

[Gebaseerd op een beroerde scan, het lukt niet om deze tekst in redelijke tijd netjes te krijgen. Het belang ervan kan ik zo ook niet goed beoordelen.
[ergens in 1974? Aantekening voor eigen gebruik]

cesuur

Beschrijving, bespreking en bewerking van Van Naerssen's studies over het combineren van studieonderdelen (model voor tentamens).

Drie publicaties van Van Naerssen behandelen dit onderwerp:

Over optimaal studeren en tentamens combineren. Openbare les, 1970,
Een model voor tentamens, Ned.T.v.d.Psychol. 1971, 26, 121132. Een model voor tentamens (vervolg) NTvdPsychol. 1971, 26, 551559. [Zie <https://benwilbrink.nl/publicaties/70vNaerssenLes.htm> voor de openbare les en voor summaries van andere publicaties over dit thema. bw 2025]

Van Naerssen's model is van belang voor het oplossen van probleemstellingen m.b.t. het combineren van verschillende studieonderdelen bij het nemen van examenbeslissingen, voor vragen over herkansingsmogelijkheden voor een bepaald tentamen (hoeveel tijd tussen tentamen en herkansing, kan niet beter het tentamen wat uitvoeriger gemaakt worden zodat een herkansing in aanzienlijke mate overbodig wordt, e.d.) en voor meeromvattende problemen rond herkansingsmogelijkheden na gezakt zijn voor de propedeuse.

Daarnaast is het evident dat de cesuurproblematiek voor het afzonderlijke tentamen in sterke mate afhangt van de strategie die gevolgd wordt bij het combineren van afzonderlijke tentamenuitslagen tot examenbeslissingen. In die zin had Wijnen zich dan ook niet mogen onttrekken aan het relateren van zijn cesurbepalingsmethode aan enkele bekende methoden voor het combineren van studieonderdelen tot examenbeslissingen (Wijnen, onder of boven de maat, 1971).

De bedoeling is om het werk van Van Naerssen te analyseren op zaken die voor de dagelijkse onderwijspraktijk van belang kunnen zijn, Op die relevantie immers zal het deel van de docentencursus dat gaat over de cesurbepaling gebaseerd moeten zijn. Daarvoor is het niet voldoende om de door een persoon aangegeven lijnen eenvoudig door te trekken tot praktische aanwijzingen voor de docent, maar zullen ook de veronderstellingen die expliciet of impliciet aan zijn model inherent zijn kritisch bekeken moeten worden. In dit verband moet dan ook niet vreemd opgekeken worden wanneer deze bespreking zou beginnen met behandeling van het allesofniets model voor leren, bijvoorbeeld.

Testleer en docimologie.

Waarom is het combineren van tentamens een speciaal probleem, we kennen uit de testleer immers diverse methoden om scores optimaal te combineren bijv. om een bepaalde criterium variabele te voorspellen?

Van Naerssen 1970:

"Er bestaat over het combineren van tests tot het verkrijgen van een zo hoog mogelijke validiteit of utiliteit een uitgebreide literatuur. Deze beperkt zich echter tot de gebruikelijke modellen van de testleer en tot de multivariate statistische technieken. We zullen zien dat nieuwe aannamen nodig zijn om deze problematiek op te lossen, assumpties die buiten de testleer vallen."

Een indruk van die 'uitgebreide literatuur' kan verkregen worden uit John R. Hills: *Use of measurement in selection and placement*. In *Educational Measurement*, edited by Robert L. Thorndike, 1971. Uit de behandeling die Hills van het onderwerp geeft, waarin ook bij herhaling gesproken wordt over combinatie problemen, multiple cutting scores e.d., blijkt echter niet veel van een onbruikbaarheid van de testleer voor het oplossen van het combinatie probleem, zoals Van Naerssen veronderstelt. Het vermoeden bestaat, dat het werk van Van Naerssen een welkome aanvulling is op wat anderszins reeds uit de testleer bekend is, maar deze geenszins vervangt. Het zou interessant zijn om uit te zoeken in hoeverre het hele apparaat van de testleer toepasbaar is nadat een analyse langs de lijnen uitgezet door Van Naerssen is uitgevoerd.

Van belang bij het kiezen van een combinatie methode is (Van Naerssen 1970) is van welk criterium men gebruik maakt. Gaat het alleen om de betrouwbaarheid van de beslissing of om het peil van de geslaagden, dan kan men terecht bij de testleer, eventueel uitgebreid met wat besliskundige begrippen. Maar gaat het om de tijd, die de student nodig heeft om voor het geheel te slagen, dan zal men de theorie verder moeten uitbreiden.

Dit is ook nodig wanneer we problemen willen oplossen als:

Hoe groot moeten de tijdsafstanden zijn tussen de tentamens?

Wat is eigenlijk het nut van een zogenaamde herkansing?

Als we de compensatorische methode gebruiken wat zijn dan de gevolgen van de vrijheid om te mogen kiezen welk tentamen men overdoet als men gezakt is voor het geheel?

Of is het soms efficiënter het geheel te laten overdoen?

Of dit wel maar dan met een lagere aftestgrens?

Wat zijn precies de consequenties van een wel vaker gebruikte methode waarbij de student voor elk tentamen een minimum score moest halen, maar bovendien een minimum score voor het geheel?"

Om deze vragen onderzoekbaar te maken, moeten we volgens Van Naerssen allereerst kunnen beschikken over een theorie waaruit toetsbare hypothesen af te leiden zijn. Als aanzet voor die theorievorming presenteert hij in zijn openbare les een mathematisch model, waarbij gebruik gemaakt wordt van een ander model uit de mathematische psychologie: het model voor allesofniets leren.

De door Van Naerssen gestelde vragen zijn waarschijnlijk ook wel zonder de door hem voorgestelde theorie te beantwoorden. Een vraag bijv. als naar de efficiëntie van het laten overdoen van alle studieonderdelen, voor een bepaald examen wanneer het laatste onvoldoende is gemaakt, laat zich ook wel analyseren en

onderzoeken met behulp van de reeds ontwikkelde testtheoretische modellen. Tenminste zou je er nog vanuit kunnen gaan dat bij het niet aantoonbaar zijn van enig positief rendement van een herhalingsprocedure, deze afgeschaft of door andere procedures vervangen moet worden. Maar goed, een modelmatige benadering zoals door Van Naerssen voorgesteld kan een vruchtbare analyse methode opleveren.

een speltheoretische benadering?

Van Naerssen ziet de situatie als een spel van studenten tegen staf. Voor één van de partijen ziet hij een duidelijke optimaliseringsstrategie: De student (Van Naerssen 1970) "minimaliseert zijn totale studeertijd. Dit althans is als zijn optimale strategie te beschouwen. Dat hij hiernaast nog vele andere doeleinden nastreeft tijdens zijn studie doet voor het onderhavige probleem niets ter zake."

Wat de staf maximaliseert is niet zo duidelijk;

"zoiets als zoveel mogelijk relevante vaardigheden overbrengen aan zoveel mogelijk studenten in zo min mogelijk tijd en met zo laag mogelijke kosten, een verwarrende hoeveelheid doeleinden."

Het opvallende van de verdere analyse van Van Naerssen is dat hij deze uitvoert zonder ook maar iets te veronderstellen over die maximalisatie die de staf zou kunnen nastreven. Daarvoor in de plaats komen zaken als het aantal items dat in een tentamen opgenomen moet worden, de tijd tussen herkansingen e.d., m.a.w. geheel andere zaken dan hij in zijn lijstje van mogelijk door de staf te maximaliseren zaken heeft opgenomen. In onze analyse zullen we dan ook moeten nagaan hoe een en ander te rijmen is met het door hem gekozen uitgangspunt. Mogelijk is bijv. dat het maximaliserings probleem voor de staf uitgesteld kan worden tot na de analyse van het combinatieprobleem.

Het maximaliseringsprobleem voor de staf, of voor de onderwijsinstelling, is gelijk aan het criteriumprobleem uit de toegepaste testleer. Dat er een verwarrende hoeveelheid doelen tegelijk nagestreefd kunnen worden is op zich geen reden om maar van een analyse van wat de criteriumvariabele is, afgezien kan worden. [sic]

Wanneer Van Naerssen op ontwapenend simpele wijze het optimaliseringsprobleem voor de student oplost door te stellen dat hij zijn totale studietijd minimaliseert, is het moeilijk in te zien wanneer niet op dezelfde wijze voor de staf gesteld kan worden dat zij hun totale onderwijstijd minimaliseren. Het is wat minder plausibel, maar is niet de plausibiliteit van de geminimaliseerde studietijd voor de student gebaseerd op een ongetoetste persoonlijke opvatting?

Er zijn overigens nog een aantal andere, geen maximaliserings-strategieën die door de staf gevolgd kunnen worden, en die van groot belang zijn voor de optimale strategie van de student: een voorbeeld is een quota selectie aan het eind van de propedeuse op basis van de studieresultaten, een procedure die door sommigen voorgesteld wordt als meest humane oplossing van het selectievraagstuk voor

numerus fixus studierichtingen. De optimale strategie van de student in die situatie is waarschijnlijk zo lang en zo hard mogelijk studeren.

Genoemde bezwaren zijn zeker geen scherpslijperijen. Toch zijn ze op dit moment niet van groot belang. Laten we eerst maar eens zien hoever we komen met de veronderstellingen die Van Naerssen hier wil hanteren, waarna altijd nog nagegaan kan worden in hoeverre de resultaten zich wijzigen bij het invoeren van enigszins andere aannamen, en in hoeverre het model dan nog staande blijft in empirische toetsing.

van Naerssen kiest een speltheoretisch uitgangspunt, wat o.a. betekent dat maximaliseren van de winst, m.a.w. van de utiliteit, van belang is. Gecontrasteerd met zijn eerdere uitspraak dat de literatuur over het combineren van tests tot het verkrijgen van een zo hoog mogelijke utiliteit de beantwoording van een aantal eenvoudige vragen (zoals geciteerd) niet toelaat, maakt duidelijk dat Van Naerssen bij zijn analyse kennelijk vast blijft houden aan de principes: zo hoog mogelijke validiteit en utiliteit, en dat hij alleen maar enige nieuwe technieken wil ontwikkelen om hetzelfde doel te bereiken. We zullen de analyse die Van Naerssen geeft dan ook bijzonder kritisch moeten bekijken op het punt van de definitie van utiliteit die (impliciet of expliciet) gehanteerd wordt, vooral ook gezien zijn uitspraak dat niet duidelijk is wat nu eigenlijk de maximale utiliteit voor de staf (= de onderwijsinstelling) is.

Maar waarom zou je een speltheoretisch uitgangspunt kiezen zonder het ook speltheoretisch uit te werken? Wie spelen tegen wie, is bijv. een voor de hand liggende vraag. Zijn dat de studenten tegen de staf, zijn dat de studenten tegen elkaar, zijn dat verschillende secties tegen elkaar (docenten tegen elkaar, eventueel over de rug van de studenten ?)

Moeten we denken aan competitieve of aan coöperatieve modellen?

Welke mogelijke strategieën kunnen gespeeld worden door bovengenoemde partijen, en welke is de winst en verlies matrix die daarbij hoort? Op welke wijze worden in die winst en verlies matrix de foute beslissingen ingecalculleerd? etcetera, etcetera.

Ongetwijfeld zal Van Naerssen sommige van deze vragen beantwoorden, en andere laten liggen. Als het model van Van Naerssen uitgebreid beschreven is, zullen we op genoemde vragen terugkomen, en nagaan in hoeverre ze in het model geïncorporeerd zijn of niet.

Een verwante vraag die zich opdringt is in hoeverre in het tentamen en examenspel sprake is of kan zijn van onderhandeling vooraf tussen beide partijen. De Groot zou kunnen vragen in hoeverre transparantie van de tentamensituatie een rol speelt, en onderhandeling tussen studenten en staf over de transparantie van de toetssituaties (het model van Van Naerssen schijnt een volledige isomorfie te veronderstellen tussen leerstof 'items' en tentamen 'items', waarvan de praktische realiteit op zijn zachtst gezegd twijfelachtig is, wat ten koste van de transparantie van de toetssituaties kan gaan). Met andere woorden, het model veronderstelt waarschijnlijk ook heel bepaalde toetsvormen, constructiemethoden voor items, en

mogelijkheden voor studenten om zich voor te bereiden op de bedoelde toetsingen.

Voorlopig ziet het er even naar uit dat Van Naerssen gaat analyseren wat de optimale strategie van de student is, dat hij veronderstelt dat deze optimale strategie gelijk blijft onder een groot aantal mogelijke strategieën die door de staf gevolgd kunnen worden, en dat hij daarom na het vastgelegd hebben van de optimale strategie van de student, gerust kan gaan zoeken naar wat in die omstandigheid de optimale strategie van de docent zou kunnen zijn. Met andere woorden, het lijkt er op dat in het examenspel van Van Naerssen een zadelpunt aanwezig is. De vraag is of hij er al van begin af aan vanuit gegaan is dat er een zadelpunt moet zijn (dat zou wel zo prettig zijn voor de analyse), of dat het bestaan van een zadelpunt deductief is vastgesteld.

Tot zover het speltheoretisch element; we zullen er later nog op terugkomen.

Optimale strategie van de student.

De student minimaliseert zijn totale studeertijd.

Deze veronderstelling van Van Naerssen is niet geheel onschuldig. Hij heeft namelijk precies deze veronderstelling nodig voor het model zoals hij dat gaat opzetten, en geen andere veronderstellingen leiden (even makkelijk) tot dat model.

Zijn er andere veronderstellingen te maken over de optimale strategie van de student, die minstens even plausibel zijn? De veronderstelling van Van Naerssen leidt weliswaar tot het kunnen gebruiken van een aardig model, maar het heeft er veel van weg dat de student van Van Naerssen net zo weinig reëel is als 'rational man' van de econoom.

Laten we eens wat proberen:

- de student wil ze prettig mogelijk studeren;
- de student wil binnen een redelijke studietijd zich zo goed mogelijk in zijn vakgebied inwerken;
- de student wil zijn eigen doelen zo goed mogelijk realiseren, ook al wijken ze sterk af van de (impliciete doelstellingen zoals die bijv. achter de hem gestelde tentamenvragen liggen;
- de student wil een grote vrijheid hebben bij de keuze van zijn vakkenpakket (op de student toegesneden studieprogramma's, geen eenheidspakketten voor grote groepen studenten);
- de student wil een selectievrij onderwijs;
- de student wil prettig kunnen samenwerken met zijn medestudenten;
- de student wil getoetst worden op een wijze die hem goed ligt (niet noodzakelijk een multiple choice vorm voor iedere toetsing in het onderwijs, er moet een grotere variëteit in toetsingsmogelijkheden komen, met in achtneming van een aantal fundamentele principes als billijkheid, transparantie, objectiviteit etc., zij het natuurlijk dat deze eisen niet absoluut zijn maar er tot op aanvaardbare hoogte aan

voldaan moet zijn;

de student wil dat hij in staat gesteld word minder dan optimale studiestrategieën te volgen, wanneer daar volgens hem redenen toe zijn,

Het lijstje zal nog wel uit te breiden zijn. Wel is nu al duidelijk dat de veronderstelling van Van Naerssen in sterke mate willekeurig is, en dat we niet met hem mee kunnen gaan waar hij zegt (Van Naerssen 1970): "Dat hij hiernaast nog vele andere doelen nastreeft tijdens zijn studie doet voor het onderhavige probeem niets ter zake." Van Naerssen pretendeert namelijk wel degelijk dat zijn model leidt tot praktische oplossingen voor een aantal problemen rond het combineren van studieonderdelen tot examenbeslissingen. Welnu, die 'praktische oplossingen' moeten dan ook getoetst worden aan de 'vele andere doeleinden' die de student kan nastreven, en ook de andere bedoelingen die docenten kunnen hebben t.a.v. de studeerwijze te volgen door de student.

Ook nu zullen we voorlopig met de veronderstellingen van Van Naerssen meegaan, kijken waar die ons naartoe voeren, en na afloop zullen we op onze schreden terugkeren en na moeten gaan in hoeverre enig tornen aan de veronderstellingen van Van Naerssen zijn model, of de uitkomsten er van, wijzigt of ondergraaft.

Van Naerssen (1970):

"Dit betekent dat we een verband moeten aannemen tussen enerzijds studeertijd en anderzijds de geleerde vaardigheid. Deze laatste kunnen we weer met enig recht gelijk stellen aan de zogenaamde ware score op de als tentamen gebruikte studietoets.

In de psychologische functieleer zijn verschillende modellen ontwikkeld die een dergelijk verband aangeven. Een aantrekkelijke theorie is die van het allesofniets leren, waarvan vele varianten bestaan, die weer in één model kunnen worden samengevoegd. Een eenvoudige variant is ontworpen voor het memoriseren van een rijtje woorden, maar is misschien ook bruikbaar voor het leren van een boek, syllabus of collegediktaat."

Van Naerssen bedoelt de variant EstesBower, m.a.w. wat normalerwijze onder het allesofniets leren model wordt verstaan (zie Donald Laming 'Mathematical Psychology' 1973. of Coombs, Dawes & Tversky 'Mathematical Psychology' 1970).

Is dit model bruikbaar voor het leren van een boek? De veronderstelling bij allesofniets leren is dat het leermateriaal bestaat uit eenvoudig materiaal, Laming (1970. blz 251:),

"One may view attended items as entering longterm memory and unattended items as nominally entering shortterm memory, inasmuch as they are presented to the sense organs. If the material to be learned is simple, one may suppose that attention is sufficient to ensure that learning takes place. But if the material is complex (nonsens lettergrepen zijn in deze zin al complex gebleken) it is conceivable that only part of each item may be learned. Herein is a reason why all or none learning should usually be found only with simple material. The unattended items in shortterm memory are discarded, possibly through overwriting by subsequent items; for such items learning begins afresh the next time they are presented."

Misschien is het mogelijk om in plaats van de moeilijk te beargumenteren aanname van Van Naerssen, gebruik te maken van de opvattingen van Gagné, zoals neergelegd in zijn Conditions of learning. Wanneer we dan kijken naar bijv. het leren van begrippen, dan kunnen we daaruit lezen dat Gagné aanneemt dat het leren van een begrip een alles-of-niets proces is mits aan de juiste voorwaarden is voldaan. In tegenstelling tot het alles-of-niets model uit de mathematische leerpsychologie echter veronderstelt Gagné dat dit leren ook op het eerste trial, m.a.w. noodzakelijkerwijs, plaats moet vinden. Veronderstellen we dat op onze tentamens alleen begrippen, zoals door Gagné gedefinieerd, getoetst worden, dan mogen we misschien omwille van ons argument veronderstellen dat het aantal te leren begrippen bekend tel en dat voor het leren van ieder van die begrippen de student eenzelfde hoeveelheid tijd nodig heeft. De tijd die gaat zitten in het leren van een begrip is de voorbereidingstijd die nodig is, omdat het leren zelf immers onmiddellijk plaats vindt. Ik heb zo'n vermoeden dat we hiermee toch niet in staat zullen zijn het model dat Van Naerssen opstelt althans wat zijn mathematische vorm betreft, te handhaven. We zullen daar straks weer op terugkomen.

Maar waarschijnlijk doet het er allemaal heel erg weinig toe, omdat de eigenschap waarvan Van Naerssen gebruik maakt, niet uniek is voor het allesofniets leren model. Immers, een groot aantal trials worden gemiddeld, over studenten, maar ook per student. In beide gevallen zijn de te verwachten leercurven consistent met zowel een mogelijk allesofniets leren model, als het oudere incremental learning model. De vraag is of dit betekent dat Van Naerssen dan ook maar zonder meer het mathematisch model voor allesofniets leren mag hanteren.

We moeten nu nog bekijken het verband tussen studeertijd en geleerde vaardigheid, het verband tussen geleerde vaardigheid en ware score op de toets, en het verband van een en ander met het gestipuleerde leermodel. Niet gering, zou je zeggen, maar toch besteedt Van Naerssen er in zijn openbare les geen verdere aandacht aan, noch verwijst hij via het notenapparaat naar enige literatuur waarop deze aannamen enigszins gebaseerd zouden kunnen zijn.

Laten we eens proberen om met behulp van Gronbach en Gleser enige helderheid in deze ongespecificeerde verbanden te brengen.

Cr & Gl 1965: 27 :

"Under any (.....)

In any decision problem, we are actually concerned chiefly with a subset of the various s's, those which are related to payoff. For any set of treatments under consideration, the test scores may be so factored as to extract those dimensions that account for variance in payoff under at least one treatment. Payoff under any treatment depends on one or more of these aptitude dimensions. Any other of the test-score dimensions will have zero coefficients in payoff functions for all the treatments (vgl local independence, b.w.)." [twee alinea's, voor de eerste slecht gescande en hier weggelaten alinea zie geciteerde boek]

De vraag die Van Naerssen stelt is welke 'treatment' de student zich zelf het beste kan geven, of met welke treatment hij zijn studietijd kan minimaliseren. De benodigde studietijd moet dan ook in een of andere relatie tot de payoff van die treatment staan: de treatment met de hoogste payoff wordt gekozen, d.w.z. de treatment die de minste studietijd vergt. Het merkwaardige aan deze redenering is dat benodigde studietijd hier als afhankelijke variabele optreedt, terwijl Van Naerssen het doet voorkomen dat de student moet beginnen met zijn studietijd te manipuleren, d.w.z. zo te kiezen dat hij op zo efficiënt mogelijke wijze, d.w.z. in zo weinig mogelijk (bestede) studietijd, door het hele examenprogramma rolt. Het lijkt dan ook juister om te zeggen dat de payoff een functie is van de bestede studietijd voor het behalen van een afzonderlijk tentamen (voorlopig zijn we dus nog niet toe aan bespreking van het probleem van combinatie van verschillende tentamens). De payoff is maximaal wanneer er geen tijd teveel besteed is aan de voorbereiding voor het tentamen; op de vraag wat teveel bestede tijd is komt Van Naerssen nog te spreken, we zullen daar op wachten. Om onduidelijkheid te vermijden: wanneer de student in eerste instantie zakt voor het tentamen, en slaagt bij de eerste herhaling, dan is op zich niet van belang dat hij de eerste keer gezakt is, maar wordt wel zijn benodigde studietijd bepaald door de tijd besteed aan de voorbereiding op het eerste tentamen, op te tellen bij de tijd besteed aan de voorbereiding op de herkansing.

Een aantal zaken blijven voorlopig nog wat in het duister. Laten we eerst maar weer teruggaan naar het geciteerde gedeelte uit Cronbach en Gleser. Onder treatment zullen we verstaan de tijd die de student besteedt aan de voorbereiding op het tentamen. Alleen op basis van die tijdbesteding is het toetsresultaat niet voorspelbaar, omdat ook een aantal andere eigenschappen van de student daarop van invloed zijn. Van Naerssen (1968) spreekt tentamens die "die combinatie van logisch redeneren, verbaliteit, ijver, efficiëntie, etcetera, die nu eenmaal nodig is om met succes te studeren" meten.

We zullen de terminologie van Van Naerssen wat moeten zuiveren. Hij spreekt over 'geleerde vaardigheid' en 'ware score op de toets'; die zijn beide abstracte grootheden die als zodanig niet waarneembaar zijn. De vraag is hoe ze te relateren zijn aan wel waarneembare grootheden, zoals bijv. scores op de toets, en bestede tijd (zoals bv. te meten in een tijdbestedingsonderzoek). Een verband tussen deze studeertijd en toetsscores zal ongetwijfeld vastgesteld kunnen worden, al zal het ook aanzienlijk vertroebeld worden door individuele verschillen op bovengenoemde 'characteristics' (waarvoor we kunnen corrigeren wanneer daarover gegevens te verkrijgen zouden zijn).

Hoe het bovenstaande gebruikt kan worden bij kritische analyse van het model van Van Naerssen is nog niet geheel duidelijk, daarvoor moeten we eerst maar eens verder gaan met de bespreking van zijn openbare les.

Van Naerssen spreekt over het berekenen van de kans dat een bepaald 'leeritem' in het geheugen wordt opgenomen (in welk geval het ook onthouden wordt en gereproduceerd kan worden tot in ieder geval het tijdstip van het tentamen; een eenmaal geleerd item blijft dus geleerd, dat is het allesofniets leren model a.h.w.).

Het is echter niet nodig om te spreken over de kans dat een bepaald leeritem in het geheugen wordt opgenomen, het gaat immers om het deel van de stof dat op een bepaald moment geleerd is, gereproduceerd kan worden. Dat deel van de stof kan voor het gemak uitgedrukt worden in het percentage van de 'leeritems' dat gekend wordt. We hebben dus niet de ingewikkelde veronderstelling nodig dat een eenmaal geleerd item ook in het geheugen blijft en op tentamen gereproduceerd kan worden: het mag best zo zijn dat dat item nog voor het tentamen weer vergeten wordt, mits het maar vervangen wordt door tenminste een ander leeritem dat dan gekend wordt.

Het gemiddeld aantal leeritems dat de student na n bestuderingen kent, noemt Van Naerssen zijn relatieve ware score:

"Want we eisen dat de toets is gebaseerd op een representatieve steekproef van de verzameling leeritems. Kennisniveau en ware score zijn dus identiek." Aan de laatste identiteit hebben we weinig, zoals al eerder betoogd; het zijn beide platonische grootheden.

"En als de n bestuderingen nu ook nog even lang duren dan hebben we het verband tussen studeertijd en ware score t ." (eerder heeft Van Naerssen verondersteld dat elk leeritem bij elke bestudering van de stof eenzelfde kans (c) heeft om in het geheugen opgenomen te worden.) Dat iedere volgende bestudering van de stof even lang duurt lijkt hoogst onwaarschijnlijk. Onder de aannamen van het alles of niets leren model zou je eerder verwachten dat de benodigde 'repetitie' tijd evenredig afneemt met het aantal al eerder geleerde leeritems (de hoeven immers niet meer gerepeteerd te worden). Maar goed, Van Naerssen's formule wordt:

$$t = 1 - (1-c)^n \quad \text{dus } n = \ln(1-t)/\ln(1-c).$$

Hoe zit die formule in elkaar ?

proportie items geleerd na 1e keer doornemen: c

„ „ bijgeleerd na 2e „ „ :
 $c(1-c)$

„ „ bijgeleerd na 3e „ „ :
 $c(1-c-c(1-c))$

etc.

De genoemde termen zijn eenvoudiger als volgt te schrijven:

c
 $c(1-c)$
 $c(1-c)^2$
 etc

De n -de term van deze serie is $c(1-c)^{n-1}$.

Onder de vooronderstellingen van het alles of niets leren model hoeft een eenmaal geleerd leerstofitem niet meer gerepeteerd te worden, kost het doornemen van een

enkel item altijd evenveel tijd (we zullen de tijd nodig voor het doornemen van een enkel item z noemen) ongeacht het aantal voorgaande doornemingen waarbij het leerstofitem niet geleerd is.

Als de tijd nodig voor het de eerste keer doornemen van alle leerstofitems Z is, dan kost het doornemen van de leerstofitems die de eerste keer niet geleerd zijn, $Z(1-c)$.

Nu is de n de term van de somreeks, dat is het aantal leerstofitems dat gekend wordt na n keer doornemen van de stof, gelijk aan

$$1 - (1-c)^n$$

De tijd die voor het $n+1$ keer doornemen van de stof proportioneel aan het aantal nog niet geleerde items

$$1 - (1 - (1-c)^n) = (1-c)^n$$

Daaruit volgt dat de tijd, nodig voor het de n de keer doornemen van de leerstofitems, onder het allesofniets leren model is

$$Z (1-c)^{n-1}$$

De tijd die nodig is om $1 - (1-c)^n$ deel van de leerstofitems te leren is dan

$$\sum Z (1-c)^{n-1} \quad n = 1, 2, 3, \dots, \text{sommering over } j = 1 \text{ tot } n.$$

Het aardige van deze somreeks is dat hij convergeert naar Z/c .

$$\begin{array}{l} \text{zij } S_n = Z + Z(1-c) + Z(1-c)^2 + \dots + Z(1-c)^{n-1} \\ (1-c) S_n = Z(1-c) + Z(1-c)^2 + \dots + Z(1-c)^{n-1} + Z(1-c)^n \\ \hline c S_n = Z - Z(1-c)^n \end{array}$$

waaruit volgt $S_n = Z(1 - (1-c)^n)/c$

Omdat $|1-c| \leq 1$, is $\lim_{n \rightarrow \infty} (1-c)^n = 0$ (voor n naar oneindig)

Waaruit volgt dat $\lim S_n = \lim (1 - (1-c)^n)/c = Z/c$.

Het verband tussen de totaal bestede tijd en de ware score t is:

$$S_n = tZ/c \quad (\text{zoals volgt uit (1) en (6)})$$

(8)

De studeertijd is dan niet n , zoals Van Naerssen veronderstelt, maar S_n .

Aftestgrens

Van Naerssen zegt over het bepalen van de aftestgrens in dit verband niets. Hij veronderstelt eenvoudig dat de docent eist dat een bepaalde proportie van de leerstof items gekend wordt (een proportie op de schaal van 0 tot 1 van 'relatieve ware scores', wat voorlopig nog open laat wat de relatie van deze aftestgrens is tot de geobserveerde scores),

Zijdelings wordt de problematiek van relatieve en absolute normen gesignaleerd, maar in haar consequenties voor het tentamenmodel althans op deze plaats niet verder uitgewerkt.

Voordat nu ons probleem van de optimale studiestrategie voor de student geattaqueerd kan worden, moet inderdaad de aan de toetsing's procedure inherente meetfout (bij Van Naerssen per definitie het verschil tussen geobserveerde en ware scores in het model geïntroduceerd worden. We volgen Van Naerssen eerst maar eens.

Van N. neemt aan "dat voor een gegeven ware score de geobserveerde scores normaal verdeeld zijn om de ware score heen." Laten we dit iets preciezer formuleren dan Van Naerssen hier doet. Laten we de ware score definiëren als de verwachte score over replicaties en verschillende itemcombinaties in de test (Lord & Novick 7.6.1), en laten we deze t noemen.

Voor een random gekozen student:

(Lord & Novick 11.9.1)

"Thus in repeated testing with different random samples of n items each, the frequency distribution of the observed score of examinee a will be

$$f(x_a) = \binom{n}{x_a} (t_a)^{x_a} (1-t_a)^{n-x_a} \quad (9)$$

This is an ordinary binomial distribution with probability of success t_a ."

Dat hier de geobserveerde scores normaal verdeeld zijn om de ware score heen, is niet het geval, zoals bijv. te zien in Lord & Novick fig 23.2.1.

Daarom ook stelt van Naerssen

"dat, gegeven een bepaalde ware score t , de variantie van de normaal verdeelde fouten gelijk is aan die van de binomiale verdeling, hetgeen een voor de berekeningen handig compromis is tussen klassiek en binomiaal foutenmodel. Dus, voor een gegeven t , is de foutenvariantie

$$se^2 = t(1-t)/k \quad k = \text{aantal items.} \quad (10)$$

Delen door k , en niet vermenigvuldigen met k , omdat we met 'relatieve' scores werken (ware scores van 0 tot 1)."

Nu schijnt de aanname van Van Naerssen over de verdeling van waargenomen

scores rond ware score niet de formulering van Lord & Novick 11.9.1 te zijn (zoals boven geciteerd), maar Lord & Novick assumption 23.2.1

"The conditional distribution of observed score X for given (proportion-correct) true score $t=t/k$ is the binomial distribution

$$h(x|t) = \binom{k}{x} t^x (1-t)^{k-x}, \quad x = 0, 1, 2, \dots, k$$

(11)

$k = \text{aantal items.}$ "

En deze verdeling is exact gelijk aan Lord & Novick 11.9.1. , onze (9).

Over de t in assumptie (11) zeggen Lord & Novick (blz 510):

"It should not be assumed that this t has all the characteristics of the true score in the itemsampling model (d.i. formule (9), nor of the generic true score of chapters 8 and 9. Actually, (11) is the only assumption necessary for the developments here relating to a single test form; the term 'true score' need not be defined here other than by (11)."

(symbolen en formuleverwijzingen aangepast, b.w.)"

Hoe komt van Naerssen aan de formule voor de foutenvariantie (10)'? In formule (9), evenals in formule (11), heeft x betrekking op het aantal goedgemaakte items. De standaarddeviatie van x is volgens de bekende formule voor de standaarddeviatie van een binomiaalverdeling, gelijk aan de wortel uit $kt(1-t)$.

Per definitie is $e=xkt$. Omdat kt een constante is, is de verdeling van e gelijk aan de verdeling van x (maar met een ander gemiddelde omdat van alle scores kt afgetrokken wordt), en in het bijzonder is de standaarddeviatie van e ook gelijk aan de wortel uit $kt(1t)$. Van Naerssen werkt echter bij voorkeur op de schaal van 0 tot 1, i.p.v. op de schaal van 0 tot k , zodat we de verdeling moeten transformeren. Lord & Novick blz 198:

"The standard error of measurement is a dimensional quality, that is expressed in the same units as the observed score. If the observed score is multiplied by a constant, the standard error of measurement is automatically multiplied by the same constant." Zodat:

$$se = \sqrt{(kt(1-t))/k} \quad \text{ofwel} \quad se^2 = t(1-t)/k$$

(12)

Slaagkans.

"Uit aftestgrens, ware score en standaardafwijking van de fouten kunnen we de slaagkans p bepalen, bijvoorbeeld met behulp van de normaal tabel." Voorlopig is dit ook nog louter theorie, omdat de ware score een onbekende is. Wat dat betreft zou het aandiger zijn om de omgekeerde weg te bewandelen, en de regressie van de ware score op de waargenomen score te bepalen. ;maar goed, waarschijnlijk zal Van Naerssen in één van zijn volgende artikelen op dit probleem nog wel ingaan. Het probleem is erg belangrijk omdat de standaardmeetfout voor verschillende waarden van t anders ligt: voor $t=0,2$ is $se = 0,04$ terwijl voor $t=0,5$ $se= 0,05$ telkens

bij $k=100$. Kiezen we $k=20$ dan vinden we voor $t=0,2$ $se=0,28$ en voor $t=0,5$ vinden we $se=0,35$. Daarnaast is het evident dat wil het model enige praktische bruikbaarheid hebben, het toegepast moet kunnen worden op waarneembare scores. Er doen zich hier vergelijkbare problemen voor als bij correcties voor raden bij meerkeuze toetsen, zie bijv. Van de Ven in Ned T vd Psychol sept 1974, en Lord & Novick.

Van Naerssen gebruikt voor de slaagkans p het logistische model, dat een goede plaatsvervanger is voor het matnematisch veel onhandiger model van de normaalverdeling, zonder dat de beide verdelingen veel van elkaar afwijken (d.w.z. de logistische kromme is voor de parameter waarde $1,7=D$ bijna gelijk aan de cumulatieve normaalverdeling).

$$p = 1/(1+e^{-1,7z}) \quad z=(ta)/se \quad (13)$$

$a = \text{aftestgrens}$

(zie bijv. Lord & Novick blz 399)

Van Naerssen: "met die slaagkans kunnen we de verwachting berekenen van de totale studeertijd, en deze verwachting kunnen we minimaliseren.

1 cesuur van lq. 12

4

specificiteitseis.

Van Naerssen: "Jouaten we nu eens kijken hoe de kennis, dus de ware score, in de tijd varieert, bij eer student, d',e de optimale strategie volgt. Jie kennis begint bij het punt nul, dat wil zeggen, de tentamenvragen, moeten zo geconstrueerd worden dat iemand, die niet studeert, gemiddeld een, eventueel voor raden gecorrigeerde, score nul krijgt. Dit is, tussen haakjes, de bekende specificiteitseis, die aan studietoetsen gesteld wordt (zie bijv. ~.L..Ü1'bel, Measuring educational achieveinent, 1969, P. 297 e.v.). Je toets moet meten wat onderwezen werd en niets meer.'1

~)it is natuurli jk ~,een onschuldig uitstapje van Van il~aerssen: het is een expliciete veronderstelling in het door hem uitgewerkte model'. Vragen die je op dit punt kunt stellen zijn: wordt er in de praktijk voldaan aan de specificiteitseis? 1b'en hoe kun ie dat nagaan, wat is ra.a.vi. 1,iezgeen een studietoets meet ? in zijn bijdrage aan het onderwijscongres 1968 spreekt Van Naerssen over tentamens die "essentieel dezelfde factor3 meten: die combinatie van logisch redeneren, verbaliteit, ijver, effici~ntie etcetera, die nu eenmaal nodig is om met ;ucces te studeren~'. "Iet lijkt er veel op dat de eis van specificiteit, en de ei~renschap van metingen die 'tentamens' genoemd worden in de formulering van Jan Aaerssen 1968, zeer moeilijk eenvoudig met elkaar te verenigen zijn. Voor het ogenblik zullen we deze problematiek niet verder aan kunnen snijden, maar het' is Jui~elrijk dat we er niet aan kunnen ontkomen oIA haar te elef~enertijd degelijk onder de loeo te nemen.

Verwachting van de totale studeertijd.

Als Q de slaagkans is, kunnen we $q=1/p$ de zakkans noemen. De verwachting van de totale studeertijd is gelijk aan de som van de oneindige meetkundige reeks van de zakkans q , hoewel deze formulering van Van Raerssen me niet helemaal juist voorkomt. Tussen herkansingen "daalt de kennis" tot een bepaalde fractie van de eerst bereikte optimale kennis. Die fractie kunnen we als parameter in onze formules opnemen, onder de naam onthoudfractie r . Voor de herkansing geldt "de leerperiode is echter korter dan de eerste, want hij heeft nog, wat onthouden. De voorin van de leercurve is dezelfde als eerst, en ook de hoogte van de kennis waarnaar hij moet streven. Je situatie is vlak voor dit tweede tentamen precies gelijk aan die vlak voor het eerste. Wat toen optimaal was is het ook nu. Weer kan de student slagen of zakken met dezelfde kansen als eerst.;"

Laat het kennisniveau daalt steeds van de optimale waarde t tot rt , en de studeertijd van nul tot mt heet n_1 . en die van mt tot t heet n_2

dan is de verwachting van de totale studeertijd

$$n_1 + n_2 + n_2q + n_2q^2 + n_2q^3 + \dots$$

$$n_1 + n_2/p = E(n) \quad (14)$$

$$\text{Nu is } n_1 = \ln(1-rt)/\ln(1-c) \quad (15) \quad (\text{volgens (2)})$$

De afleiding van formule (14):
de termen uit de somreeks kunnen geschreven worden als
 $n_1 + n_2 + n_2q + n_2q^2 + \dots$

$$= (q^j)_{j=0}^{\infty}$$

uitschrijven levert ons uiteindelijk op dat de verwachting daarvan is

$$n_1 + n_2 \sum_{j=0}^{\infty} (q/p)^j$$

waarvan de laatste term wegvalt omdat a gelijk is aan nul voor j naar oneindig.

cesuur van N. 13

t

$$n_1 = \ln(1-rt)/\ln(1-c) \quad \text{volgens (2)}$$

$$n_2 = \ln(1-rt)/\ln(1-c) - \ln(1-rt)/\ln(1-c)$$

ook uit (2) en $n^2 = n \cdot n$

$$p = 1/(1 + e^{-z}) \quad z = (t - a)/s \quad (13)$$

$$E(n) = \ln(1 + mt)/\ln(1 + c) + \ln(1 + t)\ln(1 + mt) \quad (14)$$

$$\ln(1 + e^{-z}) \quad (17)$$

Of dit dezelfde uitkomst is als Van Naerssen geeft, kan ik zo snel niet controleren, de tekst van Van Naerssen is niet eenduidig. (noot 16) Verwachte studeertijd en gekozen kennisniveau

Als het kennisniveau $t=0$, is $E(n)=0$ k/beide termen in (17) worden nul omdat er $\ln 1$ in de xcm.« teller voorkomt (en $\ln 1=0$ uiteraard).

Beschrijving van het verband tussen t en $L(n)$ zullen we uitstellen tot de bespreking van een van Van Naerssen's latere artikelen, waarin curven gegeven worden. ~Joveel kzaan eeliter wel geze"~u worden, aat er ergens een optimaal kennisniveau t bestaat waarvoor $L(n)$ minimaal is (zij het dan ook dat net een relatief minimum is, omdat een absoluut minimum voor $t=0$ immers bereikt wordt bij $t=0$)

Voor het optimale kennisniveau:

IM

Afstellen van het differentiaalquotient van $L(n)$ naar t levert de vergelijking op

$$IM \cdot (\ln(1 + mt)) / (\ln(1 + c)) = (a + 1/2) + e^{-z} \quad (15)$$

$$1 + mt = 2t$$

$$7k$$

waarin $y=1,7$

Gegeven a , m en k , kan men t hiermee iteratief benaderen,"

Het nemen van deze afgeleide is een rechttoe-rechtaan klus, die echter wel enige tijd in beslag neemt. Als ik nog eens een half uurtje over heb zal ik het narekenen).

cesuur van N. 14

Daarna berekent men de bijbehorende minimale verwachte studeertijd. Deze kan men in verband met de typische vorm van de leercurve het beste uitdrukken in zogenaamde halveerperioden. (Dit betekent dat in de formules $c=1$

2 wordt gesubstitueerd) De halveerperiode is onze eenheid van tijd, en U vermoedt het al op het woord afgaande: een halveerperiode is de studeertijd, die nodig is om de helft van de nog niet onthouden xstof in het hoofd te krijgen. Dus de eerste 50% kost 1 periode, de volgende 25% weer 1, enzovoort."

Wat hebben we nu. "We kunnen onder de aannamen van het model de minimale verwachte studeertijd berekenen, en de student kan daar zijn studiestrategie naar richten. Wat nu, als de student liever wat harder werkt om er zeker van te zijn dat hij meteen de eerste keer slaagt? In dat geval zou het model aangepast moeten worden, omdat het de student niet onverschillig is of hij extra studeertijd op dit moment er in steekt, of op een later (na gezakt te zijn) moment. De utiliteit van een en ander moet in de formules ingevoerd worden als vervanging voor de impliciete weging van studeertijd met een factor 1 voor zowel eerste tentamen als herhalingen. Goed, hier zullen we later nog eens over na gaan denken, van dit soort verfijningen o") het model zijn er immers vele te bedenken, en het risico bestaat dat door al dit soort variabelen in te voeren dat het model onoverzichtelijk wordt, en minder algemeen toepasbaar dan in zijn meer eenvoudige vormen.

In het voorgaande hebben we al gewezen op de onverenigbaarheid van een allesofniets leren model met de aanname van Van Naerssen dat iedere 'halveerperiode' evenveel studietijd zal kosten. Het zou interessant zijn om de alternatieve tijdmaat in rekenkundige voorbeelden uit te werken. Ook daarmee echter zullen we wachten tot de bespreking van de numerieke voorbeelden die Van Naerssen in een later artikel 1, heeft.

"Laten we nu de aftestgrens constant houden op de helft van het aantal vragen, en kijken hoe de minimale verwachte studeertijd volgens de berekeningen afhangt van het aantal items en komt van de onthoudfractie. Met het gekozen kennisniveau, dus met de strategie van de student hebben we niet meer te maken; die is weggewerkt doordat we veronderstellen dat hij de optimale strategie volgt."

Deze uitspraak van Van Naerssen is enigszins onnauwkeurig: immers, de verwachte studeertijd bij een optimale strategie, en bij gegeven a , k en m , wordt uitgedrukt in t : het te kiezen niveau van beheersing van de stof. Je hebben dus alles met t te maken, omdat t juist varieert afhankelijk van de gekozen waarden voor a , m en k . De berekeningen van Van Naerssen worden hier niet door beïnvloed:

Wel, we zien dan dat die totale studeertijd $xfckxxgkxxn$ weinig afhangt van het aantal items, zolang dit tenminste groter is dan het gebruikelijke minimum van 30 !)us wat de verwachte studeertijd betreft is er nauwelijks reden om de betrouwbaarheid te verhogen door de toetslengte te verdubbelen van 50 tot 100 items. Als bijvoorbeeld bij elk volgend tentamen no 505J onthouden is, dan zou de verwachte tijd door de testverlenging slechts dalen van 1,35 eenheden naar 1,27 eenheden. Beneden de 20 items maakt heil echter wel wat uit, vooral als de onthoudfractie klein is. En merkwaardig genoeg, als deze erg groot is, 90', ó', dan is een korte toets juist voordelig voor de lijntrekker, maar dat komt natuurlijk vanwege de dan relatief grote invloed van geluk.

ED

Het door Van Naerssen gegeven voorbeeld is duidelijk niet relevant voor de

gebruikelijke onderwijssituatie. Hij komt tot studeertijden van bijv. 1,5, laten we zeggen 1~ keer doornemen van de stof. _~1'r zullen

C~ 2

wel studenten zijn die deze strategie volgen, maar re~ler li." :kt toch te veronderstellen dat de stof een groot aantal malen doorgenomen moet worden ~waarbij dan niet iedere keer evenveel tijd nodi, is, maar minder naarmate de studie vordert). ue berekeningen van Van 1~aerssen zullen dus minstens aangevuld moeten worden met gegevens over wat er gebeurt wanneer de aftestgrens a aanzienlijk hoger gesteld wordt.

CD

cesuur van N.15

Strategie van de docent,.

De docent die graag denkt in absolute termen over het niveau van beheersing van de stof dat hij van zijn studenten verlangt, zou dit model kunnen gebruiken. ~ls zijn onderwijssituatie zich niet leent tot het 'gebruiken van dit model, dan kunnen studenten het ook niet gebruiken en is het hele model zinloos, behalve in zoverre dat de conclusie dat dit soort modellen niet bruikbaar is in die en die onderwijssituatie natuurlijk wel belangrijk is. ~',.;ie de uitspraken daarover geëaan door Van Naerssen in zijn artikel (NedTvdPs sept 1974)

De docent kan specificeren welke relatieve ware score hij verlangt van zijn studenten. Daarvan uitgaande, en van de veronderstelling dat zijn studenten de optimale studeerstrategie volgen, kan hij òerekenen welke afLtestgrens hij moet kiezen, bij toetsen met een bepaald aantal items. Het geheel wordt natuurlijk beïnvloed door de empirisch te bepalen onthoudfractie m. In de meeste omstandigheden zal de aftest grens lager gekozen moezen worden dan de relatieve ware score. Het model laat echter onmiddellijk zien, wat natuurlijk ook wel beredeneerd zou kunnen worden. dat er omstandigheden zijn waarin de aftestgrens noger gesteld moet worden. ("Alleen bij een combinatie van weinig items en een hoge onthoudfractie moet de aftestgrens hoger gesteld worden.")

Wanneer men denkt het modeL in zijn onderwijssituatie toe te kunnen passen, dan kunnen diverse beleidsmogelijkheden er uit afgeleid worden

li,~o kan men aan grafieken zien dat men het aantal items van een bepaald tentamen erust van 90 tot 50 terug kan brengen, zonder dat dit gemiddeld meer studeertijd zou kosten, mits men tegelijk het aantal jentawen,7e1e,enheden per jaar zodanig !aat. toenemen, dat de on'Cho~.Ldfr~~ctie stijgt van bi,ivoorbeeld 0,5 tot 0,7.i~

Omgekeerd zou je kunnen kijken wat er gebeurt wanneer herkansingen zeer snel vol- 11

gen op de eer.,e tentamengeleenheid, zodat de onthoudfractie jiána Selijk 1 wordt (en onder de veronderstellingen van alles of niets leren model ' Telijk 1 moet zijn). i".oeveel items moet een tentamen dan bevatten om er zeker van te zijn dat studenten die niet de optimale strategie volgen maar een kleiner deel van de stof 7,eleerd hebben, onderscheiden kunnen worden van degenen die tenminste de optimale strategie gevolgd hebben. jit lijkt dan al aardig op een probleem dat past in het besliskundig, model.

!ju ir: er zo over denk, is het opvallend dat Van l'jaerssen naar ik meen nertens in zijn artikelen s ireekt over de situatie waar studenten ~iet de Optimale strategie volen, maar niet enoe,~_ gestudeerd hebben om je door de docent gewenste relatiegve ware score te bereiken. uiteraard hebben deze studenten ook kansen om te slagen voor de tentamens, zij het dat hun kansen kleiner zullen zijn dan. van degenen die de optimale studiestrategie gevolgd hebben, juist omdat zij die optimale strategie niet volgen, zou misschien een redenering op te zetten zijn da"t, onder de aannamen van net model, en acties van de docent die daarop gebaseerd zijn, geen rekening ",e~Louden hoeft, te worden met de foute beslissingen die genornen worden wanneer studenten worden toegelaten die niet tot aan de gevraagde relatieve ware score gestudeerd hebben? Dat lijkt me niet het geval, omdat dat zou betekenen dat de zakslaag beslissingsstrategie van die docent zeer conservatief zou zijn, in het nadeel van de studenten dus, en in het nadeel van het onderwijs. Van Naerssen zal daarom zegen dat een en ander afhangt van de herkansingsmogelijkheden, die immers ook al in zijn model zitten, en dat de zaak daarmee dan toch prettig geregeld is. Het is niet onmogelijk, maar het is duidelijk dat we een en ander toch zwart op wit moeten zien te krijgen. Een leuk projectje dus.

cesuur van W. 16

n

combineren van tentamens.

'ibij de compensaatorische methode telt wen de tentamenuitslagen al of niet gewogen bij elkaar oP, terwijl de student bij de con
a
junctieve methode voor alle tentamens apart geslaagd moet zijn.;~

Het is niet duidelijk wat Van Naerssen bedoelt met ootellen van tenáamenuitslagen. Hij maakt nergens, ook niet in zi jn 1968 artikel, restricties in de zin dat ook bij een compensatorische methode bepaalde minimum prestaties tenminste geleverd moeten zijn. i)at zou voorlopig moeten betekenen dat wanneer van 14aerssen over een connensatorisch model spreekt, hij ieder cijfer in Drincipe compenseerbaar vindt door hogere cijfers elders. 1) Verder wordt ooenelaten hoe verschillende vakken afzonderlijk gewogen zouden moeten worden, of dat zelfs weging in de meeste onderwijssituaties geen zin zal hebben (wat niet zo onwaarschijnlijk is). Het laatste zou aan de hand van de literatuur nog wat verhelderól kunnen worden. Ook opengelaten wordt het schaalprobleem, hoewel dat in zekere zin nauw verwant is aan het wegingsprobleem (een vak met een grotere spreiding in cijfers wordt in feite

ook zwaarder meegewogen). Je zou kunnen denken aan een standaardisering van scores voordat de scores gecombineerd worden. In ieder geval speelt voor de afzonderlijke vakken niet het cesuurprobleem, omdat de cesuur pas gelegd wordt bij het examen, in casu de combinatie van afzonderlijke vakken. Het laatste houdt nog de mogelijkheid open dat lage cijfers door de student ooit nog in hogere omgezet kunnen worden door het tentamen nog eens te herhalen (wat onder de aannamen in het model van Van Naerssen ook vrij zeker betekent dat die student dan inderdaad de nodige aandacht aan de inhoud van het, vak moet besteden om die hogere waardering te kunnen bereiken). In verband met het laatste punt wordt onverklaarbaar dat Van Naerssen op enkele plaatsen flirt met de gedachte (veronderstelling) dat een compensatorische methode ook zou betekenen dat de 'enen die uiteindelijk zakken, het benodigde aantal punten niet halen, dan ook alle vakken over zouden moeten doen. Immers, ook onder het compensatorische model met de mogelijkheid om lage scores te verbeteren door herhaling van het tentamen, is het altijd mogelijk om te slagen voor de exameneisen (zoals ook voor het conjunctieve model zonder tijdbeperking geldt)

Van Naerssen 1968: "Bij de compensatorische methode is het theoretisch denkbaar dat een student van een of meer vakken geheel niets weet, maar hij zal dit tenminste moeten compenseren met goede prestaties in andere vakken."

cesuur van A. 17

Ook tentamens die verschillende zaken meten laten compenseren. ?

Door de instemmende verwijzing van Van Naerssen naar het artikel van Lord uit 1962 over cuttingscores ontstaat er toch enige verwarring. Ook in 1968 verwees hij naar dat artikel. Toch lezen we in zijn publicatie destijds: (over de compensatorische methode)

Mij lijkt vooral van belang dat de tentamens 1) in de tijd dicht bijeen liggen, en 2) essentieel dezelfde 'factor' meten: die combinatie van logisch redeneren, verbaliteit, ijver, efficiëntie, etcetera, die nu eenmaal nodig is om met succes te studeren. Het is namelijk denkbaar dat de groep vakken bestaat uit subgroepen, die elk hun eigen factor meten. Bijvoorbeeld een subgroep 'wiskundige' vakken tegenover een subgroep 'verbale' vakken. In dit geval is de compensatorische methode alleen te verdedigen binnen een subgroep. Immers, als zowel de ene als de andere factor belangrijk is voor de studie dan mag de ene score de andere niet compenseren. Ien moet dan wel conjunctief selecteren en het betrouwbaarheidsverlies goedmaken met veel langere toetsen."

Waarschijnlijk heeft Van Naerssen gelijk wanneer volledige compensatie bedoeld wordt. Maar het artikel van Lord heeft nu juist tot resultaat dat precies in die situaties waarin Van Naerssen dan zijn voorkeur uitspreekt voor de conjunctieve methode, een milde vorm van compensatie altijd de voorkeur verdient. Lord gaat bij zijn bespreking van twee tests juist uit van twee tests die onderling niet gecorreleerd zijn, en bovendien van de veronderstelling dat t.a.v. de doeleinden van de selectieprocedure een bevredigende score op beide tests a priori door degenen die

de selectie instellen, van belang geacht wordt. Desondanks, en dat is juist het verrassende resultaat van de analyse van Lord, zoals hij zelf ook nog eens uitdrukkelijk stelt, is altijd een milde vorm van compensatie te verkiezen boven een puur conjunctief model (en is een zuiver compensatie model ook niet optimaal, hoewel je je nog af kunt vragen of het verschil tussen een zuiver compensatie model en de compensatie volgens de optimale curven van Lord in de praktische situaties die zich in ons onderwijs voordoen, wel zo belangrijk is.

De volgende uitspraak in zijn openbare les lijkt dan ook in het licht van het werk van Lord niet houdbaar te zijn:

Maar de uitzondering waar ik op doelde is de situatie waarbij de kennis die bij eerdere cursussen geleerd werd noodzakelijk een bepaald peil moet hebben bereikt opdat de student de latere cursus met vrucht zal kunnen volgen. In deze situatie zal men soms conjunctief moeten toetsen, om de studenten te dwingen om iets te leren beheersen alvorens verder te gaan, een principe dat met veel succes gebruikt wordt bij geprogrammeerde instructies."

Wat Van Naerssen wél kan zeggen is dat in sommige situaties een volledige compensatie teruggebracht moet worden tot een milde compensatie.

Overigens snijdt de laatstgeciteerde uitspraak wederom een normenprobleem aan: wat is 'beheersing'?

cesuur van N. 13

Compenseren geeft betrouwbaarder beslissingen.

In zijn openbare les zegt Van Naerssen hier vrijwel niets over; we zullen daarom terug gaan naar zijn bijdrage aan het onderwijsresearch congres 1968.

Van Naerssen berekent een index h , deze is de verhouding tussen een bepaalde testlengte bij de compensatorische methode en de testlengte, die nodig is bij de conjunctieve methode om dezelfde betrouwbaarheid te verkrijgen.' "Een $h = 0,5$ betekent bijvoorbeeld dat de conjunctieve methode tot een betrouwbaarheid leidt, die men bij de compensatorische methode al zou bereiken bij de helft van het aantal vragen.

Van Naerssen geeft een aantal waarden van h bij verschillende gemiddelde intercorrelaties tussen de tentamens, en slaat, i5ereenllages

van resr. 90 en 80 voor ieder tentamen. Het aantal delen (vakken) speelt hierbij geen rol.

R 90~ó 8oglo

.20	$h =$	.23	$h =$	.40
~30		.23		39
.40		.22		37

.50	.22	35
.60	.21	33
70	.20	JO
.80	.18	.26
.90	.15	.22

~.Je gaan nu proberen om wat duidelijkheid in deze betrouwbaarheidsproblematiek te brengen door Van Naerssen's berekeningswijze en zijn aannamen daarbij te reconstrueren.

Allereerst veronderstelt hij bij de getallen uit de gegeven tabel dat de conjunctieve methode zonder herkansing, voor afzonderlijke tentamens wordt toegepast (of, wat op hetzelfde neerkomt, dat er bijv. wel een herkansing gegeven wordt, en dat de slaagpercentages dan betrekking hebben op tentamen + herkansing).

Ook zullen we er van uit gaan dat Van Naerssen t.a.v. de compensatorische methode ook veronderstelt dat er geen herkansingen zijn, en tevens dat cijfers van de vakken zonder naar bij elkaar worden opgeteld en gedeeld door het aantal vakken om het totaal examen cijfer te bepalen.

door Van Naerssen wel genoemd zijn:

C)

11) de scores op de (K) deelttoetsen correleren onderling evenveel met elkaar (= 1), hebben dezelfde standaardafwijking en L_7, J^n

$C \sim J$

multivariaat normaal verdeeld, en

2) de conjunctieve methode kan men opvatten als een χ^2 -toets, waarbij de afgestegrens ligt tussen de 'scores' van en 1,.

In beide gevallen kan men de betrouwbaarheidscoëfficiënt van de totale score berekenen door K en de onderlinge correlatie in te vullen in de Spearman-Brown formule. Invlaar bij de compensatorische methode moet men de tetrachorische correlatie gebruiken, die een schatting is van R, en bij de conjunctieve methode de 'phi', die altijd kleiner is.

Lord en Novick: (0.0) if $Y_1, Y_2, \dots, Y_n$ are n parallel measurements, then we may obtain

$r_{L7} = 9$

$1 + (n-1)r_{Y1}$

This result is known as the (general) Spearman Brown formula for the reliability of a composite test having parallel components, it

T

cesuur van 1.,. 19

Hier blijkt Van Aerssen verschillende tentamens o.j., te vatten als parallel metingen, ivat inderdaad de consequentie is van de ~veelal em~)ir~scr'i öeves~li,de)' opvatting uat verschillende tentamens hetzelfde meten.

Noemen we R de betrouwbaarheid van het nele ,jal,.'±~et metingen (en dus van het examen), en r de correlatie tussen twee tentamens (d',e voor alle paren tentamens gelijk verondersteld wordt), en gaan we er va11 uit dat er n tentamens deel uit i,~laken van het examen:

$$\frac{nr}{\sqrt{n(n-1)}}r$$

Als $r = .20$, en we nemen voor $n = 5$, dan vinden we als betrouwbaarheid bij de compensatorische methode voor het examen

$5(.20)^2$

Ik zie nog niet in hoe we hiervoor een tetrachorische correlatie voor nodig zouden hebben,.

n nu de betrouwbaarheid voor dezelfde vijf toetsen conjunctief, ecorrib~'neerd, bij een slaag, ,percentage van 90 ~b. ";e zullen veronderstellen dat alle studenten alle ten'kar.,iens hebben "~,edaan (en nietL. opgehouden zijn met studeren na zakken voor een bepaald tentamen).

Het vervelende is dat ik op dit moment in het geheel niet zie hoe ik verder kan komen. liet grote struikelblok voor mij is het berip betrouwbaarheidip en wat dat in dit verband zou kunnen betekenen, met name dus t.a.v. zakslaag besl~ssingen. Je zou kunnen denken aan de mate van overeenstemming bij zak slaag beslissingen Irenomen op basis van xmx twee aan elkaar parallelle series ulentainens, :jrgens

schijnt dat ook de idee van Van 1~'aerssen te zijn, anders be~7rijp
t

ik niet hoe hij komt op corre1a ie coefficienten die betrek.King hebben op 2 bij 2 contin,;~entietaðellen.

cesuur van Naerssen

Over d^* t in aamumptie (11) zeggen Lord & Noviek (blz 510):

"It should not be assumed that this t has all the characteristics of the true score in the item sampling model (d.i. formule (9)), nor of the generic true score of chapters 8 and 9. Actually, (11) is the only assumption necessary for the developments here relating to a single test form; the term 'true score' need not be defined here other than by (11)," (symbolen en formuleverwijzingen aangepast, b.w.)

Hoe komt van Naerssen aan de formule voor de toetsvariantie (10)?

In formule (9) evenals in formule (11) heeft x betrekking op het aantal goedgemaakte items. De standaarddeviatie van x is volgens de bekende formule voor de standaarddeviatie van een binomiale verdeling gelijk aan de wortel uit $kt(1-t)$. Per definitie is kt een constante is. is de verdeling van x gelijk aan de verdeling van x (maar met een ander gemiddelde omdat van alle scores kt afgetrokken wordt), n in het bijzonder is de standaarddeviatie van x ook gelijk aan de wortel uit $kt(1-t)$. Van Naerssen werkt echter bij voorkeur op de schaal van 0 tot 1, iop.v. op de schaal van 0 tot k zodat we de verdeling moeten transformeren, Lord & Novick blz 198:

"The standard error of measurement is a dimensional quality. that in expressed in the same units as the observed score, If the observed score is multiplied by a constant, the standard error of measurement is automatically multiplied by the same constant."

Zodat:

$$s_e^2 = t(1-t)J^2/k$$

(12)

Slaagkans.

"It is often true that the standard deviation of the errors can be determined by the use of the normal table," Voorlopig is dit ook nog louter theorie omdat de ware score een onbekende is. Wat dat betreft zou het handiger zijn om de omgekeerde weg te bewandelen, d.w.z. de regressie van de ware score op de waargenomen score te bepalen. Maar goed, waarschijnlijk zal Van Naerssen in een van zijn volgende artikelen op dit probleem nog wel ingaan. Het probleem is erg belangrijk omdat de standaardfout voor verschillende waarden van t anders ligt: voor $t=0,2$ is $a = 0,04$ terwijl voor $t=0,5$ $a = 0,05$ telkens bij $k=100$. Kiezen we $k=20$ dan vinden we voor $t=0,2$ $a = 0,05$ en voor $t=0,5$ $a = 0,035$. Daarnaast is het van belang

dat het model enige praktische bruikbaarheid hebben, het toegepast moet kunnen worden op waarneembare scores. Er doen zich hier vergelijkbare problemen voor als bij correcties voor raden bij meerkeuze toetsen zie bijv. Van de Ven in NedTvdPsychol, sept 1974, en Lord & Novick.

Van Naerssen gebruikt voor de slaagkans p het logistische model dat een goede plaatsvervanger is voor het mathematisch veel onhandiger model van de

normaalverdeling ν zonder dat d^* beide verdelingen veel van elkaar afwijken
(dowoz de logistische kromme is voor d^* narameter waarde 197D bijna gelijk aan
 d^* cumulatieve normaalverdeliug^{5*}

$$p \quad V(l+4^{1,07z}) \quad Z=(t\&)/0^* \quad (13)$$

$a = \text{aftentgrons}$
(zie bijvo Lord & Novick blz 399)
Van Naerssen: "met d^* slaagkaan kunnen we
do verwachting berekenen van do totale studoortijd. en deze verwaohting kunnen
we minimaliseren"#

specificiteitseis

Van Naerssen :

"Laten we nu eens kijken hoe d^* kennis. dus de ware score. in de tijd varioort, bij
een student, die do optimale strategie volgt, Die kennis begint bij het punt nulg dat
wij zeggen, de tentamenvragca moeten zo geconstrueerd worden dat lo~udv die
niet atudeerty gemiddeld eau* eventueel voor raden gecorrigeerde# score nul krijgt*
Dit ie# tuooen haakjes, de bekende opecificiteitaciay die aan atudietoetaan gestold
wordt (zie bijvo R*L*Eboly Mcaouring oducational achi*vementg 19659 po 297 e.v.)s
De toets moet meten wat onderwezen word en niets meerole

Dit la natuurlijk geen onschuldig uitstapje van Van Naerssen: het is een expliciete
veronderstelling In het door hem uitgewerkte model& Vragen die je op dit punt kunt
stellen zijn: wordt er In de praktijk voldaan aan de opecificiteitocie? Een ho* kun je
dat nagaan. wat is maa*wo hetgeen een studietoete meet ? In zijn bijdrage aan het
onderwijscongres 1968 spreekt Van Naerssen over tentamens die "essentieel
dezbldfe ~racto14 meten: die combinatie van logisch redeneren. verbaliteit, ijver.
offlei;ntie die nu eenmaal nodig in om met succes te studeren", Het lijkt tde eis van
specificiteit, *n de eigenschap van metingen die st genoemd worden In de
formulering van Van Naerssen 1968, eenvoudig met elkaar te verenigen zijn. Voor
hot ogenblik we deze oblematiek niet verder aan kunnen snijden# maar hot
in du 'äwlijk dat * er niet aan kunnen ontkomen om haar te
gologenortijd dogel Jkbqdor do q»Wte nomeu.

Verwachtilm de totale studoortláo p do "slaagk" g ieg kunnen we qalp de zakkww
noemen, D^* verwachting van de totale atudtortijd in gelijk ~n do oom van do
oneindige meetkundige

""reeks van de zakkanip q^* hoewel deze formulering van Van Naerssen ze niet
helemaal juist voorkomt, Tussen herkansingen "daalt de kennis tot een bepaalde
fractie van,de eerst bereikte optimale kennis* Die fractie kunnen w^* als parameter in
4;me formules opnemen* onder de naam onthoudfractie a."

Voor de herkansing geldt Me leerperiode in echter korter dan de eerste, want hij
heeft nog wat onthouden, De vorm van de leercurve is dezelfde als eerst. en ook de
~oogte van de kennis waarnaar hij moet streven,

De situatie is vlak voor dit tweede tentamen pregiez_gp14jk aan die_v vgn "i.Wat;toen op~ was in het ook nu. Weer kan de student slagen of zakken mèt't d'èzé~fde "naijĩálä"e st.

'~Stel het kenninnive*u daalt steeds van de optimale waarde t tot at. en de studeertijd van ~ul tot ot heet nl, en die van at tot t heet u

2*
dan ia de verwacht~4~g van de totale studeertijd

*n, + a 2 + n2q + n 29 + n2 q + #****#= nl + u 2/P un(n)

Nu is u 1 (volgens (2»

Do afleiding van formule (14)
de t*raen uit de nomroeks kunnewgischreven worden als
+ a + u 5 ~'~ ~

1 2 2J

vAk42.rt, da

,t ent

er moeili *nvo
ob
0

k t

a* d

Uitschrijven levert,dna uiteindelijk op dat de verwachting daarvan ie

n~A2/P (42q/P)gqj 12~_
waarvan de laatste term wegvalt omdat Sqj gelijk in aan nul voor j nam

~uur van N* 13

n 1 lu(lmt)/ln(lé)
volgens (2) (15)
n 2 = ln(J.t)/ln(Jc) ln(lmt)/ln(lc) (16)
ook uit (2) en u2 * n ni
p V(l+01v7z)z(t&)/& 0 (13)
ZW1nOotWin(1c)+ 197z

Of dit dezelfde uitkomst is als Van Naerssen geeft, kan ik zo snel niet controleren# de tekst van Van Naerssen is niet eenduidig. (noot 16)

Verwachte Studietoestand bij Rekenen» kennisniveau.~t)01

Als het kennisniveau $t=0$, $\ln X(n)=0$ (beide termen in (17) worden nul omdat er $\ln$ in de teller voorkomt)(en $\ln 0$ uiteraard). Beschrijving van het verband tussen t en $b(u)$ zullen we uitstellen tot de bespreking van een van Van Naerssen's latere artikelen, waarin curves gegeven worden, Zoveel kan center wel gezegd worden# dat er ergens een optimaal kennisniveau t bestaat waarvoor $E(U)$ minimaal is (zij het dan ook dat het een relatief minimum is omdat een absoluut minimum voor $E(n)$ immers bereikt wordt bij $t=0$)

Voor het optimale kennisniveau:

Indien stellen van het differentiaalquotient van $E(n)$ naar t levert de vergelijking op

$$10 \quad a + 1/2) + 0 \quad (18)$$

waarin

Gegeven $\log m \cdot n$ kan men t hiermee iteratief benaderen," (het nemen van deze afgeleide in een rechttoerechtaan klus. die echter wel enige tijd in beslag neemt. Als ik nog eens een half uurtje over heb zal ik het narekenen)*

"Daarna berekent men d^* bijbehorende minimale verwacht* studietoestand. Deze kan men in verband met de typische vorm van de leercurve het beste uitdrukken in zogenaamde halveerperiodes (Dit betekent dat in de formules omi wordt gesubstitueerd) De halveerperiode in onze eenheid van tijd# en U vermoedt het al op het woord afgaande: een halveerperiode is de studietoestand, die nodig is om d^* helft van de nog niet onthouden M leerstof in het hoofd te krijgen. Dus de eerste 50% komt & n periode# de volgende 25% weer 44u y onzovoort*11

Wat hebben we nu, We kunnen onder de aannamen van het model de minimale verwachte studietoestand berekenen. en d^* student kan daar zijn studiestrategie naar richten" Wat nu, als d^* student liever wat harder werkt om er zeker van te zijn dat hij meteen d^* keer slaagt? In dat geval zou het model aangepast moeten worden# omdat het de student niet overtuigt of hij extra studietoestand op dit moment erin steekt# of op een later (na gezakt te zijn) moment* De utiliteit van *en en ander moet in de formules ingevoerd worden als vervanging voor de impliciete weging van studietoestand met een factor 1 voor zowel eerste tentamen & la herhalingen, Goedo hier zullen we later nog eens over na gaan denken van dit soort vertijningen op het model zijn er immers vele te bedenken. en het risico bestaat dat door al dit soort variabelen in te voeren dat het model onoverzichtelijk wordt, en minder algemeen toepasbaar dan in zijn meer eenvoudige vormen,

In het voorgaande hebben we al gewezen op de onverenigbaarheid van een allesofniet leren model met de aanname van Van Naerssen dat iedere

'halvoorperiode* evenveel studietijd zal kosten* Het zou interessant zijn om de alternatieve tijdmaat in rekenkundige voorbeelden uit te werken, Ook daarmee echter zullen we wachten tot de bespreking van de numerieke voorbeelden die Van Naerssen in een later artikel geeft*

"Laten we nu de aftentgrens constant houden op de helft van het aantal vragen. en kijken hoe de minimale verwachte studietoortijd volgens de berekeningen afhangt van het aantal items en k van de onthoudfractie Met het gekozen $k = \text{niveau}$ dus met de strategie van de student hobbu we niet meer te maken; die is weggewerkt doordat we veronderstellen dat hij de optimale strategie volgt*" Deze uitspraak van Van Naerssen in enigszins onnauwkeurig immers» de verwachte stude*rtijd bij een optimale strategie# en bij gegeven a , k en m wordt uitgedrukt in t het te kiezen niveau van beheersing van de stof* We hebben dus alles met t te maken# omdat t juist varieert afhankelijk van de gekozen waarden voor a , k en m . De berekeningen van Van Naerssen worden hier niet door beïnvloed:

"Wel# w * zien dan dat die totale studietoortijd t weinig afhangt van het aantal items. zolang dit tenminste groter is dan het gebruikelijke minimum van 30* Dus wat de verwachte studeertijd betreft is er nauwelijks reden om de betrouwbaarheid te verho#en door de toetalengte te verdubbelen van 50 tot 100 items* Als bijvoorbeeld bij elk volgend tentamen nog 50% onthouden is# dan zou de verwachte tijd door de testverlenging slechts dalen van $1 \cdot 35$ eenheden naar $1 \cdot 27$ eenheden. Beneden de 20 items maakt het echter wel wat uit, vooral als de onthoudfractie klein is# En merkwaardig genoeg al deze erg groot is# 90%. dan in een korte toets juist voordelig voor de lijntrekker, maar dat komt natuurlijk vanwege de dan relatief grote invloed van het*gelukl."

Het door Van Naerssen gegeven voorbeeld is duidelijk niet relevant voor de gebruikelijke onderwijssituatie, Bij komt tot studietoortijden van bijv* 1935* laten w * zeggen li koor doornemen van de stof. Er zullen wel studenten zijn die deze strategie volgen, maar roer lijkt toch te veronderstellen dat de stof een groot aantal malen doorgenomen moet worden (waarbij dan niet iedere keer evenveel tijd nodig is. maar minder naarmate de studie vordert)* De berekeningen van Van Naerssen zullen dan minstens aangevuld moeten worden met gegevens over wat er gebeurt waer de aftentgrons a aanzienlijk hoger gesteld wordt*

***Uur van Ne13

Strategie van de docent.L

De docent die graag denkt in absolute termen over het niveau van behoorling van de stof dat hij van zijn studenten verlaagt. zou dit model kunnen gebruiken* Als zijn onderwijssituatie zich niet leent tot het gebruiken van dit model, dan kunnen studenten het ook niet gebruiken en áá het hele model zinloos. behalve in zoverre dat de conclusie dat dit soort modellen niet bruikbaar is is die en die onderwijssituatie natuurlijk wel belangrijk is (Zie de uitspraken daarover gedaan door Van Naerssen in zijn artikel & Ned~a sept 1974)

De docent kan specificeren welke relatieve ware score hij verlangt van zijn student. Daarvan uitgaande, en van de veronderstelling dat zijn studenten de optimale studeerstrategie volgen, kan hij berekenen welke aftestgrens hij moet kiezen, bij toetsen met een bepaald aantal items. Het geheel wordt natuurlijk beïnvloed door de empirisch te bepalen onthoudfractie a . In de meeste omstandigheden zal de aftest grens lager gekozen moeten worden dan de relatieve ware score. Het model laat echter onmiddellijk zien, wat natuurlijk ook wel beredeneerd zou kunnen worden, dat er omstandigheden zijn waarin de aftestgrens hoger gesteld moet worden. ("Alleen bij een combinatie van weinig items en een hoge onthoudfractie moet de aftestgrens hoger gesteld worden, >")

Wanneer men denkt het model in zijn onderwijssituatie toe te kunnen passen, dan kunnen diverse beleidsmogelijkheden er uit afgeleid worden

"Zo kan men aan grafieken zien dat men het aantal items van een bepaald tentamen gerust van 90 tot 50 terug kan brengen zonder dat dit gemiddeld meer studeertijd zou kosten, mits men tegelijk het aantal tentamengelegenheden per jaar zodanig laat toenemen dat de onthoudfractie stijgt van bijvoorbeeld 0,5 tot 0,7.11

Omgekeerd zou je kunnen kijken wat er gebeurt wanneer herkansingen zeer snel volgen op de eerste tentamengelegenheid, zodat de onthoudfractie bijna gelijk 1 wordt (en onder de veronderstellingen van allee of niets leren model gelijk 1 moet zijn), Hoeveel items moet een tentamen dan bevatten om er zeker van te zijn dat studenten die niet de optimale strategie volgen maar een kleiner deel van de stof geleerd hebben, onderzocht kunnen worden van degenen die tenminste de optimale strategie gevolgd hebben. Dit lijkt dan al aardig op een probleem dat past in het beoelkundig model.

Nu ik er zo over denk, is het opvallend dat Van Naerssen naar ik meen nergens in zijn artikelen spreekt over de situatie waar studenten niet de optimale strategie volgen, maar niet genoeg gestudeerd hebben om de door de docent gewenste relatieve ware score te bereiken*. Uiteraard hebben deze studenten ook kansen om te slagen voor de tentamens, zij het dat hun kansen kleiner zullen zijn dan van degenen die de optimale studiestrategie gevolgd hebben, Juist omdat zij die optimale strategie niet volgen*. Zou misschien een redenering op te zetten zijn dat onder de aannamen van het model, en acties van de docent die daarop gebaseerd zijn, geen rekening gehouden hoeft te worden met de foute beslissingen die genomen worden wanneer studenten worden toegelaten die niet tot aan de gevraagde relatieve ware score gestudeerd hebben? Dat lijkt me niet het geval, omdat dat zou betekenen dat de zakslaag beelisalagoetragie van de docent zeer conservatief zou zijn. in het nadeel van de studenten dus en in het nadeel van het onderwijs, Van Naerssen zal daarop zeggen dat een en ander afhangt van de herkansingsmogelijkheden, die immers ook al in zijn model zitten# en dat de zaak daarmee dan toch prettig geregeld is, Het is niet onmogelijk, maar het is duidelijk dat we een en ander toch zwart op wit moeten zien te krijgen, Een leuk projectje dus,

Combineren van toetsen.

"bij de compensatorische methode tolt men de tentamenuitslagen

of ziet gewogen bij elkaar op. terwijl de student bij de correctieve methode voor alle tentamens apart geslaagd moet zijn"

Het is niet duidelijk wat Van Naerssen bedoelt met optellen van tentamenuitslagen. Hij maakt nergens. ook niet in zijn 1968 artikel# restricties in de zin dat ook bij een compensatorische methode bepaalde minimum prestaties tenminste geleverd moeten zijn. Dat zou voorlopig moeten betekenen dat wanneer van Naerssen over een compensatorisch model spreekt* hij ieder cijfer in principe combineerbaar vindt door hogere cijfers elders. .1) Verder wordt opengelaten hoe verschillende v ~ en afzonderlijk gewogen zouden moeten worden# of dat zelfs weging in de meeste onderwijssituaties geen zin zal hebben (wat niet zo onwaarschijnlijk is# Het laatste zou aan de hand van de literatuur nog wat verhelderd kunnen worden* Ook opengelaten wordt het schaalprobleem*ag hoewel dat in zekere zin nauw verwant is aan het wegingsprobleem (een vak met een grotere spreiding in cijfers wordt in feite ook zwaarder meegevoegd)# Je zou kunnen denken aan een standaardloosering van scores voordat de scores gecombineerd worden. In ieder geval speelt voor de afzonderlijke vakken niet het eesuurprobleem omdat de censuren niet gelegd worden bij het examen In de combinatie van afzonderlijke vakken Het laatste houdt nog de mogelijkheid open dat lage cijfers door de student ook nog in hogere omgezet kunnen worden door het tentamen nog eens te herhalen (wat onder de aannamen in het model van Van Naerssen ook vrij zeker betekent dat die student dan inderdaad de nodige aandacht aan de inhoud van het vak moet besteden om die hoger* waardering te kunnen bereiken)* In verband met het laatste punt wordt onverklaarbaar dat Van Naerssen op enkele plaatsen flirt met de gedachte (veronderstelling) dat een compensatorische methode ook zou betekenen dat degenen die uiteindelijk zakken# het benodigde aantal punten niet halen* dan ook alle vakken over zouden moeten doen# Immers# ook onder het compensatorisch model met de mogelijkheid om lage scores te verbeteren door herhaling van het tentamen is het altijd mogelijk om te slagen voor de exameneisen (zoals ook voor het correctieve model zonder tijdbeperking geldt)

Van Naerssen 1968: "Bij de compensatorische methode is het theoretisch denkbaar dat een student van een of meer vakken gehooft niets weet# maar hij zal dit tenzinnste moeten compenseren met goede prestaties in andere vakken"

Ook tentamens die verschillende vakken mogen laten compenseren. 9