlijk uitgebreide literatuur over Criterion Referenced Measurement, Mastery Learning, Individually Prescribed Learning, of welke anglicismen daar nog meer zijn: Block 1974, Swaminathan, Hambleton en Algina 1975 (in een volgende editie van deze bibliografie zal een poging gedaan worden iets vollediger te zijn t.a.v. deze literatuur).

Categorie C

De (sub)faculteit als beslissingnemer: selectie.

Van meet af aan heeft bij de ontwikkeling van psychologische tests voor gebruik bij selectie voorop gestaan dat daarbij de belangen van de selecterende instelling gediend zouden moeten worden. Op dit gebied bestaat dan ook een zeer uitgebreide literatuur, zowel theoretisch (psychometrische verhandelingen) als wat opgedane ervaringen in de praktijk betreft (waarvan in categorie 3 al het nodige opgenomen). Pas recent is het besef doorgebroken dat informatie die nuttig is voor het selecterende bedrijf (school, leger), voor anderen (de betrokken personen, politieke beleidsorganen) wel eens minder 'nuttig' zou kunnen, ja soms zelfs rechtstreeks conflicteren met ethische en normatieve principes (categorie E).

Ook voor de rechttoe rechtaan selectie problematiek is een besiskundige aanpak nog een zelden vertoonde toer (maar zie Van Naerssen, 1962, zoals afgedrukt in Cronbach en Gleser 1965), hoewel Cronbach en Gleser al in 1957 het theoretische voorbereidende werk leverden.

In deze categorie zijn dan ook werken samengebracht die een klassieke aanpak hanteren: Angoff, 1971; Cole en Hanson 1973; Thorndike 1949; naast werken die theoretisch of praktisch de besiskundige aanpak praktiseren: Cronbach en Gleser 1965; Van Naerssen 1970; Hills 1971.

Categorie D.

Beleidsinstantie (Uraad, kabinet en parlement) als beslissingnemer: plaatsing.

Voor beleidsinstanties is de overwegende vraag wie het best welke opleiding kan volgen, een probleem dat in Cronbach en Gleser 1965 (0) als het plaatsings probleem behandeld wordt. Natuurlijk zijn er ook binnen een opleiding vaak beslissingen te nemen over plaatsing van studenten in verschillende richtingen, 'stromen', e.d., en ook dergelijke problemen komen in deze categorie aan de orde: vooral Willingham 1974.

Er zijn nogal wat technische en methodologische kantjes aan het plaatsingsprobleem, zodat veel relevante literatuur ook te vinden is onder categorieën H en J (zoals ook geldt voor selectieproblematiek C). Een aantal auteurs hebben beschikbare technieken als het ware op het lijf van de plaatsingsproblematiek beschreven: Rulon e.a. 1967; en Birnbaum en Maxwell 1965.

Ondanks een overweldigende hoeveelheid mogelijk relevante technieken (discriminant analyse, cluster analytische benaderingen, klassificatieprocedures) is er met al dit materiaal toch eigenlijk weinig werk verzet aan het plaatsingsprobleem zoals zich dat in het onderwijs op ieder niveau weer voordoet.

Verwant aan de plaatsings problematiek is de vraag hoeveel mensen in bepaalde richtingen opgeleid zouden moeten worden, gezien bepaalde maatschappelijke doelstellingen. Het zou het kader van deze bibliografie te buiten gaan om ook dit onderwerp uitvoerig op te nemen, maar een enkel item (Folger, Astin, en Bayer 1970 bijv.) kunnen we toch niet nalaten op te nemen.

Wellicht ten overvloede moet tenslotte nog de opmerking gemaakt worden dat eenzelfde drempelprobleem (bijv. de toelating tot het wetenschappelijk onderwijs al dan niet onder een numerus fixus regeling) afzonderlijk geanalyseerd kan en moet worden van de in categorieën A t/m D gerepresenteerde belangengroeperingen uit. De resultaten van de afzonderlijke analyses zullen in het algemeen niet met elkaar in overeenstemming zijn, een gegeven dat tot op heden in discussies over selectie en beoordelingsprocedures te weinig of zelfs in het geheel geen aandacht heeft gekregen.

Categorie E.

Normatieve en ethische aspecten bij beoordeling en selectie.

Het gebruik van tests voor selectie doeleinden is altijd al onderwerp van heftige discussie geweest. In de loop der tijden heeft die discussie geleid tot steeds scherper geformuleerde normen waaraan tests moeten voldoen (Standards 1974), evenals tot strikter eisen waar het gaat om waarborgen van gelijke kansen bij sollicitaties voor verschillende bevolkingsgroepen: Anderson en Rogers 1970 (G). Veelal zijn de critici zich er niet van bewust dat vrijwel al hun kritiek niet zozeer betrekking heeft op een bepaalde vorm waarin de beoordeling plaats vindt (in casu d.m.v. tests), als wel op die beoordeling zelf. Daarmee is bedoeld dat veel kritiek niet ondervangen zou worden door maar geen tests meer te gebruiken, omdat ook dan beoordelingen nog gedaan moeten worden, en wel op de oude manier via mondelinge of schriftelijke, niet gestandaardiseerde procedures. De gevolgen daarvan zouden vandaag de dag als desastreus ervaren worden.

Ook in onderwijs situaties geldt dat beoordelingen, hoe ook tot stand gekomen, in principe moeten voldoen aan kwaliteitseisen zoals gesteld in de Standards 1974 (vooral het laatste hoofdstuk daarvan, over het gebruik van tests en toetsen). Er is geen essentieel verschil tussen een beoordeling die plaats vindt op basis van een kort gesprek, en een beoordeling die gebeurt via een testbatterij die zorgvuldig samengesteld en gestandaardiseerd is: het praktische verschil is dat bij het mondelinge gesprek het gegeven oordeel vrijwel een slag in de lucht is. Welnu, iedere beoordeling in het onderwijs, of ten behoeve van selectie vooraf, moet in principe aan een groot aantal kwaliteitseisen voldoen. In de praktijk is aan dergelijke eisen vaak moeilijk of niet tegemoet te komen, wat niet wil zeggen dat men ze dan maar aan de laars kan lappen: dat zal dan ook consequenties moeten hebben voor de wijze waarop op basis van dergelijke beoordelingen beslissingen genomen worden.

In deze categorie vallen publicaties die betrekking hebben op de kwalitatieve eisen die aan tests en toetsen, en aan beoordelingen in het algemeen, gesteld moeten worden: De Groot 1972; Mosier 1947; Standards 1974. En een aantal zaken die daarnaast nog van belang zijn, zoals de rol van 'toetsfabrieken' (Holmen en Docter 1972), de aard van de inrichting van het onderwijs (Johnson en Johnson 1974) en de richting waarin tests en test gebruik ontwikkeld moeten worden: Commission on tests 1970.

Categorie F.

Rechtspositie regelingen voor onderwijsdeelnemers.

Aan rechtspositie regelingen voor onderwijsdeelnemers is voorzover mij bekend nooit iets gedaan, althans niet iets dat noemenswaard is om in de bibliografie op te nemen. Er zal uiteraard wel een stuk jurisprudentie zijn over zaken die de student of de leerling in het onderwijs, of bij examens, betreffen. Maar uit die jurisprudentie is nog zelfs geen begin ontstaan van een werken aan een rechtspositie regeling (zoals dit voor arbeiders, en in uitgebreider zin daarna ook voor werknemers, in de loop van deze eeuw wel is gebeurd).

Er is wel voor gepleit, met name door De Groot in diverse publicaties, met name ook in zijn selectierapport 1972 (N), om een dergelijke regeling althans beginnen uit te werken.

Wederom een lege categorie dus. Hoewel niet helemaal: je zou kunnen zeggen dat de wetgeving in de Verenigde Staten op het gebied van gelijke kansen bij sollicitatie procedures (Anderson en Rogers 1970 (G)) een voorbeeld is van een klein, zij het ook zijdelings, begin. Ook de idee van contract onderwijs, de laatste jaren herhaaldelijk geopperd, zou als een aanzet in die richting gezien kunnen worden; contract onderwijs, waarbij de rechten en plichten van studenten en docenten wederzijds vastgelegd worden, met de implicatie dat op een en ander een beroep gedaan moet kunnen worden wanneer de nood aan de man komt.

Dat uitwerking van een en ander verre van eenvoudig zal zijn, daaraan hoeven we niet te twifelen. In het onderwijs zal het wel nooit zo ver kunnen komen dat loutere tijdbesteding 'voldoende' zal zijn: altijd weer zal er geselecteerd worden, (hoewel dergelijke selectie aardig versluierd kan worden) er zal nooit sprake van kunnen zijn dat iedereen 'recht heeft' op bepaalde studieprestaties. Dat de problematiek lastig, en netelig (in politieke zin) is kan geen uitvlucht zijn: tenslotte besteedt iedereen tegenwoordig een groot deel van zijn leven als deelnemer van meerdere vormen van onderwijs, waarbij grote belangen evenals een belangrijk stuk welzijn, op het spel staan.

Categorie G.

Gelijke kansen problematiek

Er zitten een aantal verschillende kanten aan deze gelijke kansen. Voor een deel is het een kwestie van waarborgen van gelijke kansen bij test en selectieprocedures: Anderson en Rogers 1970; Columbia Law Review 1968. Op dit gebied is weliswaar de laatste jaren in de Verenigde Staten een stuk wetgeving verricht, maar de problemen om in de praktijk aan de gegeven richtlijnen te voldoen zijn niet eenvoudig. Hier spelen methodologische problemen voorlopig nog de spelbreker: voorgestelde procedures zijn met elkaar strijdig (Einhorn en Baas 1971; Linn 1973); of aangetoond wordt dat pogingen tot oplossing leiden tot paradoxen (Jones 1973), waarop een uitweg lijkt te zijn om utiliteiten in de modellen te betrekken, dus door een besliskundige aanpak te kiezen waarbij de wenselijkheid van verschillende mogelijke uitkomsten expliciet meegewogen kan worden (Novick e.a., publicatie in press Journal of Educational Measurement).

Aan de andere kant gaat de gelijke kansen problematiek natuurlijk verder dan de vraag of bij toetsing en selectie verschillende groeperingen of personen wel fair behandeld worden. We komen dan meer op het gebied van de onderwijs sociologie: Boudon 1974; Byrne 1975; Husén 1975; Jencks e.a. 1972; Mosteller en Moynihan 1973; Sewell en Hauser 1975.

Pogingen om selectieprocedures wat gelijke kansen betreft aan de faire kant te houden, leiden tot resultaten die niet op dezelfde lijn liggen als een rechttoerechtaan vanuit de belangen van de selecterende instantie beredeneerde selectieprocedure. Dat dat tot de nodige misverstanden kan leiden mag blijken uit Wilbrink sept 1975 en de daarop in hetzelfde blad 2 weken later gevolgde

reacties.

Zoals uit het bovenstaande mag blijken, heeft de gelijke kansen problematiek zowel haar wortels in geldende maatschappelijke opvattingen, als in methodische zwakten van de wijze waarop selectieve beoordelingen verricht worden, vandaar dat ook de literatuur in categorie E hier van direct belang is. (evenals diverse items in categorie N)

Categorie H.

Methodologische problemen.

Bij toepassing van de in de volgende categorie vermeldde technieken doen zich nogal wat methodologische problemen voor. Dat wil zeggen: bepaalde technieken en methoden of procedures kunnen alleen toegepast worden wanneer aan diverse aan die technieken ten grondslag liggende assumpties redelijk voldaan is. Daarnaast ontstaan vaak problemen omdat de gevolgen van toepassing van bepaalde technieken niet evident zijn, of resultaten zich slechts met moeite laten interpreteren.

Methodologische problemen t.a.v. ethische en normatieve zaken zijn in categorie E opgenomen: Cattell 1964; Cattell 1967; Cronbach 1971; Cronbach en Gleser 1959; Duke 1968; Mosier 1947; en Stanley 1971.

Ook voor het overige laat de bibliografie zich onderverdelen, zoals bijvoorbeeld de groep artikelen die betrekking hebben op modellen voor de wijze waarop personen zich oordelen vormen, met daaraan gekoppeld de vraag of dergelijke modellen zich lenen voor vereenvoudiging en 'objectieve' toepassing: Birnbaum 1973; Castellan 1973; Dawes en Corrigan 1974; Einhorn 1970; Goldberg 1970; Hoffman 1960; Hoffman, Slovic, en Rorer 1968; Hofstee 1974; Pankoff 1968; Sawyer 1966; Wiggins en Kohen 1971; en een goed overzicht in Wiggins 1973 (J).

Een andere groep wordt gevormd door papers over problemen rond het 'ontdekken' van verschillende subgroepen (clusteranalytische technieken) en 'belang' van verschillen tussen groepen (discriminant analytische technieken), en klassificatie problemen: Anderson 1969; Campbell 1968; Cormack 1971; Finney 1965; Levy 1967.

Een belangrijke groep betreft psychometrische problemen, inclusief validerings problemen: Anastasi c.s. 1960; Dawes 1975; Humphreys 1968; James 1973; Lewy 1973; Linn 1968; Lirin en Schoenfeldt a.o. 1974; Luecke en McGinn 1975; Messick 1970; Richards c.s. 1965.

Dan is er een categorie papers die de toepassing van causale modellen in het onderzoek naar onderwijs effecten kritisch bespreekt: Herriot en Muse 1973; Linn en Werts 1969; Werts en Linn 1969.

En tenslotte een aantal papers die op onderwijskundige (onderzoek) problematiek gericht zijn: Ahmann en Glock 1959; Baird en Feister 1972; Goldman en Hewitt 1975; Nesselroade en Reese 1973; Wilbrink 1974.

Categorie J.

Beschikbare (mogelijk bruikbare) technieken en modellen.

In deze categorie zijn opgenomen zowel technieken die gehanteerd zouden kunnen worden bij praktische beoordelings-, selectie-, en andere procedures m.b.t. beslissingen van (en over) studenten, als technieken die geëigend kunnen zijn voor onderzoek naar de werking van bestaande procedures of mogelijke betere alternatieven.

Met name geldt voor vele multivariate technieken die hier genoemd worden dat zij vrijwel uitsluitend bruikbaar zijn bij explorerend en evaluatief onderzoek, maar, gezien de schaalgrootte van ons hoger onderwijs, nauwelijks in te zetten bij de uitvoering van de dagelijkse beslissingsprocedures: Darlington, Weinberg, & Walberg 1973; Dunham & Kravetz 1975; Eyman, Meyers, & Bendel 1973; Love & Stewart 1968; Nesselroade 1970.

In principe zou alle psychometrische literatuur hier genoemd moeten worden. maar om hanteerbaar te blijven noemen we naast de standaard werken (Cronbach & Gleser 1965 (C); Lord & Novick 1968; Novick & Jackson 1974) slechts: Cattell & Butcher 1968; Cooley 1971; Cole & Hanson 1973; Conger & Lipshitz 1973; Cronbach, Gleser, Rajaratnam, & Nanda 1972; Fricke 1972; Gross 1973; Houston 1968; Mosier 1951; Whitla 1968; Wiggins, J. S., 1973; Williams, Gab, & Lindem 1969; Wilmott & Fowles 1974.

Relevant voor de problematiek van studiekeuze, begeleiding, plaatsing e.d. is de volgende, nogal heterogene collectie van bronnen: Rulon a.o. 1967 (D) en andere literatuur ondergebracht in categorie D; Cattell 1969; McQuitty & Clark 1968; Novick & Jackson 1970; Tatsuoka & Cattell 1970; Tiedeman & Field 1968; Wilbrink 1968; Winter 1974.

Toegesplitst op de meer besiskundige aspecten zijn, naast de in categorieën 1, 2, en 3 al genoemde literatuur: Bruno 1972; Pitz 1974; Rapoport & Wallsten 1972; Wilbrink 1975.

Tenslotte een enkele bron relevant voor de bestudering van structuren en processen in het onderwijssysteem: Crano, Kenny & Campbell 1972; Duncan & Featherman 1973; Feldman 1970; Gimmel 1974; Hauser 1973; Plomp 1972; Wiggins, L. M., 1973; Wold 1970 a en b.

Categorie K.

Onderwijskundige uitgangspunten.

De wijze waarop beslissingen over personen worden genomen kan in onderwijs situaties niet los gezien worden van onderwijskundige principes die mede aan de structurering van die onderwijs situaties ten grondslag liggen. Met deze uitspraak kan men het makkelijk eens zijn. De moeilijkheid er mee is echter dat het niet duidelijk is wat hier met 'onderwijskundig' bedoeld wordt. Ook de scheiding met categorie N (onderwijsfilosofie) is niet scherp aan te geven. In een latere editie van deze bibliografie zal een en ander wel uitgewerkt en meer inhoud gegeven kunnen worden, voor het moment moeten we maar genoegen nemen met de onder deze categorie onder gebrachte items als een voorlopige aanduiding van bedoelde inhoud van deze categorie.

In afzonderlijke categorieën worden de feedbackfunctie van beoordelingen, en onderwijsfilosofische uitgangspunten aan de orde gesteld (categorieën L en N respectievelijk.).

Categorie L.

Feedback functie van beoordelingen (incl. motivationele aspecten).

In ieder beoordelingsstelsel zullen beoordelingen invloed uitoefenen op het gedrag van docenten en studenten, afhankelijk van de aan de cijfers gegeven interpretaties. (Een beoordeling hoeft natuurlijk niet altijd een cijfer te zijn, voor het gemak spreken we echter gewoon van cijfers daar waar we eigenlijk over 'beoordelingen' zouden moeten spreken). Dergelijke beïnvloeding, terugkoppeling, of motivatie kan op nogal wat verschillende manieren plaats vinden, zoals ook hier de literatuurlijst weer demonstreert. Bij het opzetten van beoordelings- en selectieprocedures is het goed om een overzicht van de gelijke terugkoppelingseffecten te hebben, om mogelijk onbedoelde effecten van voorgestelde examenprocedures te kunnen anticiperen, onderzoeken, of zo mogelijk door bijkomende maatregelen weg te kunnen nemen of juist versterken.

Een eerst groepering in de literatuur is die rond de didactisch meer van belang zijnde terugkoppelingseffecten, zoals kennis van resultaten, bekrachtiging van goede antwoorden e.d.: Annett 1969; Locke 1968; Mager & Pipe 1970.

Bijna logisch daarop aansluitend de literatuur m.b.t. mastery learning en meer geïndividualiseerde onderwijsvormen, waar de student op basis van behaalde studieresultaten op kleine studieonderdelen, door het onderwijs gestuurd (geloodsd) wordt: Block 1971; Block 1974; Hambleton 1974; Johnson & Abrahamson 1968; Johnston 1975; Lloyd 1971; en Sherman 1974.

De wijze waarop verschillende beoordelingen gecombineerd worden tot eindoordeel is hier ook van belang: Biggs & Braun 1972; Van Naerssen 1972, 1975)J).

Literatuur direct in de lijn van motivatie onderzoek (naast de in categorie Q, persoonlijke verschillen in het onderwijs, genoemde): Domino 1970; en Heckhausen 1975.

Tenslotte, aan al dat getest en getoets in het onderwijs zitten ook wel bedenkelijke kanten, wat averechtse effecten op studenten en onderwijsinstelling betreft: Kirkland 1971 (voor ethische en normatieve kwesties bij toetsing in het onderwijs, zie categorie E).

Categorie N.

Onderwijsfilosofie (globale doelstellingen, maatschappelijke opvattingen).

Het kiezen van utiliteiten (zie categorie 2) is voor een niet onaanzienlijk deel een kwestie van kwantificeren van je onderwijsfilosofie, je maatschappelijke opvattingen (opvattingen over de plaats van het onderwijs in de samenleving) e.d. Onderwijsfilosofische opvattingen, om ze zo maar te blijven noemen, blijken dus geenszins in de in dit rapport aangegeven conceptie over beoordeling en selectie in het onderwijs, een abstract of wereldvreemd karakter te hebben.

Dat wil niet zeggen dat vrijwel alle publicaties van onderwijsfilosofisch karakter, tot en met de contouren nota van Van Kemenade, vaag blijven in die zin dat allerlei mooie opvattingen niet geconcretiseerd worden (op de wijze zoals met de in categorie 1, 2, en 3 gegeven technieken mogelijk zou zijn). Een aardig voorbeeld van een rapport waarin uitgesproken opvattingen over de plaats van het onderwijs in de samenleving, en voorstellen voor de wijze waarop selectie uitgevoerd moet worden, nogal los van elkaar staan, is het rapport van De Groot 1972.

Eigenlijk is deze categorie niet meer dan een verzamelbak voor al die bronnen die niet in categorieën 2, E (normatieve en ethische toestanden), F (rechtspositie regelingen) of G (gelijke kansen problematiek) geplaatst kunnen worden. Erg veel houden we dan inderdaad niet over. Daar komt bij dat ik erg huiverig sta t.o.v. de meeste literatuur die verschenen is onder het kopje 'onderwijsfilosofie' o.i.d. (en dat is nogal wat).

Een voorbeeld van een publicatie die dan eigenlijk alleen nog maar in deze categorie te plaatsen valt, is Munday & Davis 1974, die een onderzoek rapporteren over relaties tussen succes in school en in later beroep, en concluderen:

"this suggests that there are many kinds of talents related to later success which might be identified and nurtured by educational institutions. As we evaluate college outcomes in terms of postcollege student behaviors, we may have to reappraise the central role previously assigned academic talent."

Categorie O.

Beoordelingsgewoonten, selectiegewoonten (beschrijvende gegevens).

Beschrijving van bestaande selectieprocedures voor hoger onderwijs instellingen in diverse landen: Astin 1971; Bowles 1963, 1965; Hearnden 1963; Peet & De Munnik 1972; Powell 1973; Willingham 1973; Willingham 1974. Het is altijd wel aardig om te zien hoe slecht er ergens anders op de wereld geselecteerd wordt, maar het is meestal niet erg duidelijk hoe deze informatie

geïnterpreteerd moet worden.

Waar altijd erg veel belangstelling voor bestaat, hoewel ook hier een geweldig interpretatie probleem (m.b.t. vergelijking met de eigen onderwijs situatie) ligt, zijn gegevens over verband tussen voorafgaande prestaties en latere prestaties, met name rond de overgang naar wetenschappelijk (of college) onderwijs: Astin 1971; Choppin 1973; Everwijn, Koopmans & Vroon 1974; Fincher 1974; Hoyt (voorspelling succes in beroep) 1965; Kopfman 1969; Richards et al 1974; om maar de meest voor de hand liggende bronnen te noemen. De reeks kan ad libitum uitgebreid worden.

Cijfergeverij in het wetenschappelijk onderwijs blijkt ook in het buitenland al dezelfde merkwaardige trekjes te tonen die we ook in onze eigen onderwijspraktijk vandaag de dag nog steeds waar kunnen nemen: Baird & Feister 1972; Eggleston & Kerr 1969; Goldman et al 1974; De Groot 1966 (klassiek !); Ingenkamp 1971.

Ongelukkige cijferpraktijken kunnen bijdragen tot het drop out probleem: Astin & Panos 1969; Miller 1970; Parkyn 1967. Echter, het bevredigen van de nieuwsgierigheid naar hoe er op het ogenblik (in diverse landen) geselecteerd en beoordeeld wordt, levert op zich nog geen enkele bijdrage tot betere praktijken in ons eigen wetenschappelijk onderwijs. Het is dan ook met enige tegenzin dat deze literatuur in de bibliografie is opgenomen. Hier ligt niet de ingang tot een beter en prettiger onderwijs vormgeving. Wie toch van deze literatuur kennis wil nemen, zij gewaarschuwd dat bij de interpretatie van dit soort gegevens zich forse methodologische problemen voordoen (zie de in categorie H genoemde literatuur).

Categorie P.

Evaluatie van onderwijs, en curriculum beschrijving.

Het is niet alleen volledigheidshalve dat deze categorie in de lijst is opgenomen. Immers, het heeft weinig zin om onderwijsveranderingen door te voeren die niet in principe ook op hun effect te toetsen zijn. Voor zover het gaat om het evalueren van gebruikte selectie en toetsings instrumenten, is de wijze waarop zij beoordeeld moeten worden (de criteria waaraan zij moeten voldoen) vastgelegd in documenten als de Standards (1974, categorie E), en daarvoor verwijzen we naar de gehele categorie E.

Omdat echter beoordeling, plaatsing, en selectie geen activiteiten zijn die kunnen plaats vinden los van het onderwijs en de inhoud van het onderwijs, moeten zij ook mede in hun relatie tot het onderwijs gevalueerd worden. Zeker daar waar bepaalde beoordelings en selectie methoden worden voorgesteld om juist de effectiviteit of het rendement van dat onderwijs te kunnen verbeteren.

Voor een evaluatie van het onderwijs systeem is een beschrijving van het functioneren van dat onderwijs, en een beschrijving van de inhoud van dat onderwijs, van groot belang. Jammer genoeg zijn òuist op dit gebied geen of weinig technieken ontwikkeld, of zijn elders ontwikkelde technieken op hun bruikbaarheid voor dit doel onvoldoende onderzocht. Een paar aanzetten zijn in de literatuur echter wel te vinden: Posner 1974; Short 1973; Stones 1972.

Belangrijk voor iedere evaluatie poging is dat duidelijk is tussen welke beleidsalternatieven op basis van de te verzamelen en analyseren informatie een keuze gedaan moet worden: Astin & Panos 1971; Stufflebeam. et al 1971.

Het euvel van veel evaluatie projecten is dat men ondoordacht, hals over kop, aan het evalueren slaat. Het zou goed zijn in een paar ogenblikken van reflectie in ieder geval de volgende bronnen te raadplegen: Baker 1974; Costin 1971; Cronbach 1963; Johnson, Rhodes & Rumery 1975a en b; Messick 1970; Sanders 1973.

Het spreekt vanzelf dat voor evaluatie een aantal technieken uit categorie J onontbeerlijk zijn, en dat een aantal methodologische valkuilen (categorie H) het pad van de onderzoeker pokdalig maken.

Categorie Q.

(de betekenis van) Persoonlijke verschillen in het onderwijs.

Over de betekenis van persoonlijke verschillen t.a.v. de onderwijsvormgeving, beoordeling, selectie etc. bestaan doorgaans zeer verwarde ideeën. Zie als uiting daarvan de discussie over de mate van erfelijk bepaald zijn van verschillen in intelligentie (De Mulder 1974 (N)), of over de mogelijkheden om binnen een al ver geselecteerde groep kandidaten voor het w.o. nog verder te selecteren t.b.v. een numerus fixus. Veel van deze misverstanden zijn terug te voeren op methodologische misvattingen (zie categorie H).

In dezelfde lijn liggen al te enthousiaste opvattingen over de wenselijkheid en mogelijkheden van geïndividualiseerd onderwijs, mastery learning e.d., zie bijv. de kritiek van Holland 1971.

Ook ten aanzien van het onderwerp van deze categorie moet weer geconstateerd worden dat de literatuur uit alle hoeken en gaten nog bijeen gesleept, en georganiseerd moet worden: hoewel er op onderdelen erg veel en goed werk wordt gedaan door psychologen en onderwijsonderzoekers, heeft nog geen auteur zich gewaagd aan het geven van een behoorlijk overzicht.

In deze categorie horen dan thuis: kwesties rond motivatie (Atkinson & Raynor 1974; Cattell & Child 1975; Maehr & Sjogren 1971), pogingen om onderwijs aan te passen bij verschillen tussen studenten (onderzoek naar Aptitude Treatment Interactions: Cronbach & Snow in press; Salomon 1972 (K); Berliner & Cahen 1973) waarbij vooral ook verschillen in cognitief functioneren van belang zijn (Gagné 1967; Sperry 1972).

Gelijke kansen problematiek (categorie G) moet hier ook genoemd worden, omdat het er hier om gaat in hoeverre verschillende

personen of groepen, die zich onderscheiden door bepaalde persoonlijke of groeps kenmerken, al of niet terecht verschillend of juist gelijk behandeld worden. Voorbeeld van onderzoek dat beide categoriën (G en Q) overlapt, is het werk van Jensen, dat niet verder genoemd is omdat het geen betrekking heeft op hoger onderwijs.

[2002: de hele literatuurlijst opgenomen in literatuurbestand Ex]

categorie 2.

- Alkin, M. C. Evaluating the costeffectiveness of instructional programs. In M. C. Wittrock & D. E. Wiley (Eds), *The evaluation of instruction: issues and problems*. London: Holt, Rinehart, & Winston, 1970.
- Davis, J. H. Group decision and social interaction: a theory of social decision schemes. *Psychological Review*, 1973, 80, 97125.
- Lindley, D. *Making decisions*. London: Wiley, 1971.
- Marco, G. L. A comparison of selected school effectiveness measures based on longitudinal data. *Journal of Educational Measurement*, 1974, 11, 225234.
- Naerssen, R. F. van A scale for the measurement of subjective probability. *Acta Psychologica*, 1962, 26, 159166.
- Raiffa, H. *Decision analysis. Introductory lectures on choices under uncertainty*. London: AddisonWesley, 1968.
- Vlek, Ch. A. J. The fair betting game as an admissible procedure for assessment of subjective probabilities. *British Journal of Mathematical and Statistical Psychology*, 1973, 26, 1830.
- Wilbrink, B. *Het nemen van beslissingen. (interne nota)* Amsterdam: C.O.W.O. spui 21, 1975.
- Witmer, D.R. Economic benefits of college education. *Review of Educational Research*, 1970, 40, 511524.

categorie 3.

- Astin, A. W. Personal and environmental factors associated with college dropouts among high aptitude students. *Journal of Educational Psychology*, 1964, 55, 219227.
- Dawes, R. H. & Corrigan, B. Linear models in decision making. *Psychological Bulletin*, 1974, 81, 94106.
- Hackman, J. R., Wiggins, N., & Bass, A. R. Prediction of longterm success in doctoral work in psychology. *Educational and Psychological Measurement*, 1970, 30, 365374.
- Harris, D. Factors affecting college grades: a review of the literature, 1930 1937. *Psychological Bulletin*, 1940, 37, 125166.
- Kahneman, D., & Tversky, A. On the psychology of prediction. *Psychological Review*, 1973, 80, 237251.
- Mehrabian, A. Undergraduate ability factors in relationship to graduate performance. *Educational and Psychological Measurement*, 1969, 29, 409419.
- Schrader, W. B. The predictive validity of College Board admissions tests. In Angoff, W. H. (Ed). *The College Board admissions testing program*. New York: College Entrance Examination Board, 1971.
- Thacker, A. J., & Williams, R. E. The relationship of the graduate record examination to grade point average and success in graduate school. *Educational and Psychological Measurement*, 1974, 34, 939944.
- Warburton, F. W. The relationship between personality factors and scholastic attainment. In Dreger, R. M. (Ed), *Multivariate personality research*. Baton Rouge: Claitor's Publishing Division, 1972.
- Wiggins, N., & Kohen, E. S. Man versus model of man revisited: the forecasting of graduate school success. *Journal of Personality and Social Psychology*, 1971, 19, 100106.

categorie A.

- Abe, C., & Holland, J. L. A description of college freshmen. I. Students with different choices of major field. II. Students with different vocational choices. (Research Report 3 & 4). Iowa City: American College Testing Program Publications Office, 1965.
- Baird, L. L. The undecided student How different is he? (ACT Research Report 22). Iowa City: American College Testing Program Publications Office, 1967.
- Bossel, H. College student and dropout problem: a qualitative dynamic simulation. *Instructional Science*, 1974, 3, 2350.
- Cooley, W. W., & Lohnes, P. R. Predicting development of young adults. (Interim report 5 to the U. S. Office of Education, Cooperative Research Project no. 3051.) Palo Alto: Project TALENT Office, American Institutes for Research and University of Pittsburgh, 1968.
- Flanagan, J. C., Shaycoft, M. F., Richards, J. M., Jr., & Claudy, J. G. Five years after high school. (Final report to the U. S. Office of Education.) Palo Alto: Project TALENT Office, American Institutes for Research and University of Pittsburgh, 1971.
- Gelatt, H. B., Varenhorst, B., Carey, R., & Miller, G. P. *Decisions and outcomes (+ idem, A leader's guide)*. New York: College Entrance Examination Board, 1973.
- Gelatt, H. B., Varenhorst, B., & Carey, R. *Deciding (+ Deciding: A leader's guide)*. New York: College Entrance Examination Board, 1972.
- Jepsen, D. A., & Dilley, J. S. Vocational decisionmaking models: a review and comparative analysis. *Review of Educational Research*, 1974, 44, 331349.
- Marks, E. Cognitive and incentive factors involved in within university transfer. *Journal of Educational Psychology*, 1970, 61, 19.
- Moore, W., Jr., Against the odds. The highrisk student in the community college. San Francisco: Jossey-Bass, 1970.
- Saenger-Ceha, M. M. Th. *Psychological and social factors in student dropout. (proefschrift)* Amsterdam: Swets and Zeitlinger, 1970.
- Tiedeman, D. V., & Field, F. L. Measurement for guidance. In Whitla, D. K. (Ed), *Handbook of measurement and assessment in behavioral sciences*. London: AddisonWesley, 1968.
- Tinto, V. Dropout from higher education: a theoretical synthesis of current research. *Review of Educational Research*, 1975, 45, 89125.
- Watts, A. G. *Diversity and choice in higher education*. London: Routledge & Kegan Paul, 1972.
- Wish, P. A.51 & Hasazi, J. E. Motivational determinants of curricular choice in college males. *Journal of Counseling*

Psychology, 1973, 20, 127131.

categorie B.

- Block, J. (Ed.). Schools, society and mastery learning. London: Holt, Rinehart, & Winston, 1974.
- Millman, J. Criterion referenced measurement. In Popham, W. J. (Ed.), Evaluation in education; current applications. Berkeley: McCutchan, 1974.
- Swaminathan, H., Hambleton, R. K., & Algina, J. A bayesian decision theoretic procedure for use with criterionreferenced tests. Journal of Educational Measurement, 1975, 12, 8798.
- Wijnen, W. H. F. W. Onder of boven de maat; een methode voor het bepalen van de grens voldoende onvoldoende bij studietoetsen. Amsterdam: Swets & Zeitlinger, 1972.
- Wijnen, W. H. F. W. Het diagnostisch gebruik van studietoetsen. In Van Woerden, W. W., Chang, T. M., & Van Geuns-Wiegman, L. J. M., (Red.) Onderwijs in de maak. Utrecht: Het spectrum, 1973.

categorie C.

- Angoff, W. H. (Ed.). The College Board admissions testing program; a technical report on research and development activities relating to the Scholastic Aptitude Test and Achievement Tests. New York: College Entrance Examination Board, 1971.
- Cole, N. S., Hanson, G. R. Assessing students on the way to college. (2 volumes). Iowa City, Iowa: The American College Testing Program, 1973.
- Cronbach, L. J., & Gleser, G. C. Psychological tests and personnel decisions. London: University of Illinois Press, 1965. (Originally published, 1957)
- Fuess, C. M. The College Board; its first fifty years. New York: College Entrance Examination Board, 1967.
- Hills, J. R. Use of measurement in selection and placement. In Thorndike, R. L. (Ed.), Educational Measurement. Washington, D. C. American Council on Education, 1971
- Lord, F. M. Cutting scores and errors of measurement. Psychometrika, 1962, 27, 1930.
- Naerssen, R. F. van, Het nut van selectie, enkele voorbeelden. In Drenth, P. J. D., Willems, P. J., & de Wolff, Ch. J. Bedrijfspsychologie; onderzoek en evaluatie. Kluwer, Van Loghum Slaterus, 1970.
- Thorndike, R. L. Personnel selection; test and measurement techniques. New York: Wiley, 1949.

categorie D.

- Birnbaum, A., & Maxwell, A. E. Classification procedures based on Bayes's formula. In Cronbach, L.J., & Gleser, G. C. Psychological tests and personnel decisions. London: University of Illinois Press, 1965.
- Cole, N. S. Differential validity in the ACT tests. (ACT Research Report 30). Iowa City, Iowa: American College Testing Program Publications Office, 1969.
- Folger, J. K., Astin, H. S., & Bayer, A. E. Human resources and higher education. New York: Russell Sage Foundation, 1970.
- Loeb, J., & Bowers, J. Programs of study as a basis for selection, placement and guidance of college students. Journal of Educational Measurement. 1973, 10, 131139.
- Rulon, P. J., Tiedeman, D. V., Tatsuoaka, M. M., & Langmuir, C. R. Multivariate statistics for personnel classification. New York: Wiley, 1967.
- Willingham, W. W. College placement and exemption. New York: College Entrance Examination Board, 1974.

categorie E.

- Stanley, J. C. Reliability. In Thorndike, R. L. (Ed.), Educational measurement. Washington, D. C.: American Council on Education, 1971.
- Cattell, R. B. Validity and reliability, a proposed more basic set of concepts. In Mehrens, W. A., & Ebel, R. L. (Eds.), Principles of educational and psychological measurement; a book of selected readings. Chicago: Rand McNally, 1967. (Reprinted from Journal of Educational Psychology, 1964, 55, 122.)
- Cronbach, L. J. Test validation. In Thorndike, R. L. (Ed.), Educational Measurement. Washington, D. C.: American Council on Education, 1971.
- Cronbach, L. J., & Gleser, G. C. Interpretation of reliability and validity coefficients. In Cronbach, L. J., & Gleser, G. C. Psychological tests and personnel decisions. Urbana: University of Illinois Press, 1965. (Reprinted from Journal of Educational Psychology, 1959, 50, 230 237.)
- Duke, J. D. False negatives and the latent validity of aptitude tests. Educational and Psychological Measurement, 1968, 28, 389393.
- Groot, A. D. de, Acceptabiliteit van selectiemethoden. In Selectie voor en in het hoger onderwijs: een probleemanalyse. 's-Gravenhage; staatsuitgeverij, 1972.
- Holmen, M. G., & Docter, R. F. Educational and psychological testing: a study of the industry and its practices. New York: Russell Sage Foundation, 1972.
- Johnson, D. W., & Johnson, R. T. Instructional goal structure: cooperative, competitive or individualistic. Review of Educational Research, 1974 4431 213240.
- Messick, S. The standard problem. Meaning and values in measurement and evaluation. American Psychologist, 1975, 30, 955966.
- Mosier, Ch. I. A critical examination of the concepts of face validity. In Mehrens, W. A., & Ebel, R. L. (Eds.), Principles of educational and psychological measurement; a book of readings. Chicago: Rand McNally, 1967. (Reprinted from Educational and Psychological Measurement, 1947, 7, 191205.)
- Righting the balance, & Briefs. (separate volumes). New York: College Entrance Examination Board, 1970.
- Standards for educational and psychological tests. Washington, D. C., American Psychological Association, 1974.

categorie G

- Anderson, B. R., & Rogers, P. R. (Eds.), Personnel testing and equal employment opportunity. Washington, D. C. : Equal

Employment Opportunity Commission, 1970.

Boudon, R. Education, Opportunity and social inequality. Changing prospects in western society. London: Wiley, 1974.

Byrne, D., Williamson, B., & Fletcher, B. The poverty of education. A study in the politics of opportunity. London: Martin Robertson, 1975.

Einhorn, H. J., & Bass, A. R. Methodological considerations relevant to discriminations in employment testing. *Psychological Bulletin*, 1971, 75, 261-269.

Gross, A. L., Faggan, J., & McCarthy, K. The differential predictability of the college performance of males and females. *Educational and Psychological Measurement*, 1974, 34, 363-365.

Gross, A. L., & Su, W. Defining a 'fair' or 'unbiased' selection model: a question of utilities. *Journal of Applied Psychology*, 1975, 60, 345-351.

Hewitt, B. N., & Goldman, R. D. Occam's razor slices through the myth that college women overachieve. *Journal of Educational Psychology*, 1975, 67, 325-330.

Husén, T. Social influences on educational attainment. Research perspectives on educational equality. Paris: O.E.C.D., Centre for Educational Research and Innovation, 1975.

Jencks, C. (Ed.), *Inequality. A reassessment of the effect of family and schooling in America*. New York: Basic Books, 1972.

Jones, M. B. Moderated regression and equal opportunity. *Educational and Psychological Measurement*, 1973, 33, 591-602.

Kirk, D. L., & Yudof, M. G. *Educational policy and the law*. Berkeley: McCutchan, 1974.

Lakoff, S. A. *Equality in political philosophy*. Boston: Beacon Press, 1964.

Legal implications of the use of standardized ability tests in employment and education. *Columbia Law Review*, 1968, 68, 690-744.

Linn, R. L. Fair test use in selection. *Review of Educational Research*, 1973, 43, 139-161.

Mosteller, F., & Moynihan, D. P. (Eds.), *On equality of educational opportunity*. New York: Vintage Books, 1972.

Rossi, A. S., & Calderwood, A. (Eds.). *Academic women on the move*. New York: Russell Sage Foundation, 1970.

Sewell, W. H., & Hauser, R. M. *Education, occupation, and earnings*. London: Academic Press, 1975.

Silver, H. (Ed.). *Equal opportunity in education. A reader in social class and educational opportunity*. London: Methuen, 1973.

Wilbrink, B. Selectie voor numerus fixus is irrationeel. *NRC/Handelblad, Cultureel Supplement*, 27 sept. 1974.

Willingham, W. W. *Free access higher education*. New York: College Entrance Examination Board, 1970.

categorie H.

Ahmann, J. S., & Glock, M. D. An evaluation of the effectiveness of a freshman mathematics course. *Journal of Educational Psychology*, 1959, 50, 41-45.

Anastasi, A., Meade, M. J., & Schneiders, A. A. The validation of a biographical inventory as a predictor of college success. New York: College Entrance Examination Board, 1960.

Anderson, J. A. Constrained discrimination between k populations. *Journal of the Royal Statistical Society (Series B)* 1969, 31, 123-139.

Baird, L., & Feister, W. J. Grading standards: the relation of changes in average student ability to the average grades awarded. *American Educational Research Journal*, 1972, 9, 431-441.

Birnbaum, M. H. Reply to the devil's advocates: don't confound model testing and measurement. *Psychological Bulletin*, 1974, 81, 854-859. (gives references to preceding discussion papers).

Campbell, J. P. Testspace prediction of major field of study. *Educational and Psychological Measurement*, 1968, 28, 887-889.

Castellan, N. J. Comments on the 'lens model' equation and the analysis of multiplecue judgment tasks. *Psychometrika*, 1973, 38, 87-100.

Cormack, R. M. A review of classification. *Journal of the Royal Statistical Society (Series B)*, 1971, 33, 321-353.

Dawes, R. H., & Corrigan, B. Linear models in decision making. *Psychological Bulletin*, 1974, 81, 94-106.

Dawes, R. H. Graduate admission variables and future success: we cannot tell whether the standard selection measures used by graduate schools are valid. *Science*, 1975, 187, 721-723.

Einhorn, H. J. The use of nonlinear, noncompensatory models in decision making. *Psychological Bulletin*, 1970, 73, 221-230.

Finney, D. J. The statistical evaluation of educational allocation and selection. In Cronbach, L. J., & Gleser, G. C. *Psychological tests and personnel decisions*. Urbana: University of Illinois Press" 1965 (Reprinted from *Journal of the Royal Statistical Society (Series A)*, 1962, 125, 525-564.)

Goldberg, L. R. Man versus model of man: a rationale, plus some evidence, for a method of improving on clinical inferences. *Psychological Bulletin*, 1970, 73, 422-432.

Goldman, R. D., & Hewitt, B. N. Adaptation level as an explanation for differential standards in college grading. *Journal of Educational Measurement*, 1975, 123, 149-162.

Herriot, R. E., & Muse, D. H. Methodological issues in the study of school effects. In Kerlinger, F. N. (Ed.). *Review of research in education*. (volume I.) Itasca, Illinois: Peacock, 1973.

Hoffman, P. J. The paradigmatic representation of clinical judgment. *Psychological Bulletin*, 1960, 57, 116-131.

Hoffman, P. J., Slovic, P., & Rorer, L. G. An analysis of variance model for the assessment of configural cue utilization in clinical judgment. *Psychological Bulletin*, 1968, 69, 338-349.

Hofstee, W. K. B. *Psychologische uitspraken over personen; beoordeling, voorspelling, advies, test*. Deventer: Van Loghum Slaterus, 1974.

Humphreys, L. G. The fleeting nature of the prediction of academic success. *Journal of Educational Psychology*, 1968, 59, 375-380.

James, L. R. Criterion models and construct validity for criteria. *Psychological Bulletin*, 1973, 90, 75-83.

Levy, Ph. Substantive significance of significant differences between two groups. *Psychological Bulletin*, 1967, 67, 37-40.

Lewy, A. Discrimination among individuals versus discrimination among groups. *Journal of Educational Measurement*, 1973, 10, 19-24.

Linn, R. K. Range restriction problems in the use of selfselected groups for test validation. *Psychological Bulletin*, 1968, 69, 69-73.

Linn, R. K., & Werts, Ch. E. Assumptions in making causal inferences from part correlations, partial correlations, and partial regression coefficients. *Psychological Bulletin*, 1969, 72, 307-310.

Lissitz, R. W., & Schoenfeldt, L. F. Moderator subgroups for the estimation of educational performance: a comparison of prediction models. *American Educational Research Journal*, 1974, 11, 63-75. (see also immediately following discussion

papers, *American Educational Research Journal*, 1974, 11, 7792.

Luecke, D. F.J., & McGinn, N. F. Regression analyses and education production functions: can they be trusted? *Harvard Educational Review*, 1975, 45, 325350.

Messick, S. The criterion problem in the evaluation of instruction: assessing possible, not just intended, outcomes. In Wittrock, M. C., & Wiley, D. E. (Eds.). *The evaluation of instruction: issues and problems*. London: Holt, Rinehart, & Winston, 1970.

Nesselroade, J. R., & Reese, H. W. (Eds.). *Lifespan developmental psychology. Methodological issues*. London: Academic Press, 1973.

Pankoff, L. D. Bayesian synthesis of clinical and statistical prediction. *Psychological Bulletin*, 1968, 70, 762773.

Richards, J. M., Jr, Taylor, C. W., Price, P. B., & Jacobsen, T. An investigation of the criterion problem for one group of medical specialists. In Ronan, W. W. & Prien, E. P. (Eds.), *Perspectives on the measurement of human performance*. New York: Appleton-Century-Crofts, 1971. (Reprinted from *Journal of Applied Psychology*, 1965, 49, 7990.

Sawyer, J. Measurement and prediction, clinical and statistical. *Psychological Bulletin*, 1966, 66, 178200.

Stroud, T. W. F. Comparing regressions when measurement error variances are known. *Psychometrika*, 1974, 39, 5368.

Werts, Ch. E., & Linn, R. L. Analyzing school effects: how to use the same data to support different hypotheses. *American Educational Research Journal*, 1969, 6, 439447.

Wiggins, N., & Kohen, E. S. Man versus model of man revisited: the forecasting of graduate school success. *Journal of Personality and Social Psychology*, 1971, 19, 100106.

wilbrink, B. Onvoldoende: dus overdoen? Discussiestuk. Amsterdam: C.O.W.O. Spui 21, december 1974.

categorie J.

Bruno, J. E. A methodology for the evaluation of instruction or performance contracts which incorporates school district utilities and goals. *American Educational Research Journal*, 1972, 9, 175195.

Cattell, R. B. The profile similarity coefficient, r_p , in vocational guidance and diagnostic classification. *British Journal of Educational Psychology*, 1969, 39, 131142.

Cattell, R. B., & Butcher, H. J. *The prediction of achievement and creativity*. New York: BobbsMerrill, 1968.

Cooley, W. W. Techniques for considering multiple measurements. In Thorndike, R. L. (Ed.), *Educational measurement*. Washington, D. C. American Council on Education, 1971.

Cole, N. S., & Hanson, G. R. Assessing students on the way to college. Technical report from the ACT assessment program, & College student profiles: norms for the ACT assessment. Iowa City, Iowa: American College Testing Program Publications office, 1973.

Conger, A. J., & Lipshitz, R. Measures of reliability for profiles and test batteries. *Psychometrika*, 1973, 38, 411427.

Crano, W. D., Kenny, D. A., & Campbell, D. T. Does intelligence cause achievement? A crosslagged panel analysis. *Journal of Educational Psychology*, 1972, 63, 258-275.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. London: Wiley, 1972.

Darlington, R. B., Weinberg, L., & Walberg, H. J. Canonical variate analysis and related techniques. *Review of Educational Research*, 1973, 43, 433 454.

Duncan, O. D., & Featherman, D. L. Psychological and cultural factors in the process of occupational achievement. In Goldberger, A. S., & Duncan, O. D. *Structural equation models in the social sciences*. London: Seminar Press, 1973.

Dunham, R. B., & Kravetz, D. J. Canonical correlation analysis in a predictive system. *Journal of Experimental Education*, 1975, 43 no 4, 3542.

Eyman, R. K., Meyers, C., & Bendel, R. New methods for test selection and reliability assessment using stepwise multiple regression and jackknifing. *Educational and Psychological Measurement*, 1973, 33, 883894.

Feldman, K. A. Some methods for assessing college impacts. *Sociology of Education*, 1970, 44, 133150.

Fricke, R. *Ueber Messmodelle in der Schulleistungsdiagnostik*. Düsseldorf: Pädagogischer Verlag Schwann, 1972.

Gimmel, D. R. A path model for course achievement. *Jou of Experimental Education*, 1974, 43 no 4, 17.

Gross, A. L. Prediction in future samples studied in terms of the gain from selection. *Psychometrika*, 1973, 38, 151171.

Hauser, R. M. Disaggregating a socialpsychological model of educational attainment. In Goldberger, A. S., & Duncan, O. D. (Eds.), *Structural equation models in the social sciences*. London: Seminar Press, 1973.

Houston, Samuel R. Generating a projected criterion of graduate school success via normative judgment analysis. *Journal of Experimental Education*, 1968, 38 no 2, 5358.

Lord, F. M., & Novick, M. R. *Statistical theories of mental test scores*. London: AddisonWesley, 1968.

Love, W. A., Jr., Stewart, D. K. Interpreting canonical correlations: Theory and practice. (Interim report 7 to the U. S. Office of Education, Cooperative Research Project no. 3051.) Palo Alto: Project TALENT Office, American Institutes for Research and University of Pittsburgh, 1968.

McQuitty, L. L., & Clark, J. A. Hierarchical classification by reciprocal pairs of course selections in psychology. *Educational and Psychological Measurement*, 1968, 28, 659689.

Mosier, Ch. I. Batteries and profiles. In Lindquist, E. F. (Ed.), *Educational measurement*. Washington, D. C.: American Council on Education, 1951.

Naerssen, R. F. van, Een model voor tentamens. *Nederlands Tijdschrift voor de Psychologie*, 1971, 26, 121132, 551559.

Nesselroade, J. R. Application of multivariate strategies to problems of measuring and structuring longterm change. In Goulet, L. R., & Baltes, P. L. (Eds.), *Lifespan developmental psychology. Research and theory*. London, Academic Press, 1970.

Novick, M. R., & Jackson, P. H. Bayesian guidance technology. *Review of Educational Research*, 1970, 40, 459-494.

Novick, M. R., & Jackson, P. H. *Statistical methods for educational and psychological research*. London: McGraw-Hill, 1974.

Pitz, G. F. Subjective probability distributions for imperfectly known quantities. In Gregg, L. W. (Ed.), *Knowledge and Cognition*. London: Wiley, 1974.

Plomp, Tj. *Statistische technieken voor ATI onderzoek*. Enschede: Technische Hogeschool Twente, onderafdeling der toegepaste wiskunde, 1972.

Rapoport, A., & Wallston, T. S. Individual decision behavior. In Mussen, P. H., & Rosenzweig, M. R. (Eds.), *Annual review of psychology*, volume 23. Palo Alto: Annual Reviews Inc., 1972.

Tatsuoka, M. M., & Cattell, R. B. Linear equations for estimating a person's occupational adjustment, based on information on occupational profiles. *British Journal of Educational Psychology*, 1970, 40, 324334.

- Tiedeman, D. V., & Field, F. L. Measurement for guidance. In Whitla, D. K. (Ed.), Handbook of measurement and assessment in behavioral sciences. London: Addison-Wesley, 1968.
- Wang, M. W., & Stanley, J. C. Differential weighting: a review of methods and empirical results. *Review of Educational Research*, 1970, 40, 663705.
- Whitla, D. K. Evaluation of decision making: a study of college admissions. In Whitla, D. K. (Ed.), Handbook of measurement and assessment in behavioral sciences. London: AddisonWesley, 1968.
- Wiggins, J. S. Personality and prediction. London: AddisonWesley, 1973.
- Wiggins, L. M. Panel analysis. Latent probability models for attitude and behavior processee. Amsterdam: Elsevier, 1973.
- Wilbrink, B. Multiple diecriminant analyse van de Cattell Sixteen Personality Factor Questionnaire voor studenten in zeven studierichtingen aan de Technische Hogeschool Eindhoven. (concept rapport). Eindhoven: Technische Hogeschool, groep onderwijsresearch, 1968.
- Wilbrink, B. Overzicht van de selectieproblematiek bij nzcnerus fixus regelingen. Deel I. (concept rapport). Amsterdam: C.O.W.O. Spui 21, april 1975.
- Williams, J. D., Gab, D., & Lindem, A. Judgment analysis for assessing doctoral admission policies. *Journal of Experimental Education*. 1969, 38 no 2. 9296.
- Wilmott, A. S., & Fowles, D. E. The objective interpretation of test performance: the WCH model applied. Windsor, Berks: NFER Publishing Company Ltd, 1974.
- Winter, B. B. Nonparametric density estimation and statistical discrimination. *Psychological Bulletin*, 1974, 81, 371379.
- Wold, H. O. A. Causal inference from observational data: A review of ends and means. *Journal of the Royal Statistical Society (Series A)*, 1956, 119. (Reprinted in Wittrock, M. C., & Wiley, D. (Eds.), *The evaluation of instruction: issues and problems*. London: Holt, Rinehart, & Winston. 1970.)
- Wold, H. O. A. The approach of model building: crossroads of probability theory, statistics, and theory of knowledge. In Peltier, R., & Wold, H. (eds.), *Entretiens de Monaco en sciences humaines, session 1964*. Monaco: Centre International d'Etude des Problèmes Humains, 1966. (Reprinted in Wittrock, M. C., & Wiley, D. (Eds.), *The evaluation of instruction: issues and problems*. London: Holt, Rinehart, & Winston, 1970.)

categorie K.

- Bloom, B. S. Toward a theory of testing which includes measurementevaluation assessment. In Wittrock, M. C., & Wiley, D. (Eds.), *The evaluation of instruction: issues and problems*. London: Holt, Rinehart, & Winston, 1970.
- Bloom, B. S., Hastings, J. T., & Madaus, G. F. Handbook on formative and suffnative eva~uation of student learning. New York: McGrawHill, 1971.
- Day, H. I., & Berlyne, D. A. Intrinsic motivation. In Lesser, G. S. (Ed.), *Psychology and educational practice*. London: Scott, Foresman and Company, 1971.
- Esposito, D. Homogeneous and heterogeneous ability grouping: principal findings and implications for evaluating and designing more effective educational environments. *Review of Educational Research*, 1973, 43, 163179.
- Gagné, R. M. Instructional variables and learning outcomes. In Wittrock, M. C.9 & Wiley, D. (Eds.), *The evaluation of instruction: issues and problems*. London: Holt, Rinehart, & Winston, 1970.
- Glaser, R. Ten untenable assumptions of college instruction. *Educational Record*, 1968, 49. (Reprinted in Roueche, J. E., & Herrscher, B. R. (Eds.), *Toward instructional accountability: a practical guide to educational change*. Palo Alto: Westinghouse Learning Press, 1973.)
- Groot, A. D. de, Some badly needed nonstatistical concepts in applied psychometrics. *Nederlands Tijdschrift voor de Psychologie*, 1970, 25, 360376.
- Mayer, R. F. Measuring instructional intent, or 'Got a match?' Belmont, California: Fearon Publishers, 1973.
- Solomon, G. Heuristic models for the generation of aptitude treatment interaction hypotheses. *Review of Educational Research*, 1972, 42, 327343.

categorie L.

- Annett, J. Feedback and human behavior. The effects of knowledge of results, incentives and reinforcement on learning and performance. Harmondsworth, Middlesex: Penguin Books, 1969.
- Biggs, J. B., & Braun, P. H. Models of evaluation and their relation to student characteristics. *Journal of Educational Measurement*, 1972, 9, 303309.
- Block, J. H. (Ed.), *Mastery learning: Theory and practice (includes annotated bibliography)*. London: Holt, Rinehart, & Winston, 1971.
- Block, J. H. (Ed.), *Schools, society, and mastery learning*. London: Holt, Rinehart, & Winston, 1974.
- Celinski, O. Announced repetitive tests as a basis for self directed study and evaluation. *Journal of Experimental Education*, 1968, 36 no 2, 1726.
- Domino, G. Interactive effects of achievement orientation and teaching style on academic achievement. (ACT Research Report 39). Iowa City, Iowa: American College Testing Publications Office, 1970.
- Hambleton, R. K. Testing and decision making procedures for selected individualized instructional programs. *Review of Educational Research*, 1974, 4451 371400.
- Heckhausen, H. Fear of failure as a self-reinforcing motive system. In Sarason, I. G., & Spielberger, Ch. D. (eds.), *Stress and anxiety (volume 2)*. London: Wiley, 1975.
- Johnson, D. W., & Johnson, R. T. Instructional goal structure: cooperative, competitive or individualistic. *Review of Educational Research*, 1974, 44, 213240.
- Johnson, P. C., & Abrahamson, S. The effects of grades and examinations on selfdirected learning. *Journal of Medical Education*, 1968, 43, 360 366.
- Johnston, J. M. (Ed.), *Behavior research and techno~ogy in higher education*. Springfield, Illinois: Charles C. Thomas, 1975.
- Kirkland, M. C. The effects of tests on students and schools. *Review of Educational Research*, 1971, 41, 303350.
- Lloyd, K. E. Contingency management in university courses. *Educational Technology*, 1971, 11, 1823. (Reprinted in Ulrich, R, Stachnik, T., & Mabry, J. (Eds.), *Control of human behavior, volume three*. Glenview, Illinois: Scott, Foresman and Company, 1974.)
- Locke, E. A., Cartledge, N., & Koeppel, J. Motivational effects of knowledge of results: a goalsetting phenomenon?

Psychological Bulletin, 1968, 70 , 474485.

- Mager, R. F., & Pipe, P. Analyzing performance problems or 'You really oughta wanna'. Belmont, California: Fearon, 1970.
- Michael, J. L. The essential components of effective instruction and why most teaching is not. In Keller, F. S., & RibesInesta, E. (Eds.), Behavior modification: applications to education. London: Academic Press. 1974.
- Sherman, J. G. (Ed.), Personalized System of Instruction: 41 germinal papers, a selection of readings on the Keller plan. Amsterdam: Benjamin Inc., 1974.

categorie N.

- Feinberg, Walter. Reason & rhetoric. The intellectual foundations of twentieth century liberal educational policy. London: Wiley, 1975.
- Groot, A. D. de, Selectie voor en in het hoger onderwijs: een probleem analyse. 'sGravenhage: Staatsuitgeverij, 1972.
- Mulder, J. de, (Red.), Selectie van studenten. Achtergrond, effecten. Utrecht: OTO hoger onderwijs cahiers 13, 1974.
- Munday, L. A., & Davis, J. C. Varieties of accomplishment after college: perspectives on the meaning of academic talent. (ACT Research Report 62), Iowa City, Iowa: The American College Testing Program Publications Office, 1974.

categorie O.

- Astin, A. W. Predicting academic performance in college; sglectivity data for 2300 American colleges. London: Collier Macmillan, 1971.
- Astin, A. W., & Panos, R. J. The educational and vocational development of college students. Washington, D. C: American Council on Education, 1069.
- Baird, L., & Feister, W. J. Grading standards: the relation of changes in average student ability to the average grades awarded. American Educational Research Journal, 1972, 9, 431441.
- Bowles, F. Access to higher education (2 volumes). Paris: Unesco, 1963, 1965.
- Braak, L. H. Geïndividualiseerde onderwijssystemen: konstruktie en besturing. (proefschrift) Eindhoven: Technische Hogeschool, 1974.
- Choppin, B. H. L., Orr, L., Kurle, S.D.M., Fara, P., & James, G. The prediction of academic success. Windsor, Berks: NFER Publishing Company Ltd, 1973.
- Eggleston, J. F., & Kerr, J. F. gtudies in assessment. London: The English Universities Press Limited, 1969.
- Everwijn, S. E. M., Koopmans, F., & Vroon, A. G. Rendement van een toelatingsselectie. (Mededeling nr 22). Utrecht: Onderzoek en Ontwikkeling van Onderwijs, 1974.
- Fincher, C. Is the SAT worth its salt? An evaluation of the use of the Scholastic Aptitude Test in the university system of Georgia over a thirteenyear period. Review of Educational Research, 1974, 44, 293305.
- Goldman, R. D., Schmidt, D. E., Hewitt, B. N., & Fisher, R. Grading practices in different major fields. American Educational Research Journal, 1974, 11, 343~357.
- Groot, A. D. de, Vijven en zessen. Groningen: WoltersNoordhoff, 1966.
- Hearnden, A. Paths to university: preparation, assessment, selection. London: Macmillan Education Ltd, 1973.
- Hoyt, D. P. The relationship between college grades and adult achievement. A review of the literature. (ACT Research Report 7) Iowa City, Iowa: The American College Testing Program Publications Office. 1965.
- Ingenkamp, K. (Hrsg.) Die Fragwürdigkeit der Zensurengebung. Weinheim and Basel: Beltz Verlag, 1971.
- Kopfman, G. lusainmenhang zwischen Abiturnoten von Physikstudenten und ihrer Leistung an der T. H. München. Zeitschrift zur Experimentelle und Angewandte Psychologie, 1969, 16, 434451.
- Miller, G. W. Success, failure and wastage in higher education. London: George Harrap, 1970.
- Parkyn, G. W. Success and failure at the university. Volume II The problem of failure. Wellington: New lealand Council for Educational Research, 1967.
- Peet, A. A. J. van, & Munnik, M. de, Onderwijssystemen, selectieregelingen en doorstromingscijfers in verschillende landen. In Groot, A. D. de, Selectie voor en in het hoger onderwijs, een probeemanalyse. 'sGravenhage: Staatsuitgeverij, 1972.
- Powell, J. L. Selection for university in Scotland. London: University of London Press, 1973.
- Richards, J. M., Jr., Calkins, E. V., McCanse, A., & Burgess, M. M. Predicting performance in a combined undergraduate and medical education program. Educational and Psychological Measurement, 1974, 34, 923931.
- Willingham, W. W. The sourcebook for higher education. A critical guide to literature and information on access to higher education. New York: College Entrance Examination Board, 1973.
- Willingham, W. W. College placement and exemption. New York: College Entrance Examination Board, 1974.

categorie P

- Astin, A. W., & Panos, R. J. The evaluation of educational programs. In Thorndike, R. L. (Ed.), Educational Measurement. Washington, D. C.: American Council on Education, 1971.
- Baker, E. L. Formative evaluation of instruction. In Popham, W. J. (Ed.), Evaluation in education: current applications. Berkeley, California: McCutchan, 1974.
- Brennan, R. L. A method for the use of achievement data and time data in an instructional system. Instructional Science 1975, 4, 113136.
- Carver, R. P. A model for using the final examination as a measure of the amount learned in classroom learning. Journal of Educational Measurement, 1969, 6, 5968.
- Costin, F., Greenough, W. T., & Menges, R. J. Student ratings of college teaching: reliability, validity, and usefulness. Review of Educational Research, 1971, 41, 511535.
- Cronbach, L. J. Evaluation for course improvement. Teachers Record, 1963, 64, 672683. (Reprinted in Gronlund, N. E. (Ed.), Readings in measurement and evaluation. New York: MacMillan, 1968.)
- Entwistle, N. J., & Brennan, T. The academic performance of students: types of successful students. British Journal of Educational Psychology, 1971, 41, 268276.
- Johnson, H. C., Jr., Rhodes, D. M., & Rumery, R. E. The assessment of teaching in higher education: a critical retrospect and a proposal. Higher Education, 1975, 4, 173199, 273303.
- Lindvall, C. M., & Cox, R. C. The IPI evaluation program. (AERA monograph series on curriculum evaluation). Chicago: Rand

McNally, 1970.

Marco, G. L. A comparison of selected school effectiveness measures based on longitudinal data. *Journal of Educational Measurement*, 1974, 11, 225-234.

Messick, S. P. The criterion problem in the evaluation of instruction: assessing possible, not just intended, outcomes. In Wittrock, M. C., & Wiley, D. (Eds.), *The evaluation of instruction: issues and problems*. London: Holt, Rinehart, & Winston, 1970.

Miller, R. I. *Developing programs for faculty evaluation*. (includes annotated bibliography) San Francisco: Jossey-Bass, 1974.

Posner, G. J. The extensiveness of curriculum structure: a conceptual scheme. *Review of Educational Research*, 1974, 44, 401-407.

Sanders, J. R., & Cunningham, D. J. A structure for formative evaluation in product development. *Review of Educational Research*, 1973, 43, 217-236.

Short, E. C. Knowledge production and utilization in curriculum: a special case of the general phenomenon. *Review of Educational Research*, 1973, 43, 237-301.

Stones, E. *Educational objectives and the teaching of educational psychology*. London: Methuen, 1972.

Stufflebeam, D. L., Foley, W. J., Gephart, W. J., Guba, E. G., Hammond, R. L., Merriman, H. O., & Provus, M. M. *Educational evaluation & decision making*. Itasca, Illinois: Peacock, 1971.

categorie Q

Atkinson, J. W., & Raynor, J. O. (Eds.), *Motivation and achievement*. London: Wiley, 1974.

Berliner, D. C., & Cahen, L. S. Trait-treatment interaction and learning. In Kerlinger, N. (ed.), *Review of educational research*, volume 1. Itasca, Illinois: Peacock, 1973.

Cattell, R. B., & Child, D. *Motivation and dynamic structure*. London: Holt, Rinehart, & Winston, 1975.

Cronbach, L. J., & Snow, R. E. *Aptitudes and instructional methods*. New York: Appleton-Century-Crofts, in press.

Gagné, R. M. (Ed.), *Learning and individual differences*. Columbus, Ohio: Merrill, 1967.

Holland, J. G. The misplaced adaptation to individual differences. APA symposium paper, 1969. (Reprinted in Weisgerber, R. A. (Ed.), *Perspectives in individualized learning*. Itasca, Illinois: Peacock, 1971.)

Humphreys, L. G. The fleeting nature of the prediction of academic success. *Journal of Educational Psychology*, 1968, 59, 375-380.

Maehr, M. L., & Sjogren, D. D. Atkinson's theory of achievement motivation: first step toward a theory of academic motivation? *Review of Educational Research*, 1971, 41, 143-161.

Plomp, Tj. *De ontwikkeling van een individueel studie systeem*. Groningen: Tjeenk Willink, 1974.

Sperry, L. (Ed.), *Learning performance and individual differences: essays and readings*. London: Scott, Foresman and Company, 1972.

Hoe is het 50 jaar later? Ik neem die literatuurlijst even door, ben toch wel benieuwd. De lijst is wel heel lang, ik heb alleen het begin doorgenomen.

categorie 2.

Alkin, M. C. Evaluating the cost-effectiveness of instructional programs. In M. C. Wittrock & D. E. Wiley (Eds.), *The evaluation of instruction: issues and problems*. London: Holt, Rinehart, & Winston, 1970. Evaluatie is natuurlijk altijd wel een thema gebleven, maar het is een tikje ongrijpbaar. Ik heb in 1987 een evaluatie van een opleiding moeten uitvoeren, samen met Joep Verbeek. Op dat moment kon ik het af zonder een theoretisch kader op basis van de literatuur te schetsen, als ik mij goed herinner.

Davis, J. H. Group decision and social interaction: a theory of social decision schemes. *Psychological Review*, 1973, 80, 97-125. Herinner ik me niets van. Ik geef dat maar aan met ??

Lindley, D. *Making decisions*. London: Wiley, 1971. Een veel te technisch boek, ik heb het niet kunnen gebruiken geloof ik.

Marco, G. L. A comparison of selected school effectiveness measures based on longitudinal data. *Journal of Educational Measurement*, 1974, 11, 225-234. Nou ja, die komt in mijn vroege loopbaan nog vaak voor: random publicaties die ik wel noteer, maar die achteraf gezien rustig vergeten kunnen worden. Ik duid dat maar aan met --

Naerssen, R. F. van A scale for the measurement of subjective probability. *Acta Psychologica*, 1962, 26, 159-166. Ook dit heb ik niet kunnen benutten, hoewel het nuttig was te weten dat er een literatuur over subjectieve waarschijnlijkheden bestaat.

Raiffa, H. *Decision analysis. Introductory lectures on choices under uncertainty*. London: Addison-Wesley, 1968. Het was goed om het boek bij de hand te hebben, uiteindelijk bleek er een toch wel serieus probleem me samen te hangen: gebruik je de statistische beslissonderwijs, zoals Wim van der Linden heeft gedaan, met eindeloos gecompliceerde formules, of de economische beslissonderwijs, zoals in mijn TOR-artikelen in 1980 heb gedaan, met eenvoudige formules en dataverzameling?

Vlek, Ch. A. J. The fair betting game as an admissible procedure for assessment of subjective probabilities. *British Journal of Mathematical and Statistical Psychology*, 1973, 26, 183-190. Ik ben deze weg dus niet gegaan.

Wilbrink, B. Het nemen van beslissingen. (interne nota) Amsterdam: C.O.W.O. spui 21, 1975. https://benwilbrink.nl/projecten/75_Beslissen_dichotoom_1-8-75.pdf

Witmer, D.R. Economic benefits of college education. *Review of Educational Research*, 1970, 40, 511-524. -- maar menselijk kapitaal theorie zou wel een belangrijke rol gaan spelen in mijn denken, en mijn boekenkast.

categorie 3.

Astin, A. W. Personal and environmental factors associated with college dropouts among high aptitude students. *Journal of Educational Psychology*, 1964, 55, 219-227. Astin is altijd interessant. Ik heb een aantal boeken.

Dawes, R. H. & Corrigan, B. Linear models in decision making. *Psychological Bulletin*, 1974, 81, 94106. Key!

Hackman, J. R., Wiggins, N., & Bass, A. R. Prediction of longterm success in doctoral work in psychology. *Educational and Psychological Measurement*, 1970, 30, 365374. --

Harris, D. Factors affecting college grades: a review of the literature, 1930 1937. *Psychological Bulletin*, 1940, 37, 125-166. Dat was een andere wereld.

Kahneman, D., & Tversky, A. On the psychology of prediction. *Psychological Review*, 1973, 80, 237251. Key

Mehrabian, A. Undergraduate ability factors in relationship to graduate performance. *Educational and Psychological Measurement*, 1969, 29, 409419. --

Schrader, W. B. The predictive validity of College Board admissions tests. In Angoff, W. H. (Ed). *The College Board admissions testing program*. New York: College Entrance Examination Board, 1971. --

Thacker, A. J., & Williams, R. E. The relationship of the graduate record examination to grade point average and success in graduate school. *Educational and Psychological Measurement*, 1974, 34, 939944. --

Warburton, F. W. The relationship between personality factors and scholastic attainment. In Dreger, R. M. (Ed), *Multivariate personality research*. Baton Rouge: Claitor's Publishing Division, 1972. --

Wiggins, N., & Kohen, E. S. Man versus model of man revisited: the forecasting of graduate school success. *Journal of Personality and Social Psychology*, 1971, 19, 100106. --

categorie A.

Abe, C., & Holland, J. L. A description of college freshmen. I. Students with different choices of major field. II. Students with different vocational choices. (Research Report 3 & 4). Iowa City: American College Testing Program Publications Office, 1965. Hoort bij mijn 1968 rapportage 16PF MDA.

Baird, L. L. The undecided student How different is he? (ACT Research Report 22). Iowa City: American College Testing Program Publications Office, 1967. --

Bossel, H. College student and dropout problem: a qualitative dynamic simulation. *Instructional Science*, 1974, 3, 2350. --

Cooley, W. W., & Lohnes, P. R. Predicting development of young adults. (Interim report 5 to the U. S. Office of Education, Cooperative Research Project no. 3051.) Palo Alto: Project TALENT Office, American Institutes for Research and University of Pittsburgh, 1968. --

Flanagan, J. C., Shaycoft, M. F., Richards, J. M., Jr., & Claudy, J. G. Five years after high school. (Final report to the U. S. Office of Education.) Palo Alto: Project TALENT Office, American Institutes for Research and University of Pittsburgh, 1971. --

Gelatt, H. B., Varenhorst, B., Carey, R., & Miller, G. P. Decisions and outcomes (+ idem, A leader's guide). New York: College Entrance Examination Board, 1973. -- Lang bewaard, dat wel.

Gelatt, H. B., Varenhorst, B., & Carey, R. Deciding (+ Deciding: A leader's guide). New York: College Entrance Examination Board, 1972. --

Jepsen, D. A., & Dilley, J. S. Vocational decisionmaking models: a review and comparative analysis. *Review of Educational Research*, 1974, 44, 331349. --

Marks, E. Cognitive and incentive factors involved in within university transfer. *Journal of Educational Psychology*, 1970, 61, 19. --

Moore, W., Jr., Against the odds. The highrisk student in the community college. San Francisco: Jossey-Bass, 1970. --

Saenger-Ceha, M. M. Th. Psychological and social factors in student dropout. (proefschrift) Amsterdam: Swets and Zeitlinger, 1970. Er was toch wel een schaarste aan uitvalonderzoek. En wat er was, gaf geen aandacht aan positieve motieven om iets anders te gaan doen.

Tiedeman, D. V., & Field, F. L. Measurement for guidance. In Whitla, D. K. (Ed), *Handbook of measurement and assessment in behavioral sciences*. London: AddisonWesley, 1968.

Tinto, V. Dropout from higher education: a theoretical synthesis of current research. *Review of Educational Research*, 1975, 45, 89125. Tinto zou nog lang de expert blijven; heeft jater ook Amsterdam bezocht en daar enige tijd gewerkt.

Watts, A. G. Diversity and choice in higher education. London: Routledge & Kegan Paul, 1972. ?? MDA 16PF

Wish, P. A.51 & Hasazi, J. E. Motivational determinants of curricular choice in college males. *Journal of Counseling Psychology*, 1973, 20, 127131. --

categorie B.

Block, J. (Ed.). *Schools, society and mastery learning*. London: Holt, Rinehart, & Winston, 1974. Maar wat stond er nou precies in?

Millman, J. Criterion referenced measurement. In Popham, W. J. (Ed.), *Evaluation in education; current applications*. Berkeley: McCutchan, 1974. --

Swaminathan, H., Hambleton, R. K., & Algina, J. A bayesian decision theoretic procedure for use with criterionreferenced tests. *Journal of Educational Measurement*, 1975, 12, 87-98. Regelmatig geraadpleegd, maar heb ik het ook kunnen gebruiken?

Wijnen, W. H. F. W. Onder of boven de maat; een methode voor het bepalen van de grens voldoende onvoldoende bij studietoetsen. Amsterdam: Swets & Zeitlinger, 1972. We (COWO) hebben er uitvoerig met Wynand over gesproken, als ik mij goed herinner. Helemaal zeker weet ik het niet meer. Misschien komt de correspondentie nog boven water.

Wijnen, W. H. F. W. Het diagnostisch gebruik van studietoetsen. In Van Woerden, W. W., Chang, T. M., & Van Geuns-Wiegman, L. J. M., (Red.) *Onderwijs in de maak*. Utrecht: Het spectrum, 1973. Aan dat boek zou ik ook meewerken, maar dat is op de een of andere manier verwaarloosd. Ik zou zo niet kunnen weggen welke cruciale bijdragen er in dat boek staan.

categorie C.

- Angoff, W. H. (Ed.). The College Board admissions testing program; a technical report on research and development activities relating to the Scholastic Aptitude Test and Achievement Tests. New York: College Entrance Examination Board, 1971. --
- Cole, N. S., Hanson, G. R. Assessing students on the way to college. (2 volumes). Iowa City, Iowa: The American College Testing Program, 1973. ??
- Cronbach, L. J., & Gleser, G. C. Psychological tests and personnel decisions. London: University of Illinois Press, 1965. (Originally published, 1957) Key Key Key!
- Fuess, C. M. The College Board; its first fifty years. New York: College Entrance Examination Board, 1967. --
- Hills, J. R. Use of measurement in selection and placement. In Thorndike, R. L. (Ed.), Educational Measurement. Washington, D. C. American Council on Education, 1971. Uit dit boek heb ik toch wel veel kennis gehaald.
- Lord, F. M. Cutting scores and errors of measurement. Psychometrika, 1962, 27, 1930. Key voor compensatorische examens.
- Naerssen, R. F. van, Het nut van selectie, enkele voorbeelden. In Drenth, P. J. D., Willems, P. J., & de Wolff, Ch. J. Bedrijfspsychologie; onderzoek en evaluatie. Kluwer, Van Loghum Slaterus, 1970. Over Taylor Russel tabellen ook. Het is op de achtergrond wel een inspiratie geweest voor mijn 1978 Examens deel A (B nooit geschreven).
- Thorndike, R. L. Personnel selection; test and measurement techniques. New York: Wiley, 1949. Key voor selectie en bijv. restriction of range.

categorie D.

- Birnbaum, A., & Maxwell, A. E. Classification procedures based on Bayes's formula. In Cronbach, L.J., & Gleser, G. C. Psychological tests and personnel decisions. London: University of Illinois Press, 1965. Ik heb deze niet bestudeerd, geloof ik.
- Cole, N. S. Differential validity in the ACT tests. (ACT Research Report 30). Iowa City, Iowa: American College Testing Program Publications Office, 1969. --
- Folger, J. K., Astin, H. S., & Bayer, A. E. Human resources and higher education. New York: Russell Sage Foundation, 1970. Menselijk kapitaal.
- Loeb, J., & Bowers, J. Programs of study as a basis for selection, placement and guidance of college students. Journal of Educational Measurement. 1973, 10, 131139. --
- Rulon, P. J., Tiedeman, D. V., Tatsuoaka, M. M., & Langmuir, C. R. Multivariate statistics for personnel classification. New York: Wiley, 1967. MDA 16PF, mijn 1968
- Willingham, W. W. College placement and exemption. New York: College Entrance Examination Board, 1974. ??