

# DIE EFFEK VAN PUNTTOKENNINGSTANDAARDE OP DIE KORRELASIE TUSSEN Matriek- EN UNIVERSITEITSPRESTASIE

G. K. HUYSAMEN

*Departement Sielkunde  
Universiteit van die Oranje-Vrystaat*

## ABSTRACT

**The effect of marking standards on the correlation between matriculation and university performance.** Marking standards differ between courses in the sense that the same group of students shows different course means in different courses. Goldman and Widawski (1976) proposed a within-persons procedure to adjust for such differences between university departments and Elliott and Strenta (1988) adjusted this procedure to perform such corrections between courses within departments. In this study the correlation between matriculation performance and university performance has been increased by applying this adjustment not only to university courses but also to matriculation subjects.

## OPSOMMING

Punttoekenningstandaarde verskil tussen kursusse in die opsig dat dieselfde groep studente verskillende kursusgemiddeldes in verskillende kursusse toon. Goldman en Widawski (1976) het 'n tussenproefpersoonprosedure voorgestel om vir sodanige verskille tussen universiteitsdepartemente te korrigeer en Elliott en Strenta (1988) het dit aangepas om sodanige korreksies tussen kursusse binne departemente uit te voer. In hierdie ondersoek is die korrelasie tussen matriekprestasie en universiteitsprestasie verhoog deur hierdie aanpassing nie slegs op universiteitskursusse nie, maar ook op matriekvakke toe te pas.

Sedert J. McKean Cattell ongeveer 'n eeu gelede geprobeer het om akademiese prestasie aan die Columbia-Universiteit in New York te voorspel (Gulliksen, 1950), het sodanige voorspellingsondersoeke gefokus op die soektog na geldige voorspellers en na wyses om hulle optimaal te kombineer. Plaaslik en internasionaal is oor die algemeen gevind dat aanwysers van vorige akademiese prestasie (bv. hoërskoolprestasie) die beste en aanlegtoetsprestasie die naasbeste voorspeller van eerstejaarsprestasie op universiteit is (Ramist, Lewis & McCamley-Jenkins, 1994; Stoker, Engelbrecht, Crowther, Du Toit & Herbst, 1985). Hoewel hierdie veranderlikes die beste voorspellers was, was hul voorspellingvermoë nie perfek nie. Op grond van hierdie nie-perfekte korrelasies het verskeie navorsers die gebruik van hierdie voorspellers gekritiseer – sonder om ooit met doeltreffender voorspellers vorendag te kom. Sedert die sestigerjare het die rol van die aanwyser van universiteitsprestasie wat as kriterium gebruik word, in die Verenigde State van Amerika (VSA) onder die soeklig gekom. Daar is immers twee veranderlikes betrokke by die onderhawige korrelasie en die verkreeë waarde van die korrelasie is nie uitsluitlik 'n funksie van enige van die twee nie. Die maksimum korrelasie tussen 'n voorspeller en 'n kriterium kan byvoorbeeld nie die produk van die vierkantswortels van hul onderskeie betroubaarhede oorskry nie (Lord & Novick, 1968).

Die mees algemene maatstaf van universiteitsprestasie in die VSA is die sogenaemde *freshman grade point average* (FGPA). Daar moet in gedagte gehou word dat in die VSA nie persentasiepunte (soos aan Suid-Afrikaanse skole en universiteite) toegeken word nie, maar simbole (gewoonlik *A, B, C, D* en *F*, waar *F* op 'n druippunt dui). Om die FGPA te bereken, word syfers 4 tot 0 gewoonlik aan hierdie simbole in elke kursus toegeken en vir enige bepaalde student word die gemiddelde van hierdie syfers oor die kursusse wat hy of sy gevolg het, bereken.

Die ooreenstemmende kriterium aan Suid-Afrikaanse universiteite is die gemiddelde kurrikulumpunt (GKMP), oftewel, die gemiddelde van die persentasiepunte wat 'n student oor die kursusse in sy of haar kurrikulum behaal het. Hierdie gemiddelde moet onderskei word van die kursusgemiddelde wat vir 'n groep studente in 'n bepaalde kursus bereken word.

In navolging van Goldman en Slaughter (1976) het verskeie

Amerikaanse navorsers (Elliott & Strenta, 1988; Goldman & Widawski, 1976; Strenta & Elliott, 1987; Stricker, Rock, Burton, Muraki & Jirele, 1994) die rol van universiteitspunttoekenningstandaarde op die korrelasie met universiteitsprestasie ondersoek. Met verskille in punttoekenningstandaarde (Eng.: *grading of marking standards*) word bedoel dat dieselfde groep studente verskillende kursusgemiddeldes in verskillende kursusse behaal. Punttoekenningstandaarde soos hier gedefinieer verskil ongetwyfeld op universiteit van kursus tot kursus en van departement tot departement. Goldman en Widawski (1976) het byvoorbeeld gevind dat terwyl die gemiddelde simboolpunt vir alle studente aan 'n Kaliforniese universiteit gelyk was aan B-minus, die deursnee-student slegs 'n C-plus in die natuurwetenskappe, maar 'n B of hoër in die sosiale wetenskappe sou kon verwag. In 'n ondersoek met 956 eerstejaarstudente in die destydse Transvaal en Natal, het Crouse, Erens en Steffens (1982) iets soortgelyks gevind: Om dieselfde vlak van sukses te bereik, word 'n hoër matriekgroototaal vereis om bepaalde vlakke van sukses in die B.Sc- en B.Comm-richtings as in die B.A.-richting te behaal.

Die onderhawige ondersoek sluit aan by hierdie onderwerp deurdat dit die rol van punttoekenningstandaarde by nie slegs universiteitsprestasie soos voorheen nie, maar ook by hoërskoolprestasie, op die korrelasie tussen hierdie twee veranderlikes ondersoek. Klaarblyklik sou die bedryfs- of personeelsielkundige by die keuse tussen kandidate op grond van hul matriek- en/of universiteitspunte enige verskille in die punttoekenningstandaarde van hul onderskeie vakkombinasies in ag moet neem indien sodanige verskille hul onderskeie gemiddelde prestasies sou affekteer. Insgelyks sou universiteitstoelatingspersoneel in so 'n geval daarop bedag moet wees dat die voorspellingsgeldigheid van matriekprestasie deur die punttoekenningstandaarde van matriekpunte en/of universiteitspunte geaffekteer word.

## Benoeming en implikasies van interdepartementele punttoekenningsverskille

Daar sou geargumenteer kon word dat wat hier as verskille in punttoekenningstandaarde genoem word bloot daaraan te wyte kan wees dat die onderrigmetodes in die kursusse wat hoër kursusgemiddeldes vertoon, doeltreffender is as dié in kursusse waarin laer kursusgemiddeldes gerealiseer word. Goldman en Widawski (1976) het egter aangetoon dat studente in (universi-

*Versoeke om afskrifte moet gerig word aan G. K. Huysamen, Departement Sielkunde, UOVS, Posbus 339, Bloemfontein 9300.*

teits-) departemente met relatief streng puntoekenningstandaarde (volgens die onderhawige definisie) gemiddeld hoër tellings op die Scholastic Aptitude Test (SAT) behaal het as dié in departemente met relatief toegeeffiker bepuntingstandaarde. Dis onwaarskynlik dat die doeltreffender onderrigmetodes juis beperk sou wees tot kursusse met gemiddeld swakker studente en dat hierdie onderrigmetodes so doeltreffend sou wees dat dit die swakker studente beter sou laat presteer as die sterker studente in die ondoeltreffende kursusse.

Verder sou aangevoer kon word dat die onderhawige verskille eintlik verskille in die moeilikheidsgrade van verskillende vakke of kursusse is. In die psigometrika word die moeilikheidswaardes van twee (digotomiese) toetsitems bepaal deur die proporsies proefpersone in dieselfde groep wat die twee items onderskeidelik geslaag het, te bereken. As in ag geneem word dat 'n proporsie ooreenstem met die gemiddelde van 'n digotomiese veranderlike, kom die ooreenkoms met die definisie van puntoekenningstandaarde aan die hand van die kursusgemiddeldes van dieselfde groep persone na vore. Daar is waarskynlik 'n diplomatieuse rede waarom oorspronklik op die term *grading standards* besluit is: Dosente in een vak sou beswaarlik daarmee verlies neem as hul vakgebied as makliker as 'n ander geëtiketteer sou word. Moontlik sou die benaming puntoekenningverskille (*grading differences*) 'n meer neutrale term wees, maar *grading standards* het reeds oor die afgelope paar dekades in die psigometrie literatuur ingeburger geraak (vgl. die titels van die artikels in die verwysingslys). Hierdie toedrag van sake herinner aan die beskrywing van die proporsie proefpersone wat 'n digotomiese toetsitem geslaag het as sy *moeilikhed* (*p*-waarde) terwyl *itemmaklikheid* 'n toepasliker benaming sou wees. Nogtans is die benoeming van die onderhawige verskynsel nie heeltemal irrasioneel nie. Al waarmee die navorser te make het, is verskillende kursusgemiddeldes vir dieselfde groep proefpersone. Om sodanige verskille uitsluitlik aan die verskillende moeilikheidsgrade van verskillende vakinhoud te skryf, sou eweneens ongeoorloof wees.

Die onderhawige onderwerp het natuurlik politieke ondertone in die opsig dat inskrywingsgetalle implikasies het vir departementele personeeltoekenning. Strenta en Elliott (1987, p.282) beweer: "Open discussion of differential grading standards might be as unwelcome as open discussion of differential teaching loads or differential salaries: All such discussions invite invidious comparisons and accompanying resentment". Nogtans kan die rol van bepuntingstandaarde in studente se vakkeuses nie geïgnoreer word nie. Die rol van hierdie veranderlike het Ramist et al. (1994, p.12) genoop om onomwonde te verklaar dat "(t)he three factors contributing to a high (freshman grade point average) are high aptitude for college work, good performance in high school, and selection of courses that are graded leniently".

Volgens die Spearman-Brown-formule en die formule vir Cronbach se alfa-koëffisiënt neem die betroubaarheid van 'n veranderlike toe namate dit met onderskeidelik parallelle en tau-ekwivalente komponente verleng word (Lord & Novick, 1968). Dit sou egter nie deug om na aanleiding hiervan te redeneer dat die GKMP noodwendig betroubaarder moet wees as die persentasiepunt in 'n enkele kursus nie. Met die uitsondering van professioneel-georiënteerde kurrikula (soos dié vir maatskaplike werk, verpleegkunde, medies, fisioterapie) verskil die versameling kursusse van student tot student sodat die GKMP nie dieselfde veranderlike vir studente is nie. Selfs in kurrikula wat uit dieselfde kursusse saamgestel is, kan puntoekenningverskille tussen kursusse verhoed dat die persentasiepunte in verskillende kursusse hulle in ooreenstemming met parallelle of tau-ekwivalente komponente gedra. As gevolg van verskille in puntoekenningstandaarde tussen kursusse en van verskille in die samestelling van die kurrikula van verskillende studente kan die resultaat van die verlenging met genoemde soort komponente dus beswaarlik op die onderhawige situasie van toepassing wees.

### Tussenproefpersoonprosedures om te korreger vir verskille in puntoekenningstandaarde

Young (1993) en Stricker et al. (1994) het 'n oorsig gelewer van verskeie metodes om die FGPA vir verskillende puntoekenningstandaarde te korreger en laasgenoemde navorsers het ook die relatiewe doeltreffendheid van hierdie metodes empiries ondersoek. Hulle het onderskei tussen metodes wat op interne gegewens gebaseer is en dié wat op eksterne gegewens gegrond is. Onder eersgenoemdes is daar byvoorbeeld die metode wat aanvanklik deur Goldman en Widawski (1976) ontwikkel is. Essensieel behels hierdie tussenproefpersoonprosedure die volgende stappe:

(i) Eerstens word vir elke paar kursusse die groep studente wat albei voltooi het, geïdentifiseer. (Vir 6 kursusse, sal daar dus  $6(6-1)/2 = 15$  sulke verskillende pare kursusse wees.) Op hierdie wyse word die variansie wat aan individuele verskille in algemene vermoë, motivering en studiemetodes toegeskryf kan word, beheer omdat presies dieselfde groep proefpersone by enige sodanige paar kursusse betrokke is.

Van belang op hierdie stadium, is 'n tabel waarin sowel die kolom- as die ryopskrifte uit die benamings van dieselfde versameling kursusse bestaan en enige (nie-diagonale) sel die getal proefpersone aantoon wat sowel die vak in sy kolomopskrif as die een in sy ryopskrif aangebied het. (In die sel waarvan die kolomopskrif *Aardrykskunde* en die ryopskrif *Wiskunde* is, verskyn dus die getal studente wat sowel *Aardrykskunde* as *Wiskunde* as vakke gevolg en daarin punte behaal het.) Vervolgens sal na hierdie tabel as die *N*-matriks verwys word.

(ii) Vir elke groep wat enige paar kursusse gevolg het, word die verskil tussen hul gemiddeldes in die twee kursusse bereken. Die resultaat is 'n matriks soortgelyk as die *N*-matriks, maar met verskille tussen gemiddeldes in plaas van steekproefgroottes in die onderskeie nie-diagonale selle. (In die sel wat ooreenstem met die een wat in die vorige paragraaf ter sprake was, sal dus die verskil verskyn tussen die gemiddeldes wat dieselfde groep proefpersone vir onderskeidelik *Aardrykskunde* en *Wiskunde* behaal het.) Die argument is dat indien dieselfde groep studente 'n hoër gemiddelde punt in een kursus as in 'n ander behaal, laasgenoemde kursus onderhewig aan strenger puntoekenningstandaarde as eersgenoemde is en, omgekeerd, toegeeffiker bepuntingstandaarde in eersgenoemde kursus as in laasgenoemde geld.

(iii) Vir elke kursus word die gemiddelde van hierdie verskille met al die ander kursusse bereken en die verkreeë gemiddelde verskil is dan die indeks wat die relatiewe strengheid of toegeeffikheid van die puntoekenningstandaard in daardie kursus verteenwoordig: 'n Positiewe indeks sou dui op 'n kursus wat gemiddeld toegeeffiker as die ander kursusse nagesien word; omgekeerd, 'n negatiewe waarde vir hierdie indeks sou betrekking hê op 'n kursus wat relatief strenger bepunt word as ander kursusse.

(iv) Studente se tellings in enige gegewe kursus kan vervolgens aangepas word deur die indeks wat in die vorige stap bepaal is, van hul tellings af te trek. Die tellings in kursusse met 'n positiewe indeks word dus verminder met die grootte van daardie indeks, terwyl die tellings in kursusse met 'n negatiewe indeks, verhoog word met die grootte van dié indeks.

Goldman en Widawski (1976) het hierdie metode op universiteitsdepartemente toegepas deur die punte van alle studente wat een of meer kursusse in elk van enige paar departemente gevolg het, te gebruik vir die berekening van die aanpassingsindekse van sodanige departemente. Elliott en Strenta (1988) het 'n korrelasie van 0,55 gevind tussen die rangordes van hierdie indekse wat vir 16 departemente by Dartmouth College bepaal is en die rangordes wat 10 jaar later daarvoor vir die gelyknamige departemente aan die Riverside-kampus van die Universiteit van Kalifornië afgelei is. Strenta en Elliott (1987) het gevind dat die korrelasie tussen die SAT en FGPA vir 1 129

studente van 0,43 tot 0,48 gestyg het nadat die FGPA volgens die onderhawige metode aangepas is.<sup>1</sup>

In 'n opvolgondersoek wou Elliott en Strenta (1988) die oorspronklike metode uitbrei om nie slegs bepuntingstandaarde tussen departemente nie, maar ook dié tussen kursusse binne departemente te vergelyk. In hierdie ondersoek het hulle te staan gekom voor die probleem dat, as gevolg van die klein getal studente in die gevorderde kursusse, daar te min studente in bepaalde kombinasies kursusse was wat in stap (i) hierbo figureer. Hulle het hierdie probleem oorkom deur die onderhawige tussenproefpersoonprosedure soos volg uit te brei om aanpassingsindekse vir kursusse binne departemente af te lei:

(i) Vir elke kursus wat 'n student in 'n gegewe departement voltooi het, word sy of haar gemiddelde oor die oorblywende kursusse wat hy of sy in daardie departement voltooi het, bereken. Byvoorbeeld, indien 'n student kursusse PSY101, PSY108 en PSY202 in 'n Departement Sielkunde gevolg het, is ook gemiddeldes verkry vir die volgende: alle kursusse uitgesonderd PSY101 (d.w.s. PSY108 en PSY202); alle kursusse uitgesonderd PSY108 (d.w.s. PSY101 en PSY202); en alle kursusse uitgesonderd PSY202 (d.w.s. PSY101 en PSY108). Vir elke student word daar dus net soveel bykomende tellings (gemiddeldes oor oorblywende kursusse in die departement) bereken as wat hy of sy kursusse in daardie departement voltooi het en hierdie getal kan uiteraard van student tot student verskil.

(ii) Die indeks vir enige gegewe kursus in 'n bepaalde departement word bepaal as die verskil tussen die gemiddelde van die studente wat vir daardie kursus geregistreer het en die gemiddelde van hierdie studente se onderskeie gemiddeldes oor hul oorblywende kursusse in daardie departement. Die indeks vir PSY101 sou byvoorbeeld wees die verskil in die gemiddelde van die PSY101 tellings van diegene wat vir hierdie kursus geregistreer het en die gemiddelde van die gemiddeldes wat hierdie studente in hul ander PSY-kursusse behaal het.

Terwyl die oorspronklike metode van Goldman en Widawski (1976) berus op die vergelyking tussen die prestasies wat dieselfde groep proefpersone in twee gemeenskaplike departemente behaal het, maak die Elliott-Strenta-aanpassing (1988) daarvan voorsiening vir 'n vergelyking tussen 'n groep se prestasie in een kursus en hul prestasies in hul oorblywende kursusse in dieselfde departement. Voortaan sal na hierdie twee metodes as onderskeidelik die beperkte en die onbeperkte metode verwys word.

Elliott en Strenta (1988) het 927 studente se punte in enige kursus in 'n bepaalde departement aangepas met sowel die indeks vir daardie departement (volgens die beperkte metode) as die indeks vir die kursus in daardie departement (volgens die onbeperkte metode). Terwyl die SAT 0,44 en 0,43 onderskeidelik met FGPA en 'n kumulatiewe GPA oor vier jaar gekorreleer het, het hierdie korrelasies tot onderskeidelik 0,49 en 0,50 gestyg nadat dit met hierdie uitgebreide prosedure gekorrigeer is. Hierdie resultaat is herhaal in 'n ondersoek deur Stricker et al. (1994) waarin hierdie aanpassings die korrelasie tussen die hoërskoolrangorde en FGPA van 4 351 studente van 0,34 tot 0,41 laat styg het.

Uit bostaande beskrywings is dit duidelik dat, veral die beperkte metode, optimaal tot sy reg sal kom indien studente min of meer vir dieselfde versameling kursusse registreer. As een groep studente bv. kursusse A, B en C volg en 'n ander

groep, kursusse D, E en F, is die aanpassingsindekse wat vir eersgenoemde groep kursusse verkry word nie noodwendig regstreeks vergelykbaar met dié wat vir laasgenoemde groep bepaal is nie. Namate die getal kursusse toeneem wat gemeenskaplik deur alle studente aangebied word, verbeter die toestand om vergelykbare aanpassingsindekse vir alle kursusse af te lei. Verder volg dit dat hierdie korreksieprosedures die toename in die korrelasie tussen enige voorspeller en universiteitsprestasie sal verhoog namate dieselfde groep studente aansienlik verskillende gemiddeldes in die verskillende kursusse behaal. Trouens, indien dieselfde groep studente presies dieselfde gemiddelde vir verskillende kursusse vertoon, sou dit op die afwesigheid van verskille in punttoekenningstandaarde volgens die onderhawige definisie dui.

### Punttoekenningstandaarde op skoolvlak

Dieselfde bepuntingverskille wat tussen universiteitsdepartemente en -kursusse bestaan, sou natuurlik ook by die bepunting van skoolvakke verwag kon word. In hul omvattende ondersoek waarby onder meer 16 Suid-Afrikaanse universiteite betrokke was, het Stoker et al. (1985) gevind dat naas matriekgroototaal, matriekowerheid een van die beste voorspellers van universiteitsprestasie in verskeie vakrigtings was. Indien dieselfde punttoekenningstandaarde deur verskillende matriekowerhede gehandhaaf sou word, sou matriekowerheid geen voorspellingsvermoë getoon het nie.

Hoewel Amerikaanse navorsers meer veranderlike statistiese tegnieke aangewend het om die korrelasie tussen hoërskoolpunte en FGPA te verhoog (vgl. die oorsig van Linn, 1966), kon geen ondersoek gevind word waarin die onderhawige korreksieprosedure op hoërskoolpunte aangewend is nie. Met hul gedentraliseerde onderwysstelsel, is die afwesigheid van sodanige ondersoeke verstaanbaar. Elke skooldistrik, indien nie elke hoërskool nie, neem sy eie skoolindeksamen af. Gevolglik sou dit nie sinvol wees om die simboolpunte in enige paar vakke van aansoekers, wat letterlik van honderde verskillende hoërskole afkomstig kan wees, volgens die onderhawige metodes met mekaar te vergelyk nie. In Suid-Afrika, soos in sommige Europese lande en Israel, is daar egter 'n veel kleiner getal matrikulasieowerhede as in die VSA. Tot onlangs was daar byvoorbeeld slegs vier provinsiale onderwys-owerhede vir wit matrikulante in Suid-Afrika wat elk die opstel, afneem en nasien van die matriekeksamen vir alle skole onder sy jurisdiksie geadministreer het. Soos die jaarverslae van die rektore van Suid-Afrikaanse residensiële universiteite kan getuig, is hierdie universiteite boonop geneig om hoofsaaklik matriekleerlinge te trek uit die onderskeie provinsies waarin hulle geleë is sodat die meeste eerstejaars by 'n bepaalde universiteit geneig is om dieselfde matriekeksamen te geskryf het. Aangesien die meerderheid wit eerstejaars aan byvoorbeeld die Universiteit van die Oranje-Vrystaat (UOVS) die matriekeksamen van die gelyknamige provinsiale onderwysdepartement afgelê het, doen 'n geleentheid hom voor vir die toepassing van die onderhawige prosedures wat nie in dieselfde mate in die Verenigde State van Amerika bestaan nie.

Die doel met die onderhawige projek was om te ondersoek in welke mate die korrelasie tussen matriekprestasie en eerste-semesterprestasie aan die genoemde universiteit sou verhoog indien albei hierdie veranderlikes vir verskille in punttoekenningstandaarde volgens bogenoemde tussenproefpersoonprosedures gekorrigeer sou word.

### METODES

Vir die doeleindes van hierdie ondersoek is die matrieksimboolpunte en die eerste-semesterpunte van die wit studente wat in 1996 as eerstejaars aan die UOVS geregistreer het, gebruik. Matriekprestasie is geoperasionaliseer as 'n gemiddelde matrieksimboolpunt (GMS) deur die simbole A tot en met F in numeriese syfers van onderskeidelik 6 tot 0 om te skakel (afgesien daarvan of 'n vak op hoër of standaardgraad aangebied

<sup>1</sup> In die lig van die groot steekproewe wat tipies op die spel is in universiteitsvoorspellings-ondersoeke, word geen statistiese beduidendheidstoetse op die korrelasies in enige van die ondersoeke soos dié waarna hier verwys word, gewoonlik gerapporteer nie. Met 1 129 proefpersone is 'n korrelasie van byvoorbeeld slegs 0,07 statisties beduidend op die 1 %-peil. Statistiese beduidendheidstoetse vir die verskil tussen die korrelasies van twee voorspellers met dieselfde kriterium (Ferguson, 1966), vereis dat die twee voorspellers eksperimenteel onafhanklik van mekaar moet wees. Aangesien die een voorspeller in die onderhawige ondersoek 'n statistiese aanpassing van die ander is, word daar nie aan laasgenoemde vereiste voldoen nie. Nogtans: onder die veronderstelling dat die korrelasie tussen die oonaangepaste en die aangepaste voorspellers in hierdie voorbeeld gelyk aan 0,70 sou wees, is die verskil van 0,05 in hul korrelasies met die gemeenskaplike kriterium statisties beduidend op die 1 %-peil.



is) en die gemiddelde van die verkreë syfers te bereken. Die korreksieprosedures wat hier op die spel is, vereis dat die simboolpunt in een vak met die gemiddelde oor ander vakke bereken word. Indien die simboolpunte vir hoërgraadvakke in hoër syfers omgeskakel sou word as dié vir standaardgraadvakke, sou die gemiddeldes oor verskillende kombinasies van vakke nie vergelykbaar wees nie (tensy dieselfde getal hoër- en standaardgraadvakke telkens op die spel sou wees). (Boonop word alle matrikulante verplig om minstens twee-derdes van die verpligte ses vakke op die hoër graad te volg, sodat almal in dieselfde mate deur so 'n reëling geaffekteer sou word.) In die geval van studente wat meer as ses vakke aangebied het, is die (ekstra) vakke met die laagste simboolpunte geïgnoreer, tensy dit Afrikaans (Eerste Taal) was (in welke geval die vak met die sesde laagste simboolpunt uitgeskakel is). Met hierdie stap is beoog om die GMS van studente wat verskillende kombinasies vakke aangebied het, meer vergelykbaar met mekaar te maak. Die GMS is gekorrigeer volgens sowel die oorspronklike, beperkte metode as met die onbeperkte een en die GMS wat daarvolgens gekorrigeer is, sal aangedui word met onderskeidelik GMS(B) en GMS(O). Die toepassing van die beperkte metode het behels dat dieselfde groep se prestasie in een vak telkens met hul prestasie in 'n ander vak, vergelyk is, terwyl by die onbeperkte metode hul prestasie in een vak met hul gemiddelde prestasie oor hul onderskeie oorblywende vakke vergelyk is.

Om die beperkte metode tot sy reg te laat kom, is besluit om dit op 'n subgroep matriekvakke toe te pas waarvan elke moontlike paar deur 'n voldoende getal matrikulante van die Vrystaatse Onderwysdepartement gevolg is. Uiteraard kon daar geen leerling wees wat 'n bepaalde vak op sowel die hoër- as die standaardgraad gevolg het nie sodat die sel met die hoërgraadweergawe van enige vak as kolomopskrif en die standaardgraadweergawe van dieselfde vak as ryopskrif in die  $N$ - en die daaropvolgende matrikse per definisie leeg moes wees. Arbitrêr is besluit dat die matriekvakke wat vir hierdie doel geselekteer word, geen leë (nie-diagonale) selle in die  $N$ -matriks moes toon nie (afgesien van die nie-bestaande hoërgraad-standaardgraadkombinasies) en dat nie meer as 25 % van alle (nie-diagonale) inskrywings (insluitend lg. kombinasies) kleiner as 10 moes wees nie.

Die afwesigheid van hoërgraad-standaardgraadkombinasies van vakke is natuurlik nie bevorderlik vir die sukses van die beperkte metode nie. Die oorgrote meerderheid matriekleerlinge van die Vrystaatse Onderwysdepartement het egter Afrikaans, Eerste Taal, op die hoër graad (AFE-H) gevolg. Omdat die groep leerlinge met die hoërgraadweergawe van enige vak en dié met die standaardgraadweergawe van dieselfde vak dus albei AFE-H gevolg het, verskaf hul onderskeie prestasies in AFE-H inligting oor die punttoekenning in die hoërgraad- en standaardgraadweergawes van eersgenoemde vak. Boonop het ten minste sommige leerlinge van die twee groepe in so 'n vak, saam vir sommige van die ander vakke, buiten AFE-H, geregistreer, sodat hul prestasies in sodanige vakke bykomende inligting sou verskaf oor die onderskeie bepuntingstandaarde by die hoër- en standaardweergawes van eersgenoemde vak. Daar is aangeneem dat die probleem van geen hoërgraad-standaardgraadkombinasies op hierdie wyse oorbrug sou word. Boonop was daar uiteindelik slegs twee vakke, te wete Skeien Natuurkunde en Wiskunde, waarvan sowel die hoër- as die standaardgraad onder die groep van 13 vakke voorgekom het wat volgens hierdie prosedure geselekteer is (vgl. Tabel 1).

Die (oorspronklike) kriterium was die gemiddelde kurrikulumpersentasiepunte (GKMP), dit wil sê die gemiddelde van die persentasiepunte wat studente oor die kursusse in hul onderskeie eerste semesterkurrikula behaal het. Wat die universiteitskursusse betref, is 'n seleksie van vakke waarvan alle kombinasies deur 'n voldoende getal studente gevolg is in die Suid-Afrikaanse opset feitlik onmoontlik. Anders as in die VSA waar dit nie ongewoon is vir 'n student om byvoorbeeld vir Wysbegeerte, Ekonomie en Wiskunde in te skryf nie, word hierdie vakke aan Suid-Afrikaanse universiteite gewoonlik in afsonderlike fakulteite (Geesteswetenskappe, Ekonomiese

Wetenskappe en Natuurwetenskappe) aangebied. Oor die algemeen word studente gedwing om hul vakkeuses te beperk tot die vakke wat in hul onderskeie fakulteite ressorteer. Hoewel sommige vakke (bv. Sielkunde, Wiskunde, Bedryfsielkunde, Operasionele Bestuur) aan die UOVS in meer as een fakulteit aangebied kan word, is sodanige vakke by verre in die minderheid. Gevolglik sou die nut van selfs die onbeperkte metode hoofsaaklik lê in die korreksies wat dit binne fakulteite as tussen fakulteite teweegbring. In die lig van hierdie oorweging is die eerste semesterpunte slegs volgens die onbeperkte metode aangepas – om GKMP(O) te vorm. Omdat studente aan Suid-Afrikaanse universiteite, die UOVS ingesluit, gewoonlik 'n enkele kursus per departement per semester volg, was dit nie nodig om, soos Elliott en Strenta (1988), ook afsonderlike aanpassingsindekse vir kursusse binne departemente te ontwikkel nie.

Elk van die (ongekorrigeerde) GMS en die (gekorrigeerde) GMS(B) en GMS(O) is met sowel die (ongekorrigeerde) GKMP as die (gekorrigeerde) GKMP(O) gekorreleer. Hierdie korrelasies is vir drie groepe studente bereken:

- i) Studente wat die matriekexamens van die Provinsiale Administrasie Oranje-Vrystaat geslaag het en wie se matriekvakke almal onder die 13 geselekteerde vakke geval het;
- ii) Alle studente wat die OVS-matriekexamens geslaag het, ongeag hul kombinasie van matriekvakke; en
- iii) Alle studente, ongeag matrikulasieowerheid en ongeag matriekvakkombinasie.

Geen statistiese beduidendheidstoets is op die toenames in korrelasies wat op die aanpassings van die matriek- en/of eerste semesterpunte gevolg het, toegepas nie (vgl. voetnoot 1).

## RESULTATE EN BESPREKING

Tabel 1 toon die aanpassingsindekse wat vir die 13 matriekvakke met voldoende gemeenskaplike inskrywings, volgens onderskeidelik die beperkte metode en die onbeperkte metode verkry is. Van die totaal van 550 OVS-matrikulante was daar 248 leerlinge van wie al ses matriekvakke onder hierdie 13 vakke voorgekom het. Die korrelasie tussen die aanpassingsindekse wat volgens die beperkte en die onbeperkte metode vir die 13 geselekteerde vakke vir hierdie groep bereken is, was 0,988. Die ooreenstemmende korrelasie vir alle vakke wat op die 550 OVS-matrikulante bereken is, was 0,985.

**Tabel 1**  
Aanpassingsindekse vir 13 matriekvakke by dieselfde onderwysdepartement geskryf.

Vak	Slegs 13 vakke		Al 57 Vakke	
	Beperkte	Onbeperkte	Beperkte	Onbeperkte
Afr., 1e Taal, H	0,288	0,308	-0,211	0,240
Aardrykskunde, H	0,662	0,487	0,120	0,242
Bedryfsiekonomie, H	0,409	0,375	0,276	0,197
Biologie, H	0,037	0,153	-0,345	0,066
Duits, H	0,340	0,215	0,164	0,071
Ekonomie, H	-0,083	0,031	-0,594	-0,262
Eng., 2e Taal, H	0,292	0,303	-0,148	0,197
Geskiedenis, H	0,316	0,180	-0,127	0,064
Rekeningkunde, H	-0,367	-0,243	-0,436	-0,275
Skei-Nat, H	-0,549	-0,350	-0,774	-0,365
Skei-Nat, S	-0,337	-0,259	-0,924	-0,347
Wiskunde, H	-1,132	-0,832	-0,776	-0,614
Wiskunde, S	-0,086	-0,067	-0,089	0,079

Volgens Tabel 2 was die korrelasie tussen die GMS en die GKMP van die totale groep wit eerstejaars ( $N = 976$ ) gelyk aan 0,699. Soos verwag kon word, is die ooreenstemmende GMS-GKMP-korrelasie vir die groep oud-Vrystaat matrikulante van wie al ses matriekvakke onder die geselekteerde groep van 13 vakke geval het, hoër (0,729). Hoewel hierdie groep nie almal presies dieselfde vakke gevolg het nie, het hulle almal dieselfde AFE-H-vraestel geskryf en was daar groot

oorvleueling tussen hul oorblywende vakke. Vir hierdie groep oud-Vrystaat matrikulante styg die GMS-GKMP-korrelasie tot 0,742 indien óf die beperkte metode, óf die onbeperkte metode op slegs die matriekvakke, uitgevoer word. Hierdie toenames suggereer die uitwerking van verskillende puntoekenningstandaarde in verskillende provinsies en in verskillende vakke binne dieselfde provinsie op die korrelasie tussen matriek- en eerstejaarprestasie. Wanneer die onbeperkte metode terselfdertyd op die universiteitsvakke uitgevoer word, styg die ooreenstemmende korrelasies vir die groep OVS-matrikulante tot 0,750 en 0,749, onderskeidelik. 'n Toename in korrelasie van 0,699 tot 0,75 stem ooreen met 'n toename in persentasie verklaarde kriteriumvariansie van 7,39. Dieselfde patroon van toenames word waargeneem indien die korreksie op alle OVS oud-matrikulante afgesien van matriekvakkombinasie toegepas word.

**Table 2**  
**Korrelasies tussen aangepaste en onaangepaste matrieksimboolpunte en kurrikulumpersentasiepunte**

		GKMP	GKMP(O)	N
13 vakke OVS-matrikulante	GMS	0,729	0,735	248
	GMS(B)	0,742	0,750	248
	GMS(O)	0,742	0,749	248
Alle vakke OVS-matrikulante	GMS	0,731	0,734	550
	GMS(B)	0,737	0,742	550
	GMS(O)	0,742	0,746	550
Alle vakke Alle studente	GMS	0,699	0,711	976
	GMS(B)	0,728	0,739	976
	GMS(O)	0,728	0,739	976

Nota: GMS: Gemiddelde matrieksimboolpunt; GMS(B): Gemiddelde matrieksimboolpunt aangepas volgens beperkte metode; GMS(O): Gemiddelde matrieksimboolpunt aangepas volgens onbeperkte metode; GKMP: Gemiddelde kurrikulumpersentasiepunt; GKMP(O): Gemiddelde kurrikulumpersentasiepunt aangepas volgens onbeperkte metode

Wanneer die korrelasies in die laaste gedeelte van Tabel 2 in oënskou geneem word, moet sekere beperkende faktore in gedagte gehou word. Eerstens was daar ten minste vyf onderwysdepartemente onder die eerstejaarstudente verteenwoordig, sodat daar 'n afname was in die getal studente wie se prestasie in een matriekvak met hul prestasie in 'n ander matriekvak onder dieselfde onderwysowerheid vergelyk kon word. Tweedens was daar altesaam 94 eerste semesterkursusse uit nege fakulteite wat elk geneig is om sy studente tot 'n afsonderlike groep kursusse te beperk. Die N-matriks het dus 'n meerderheid leë selle vertoon. Ten spyte van hierdie beperkende faktore styg die GMS-GKMP-korrelasie van 0,699 tot 0,711 indien slegs GKMP gekorrigeer word en verhoog dit nog verder tot 0,739 indien die korreksies op sowel GMS as GKMP uitgevoer word – 'n uiteindelijke toename in verklaarde kriteriumvariansie van 5,75%. Hierdie resultate suggereer dat die onderhawige korreksiemetodes taamlik robuus is ten opsigte van die vereiste dat daar min leë nie-diagonale selle in die N-matriks moet wees.

Om die grootte van die toenames in die GMS-GKMP-korrelasie ná die onderhawige korreksies in perspektief te plaas, is 'n meervoudige regressie uitgevoer op die groep OVS-matrikulante vir wie Algemene Skolastiese Aanlegtoetstellings (ASAT) (Claassen, De Beer, Hugo & Meyer, 1991) beskikbaar was. Onder die groep van 248 OVS-matrikulante wie se matriekvakke almal onder die 13 geselekteerdes geval het, was daar 155 vir wie ASAT-tellings beskikbaar was. Vir hierdie groep het die GMS-GKMP-korrelasie van 0,711 tot ('n meervoudige korrelasie van) 0,717 gestyg, teenoor die styging tot 0,736 indien die matriekpunte volgens die beperkte metode gekorrigeer word. Vir die subgroep van 326 studente vir wie ASAT-tellings beskikbaar was onder die totale groep van 550 OVS-matrikulante, was daar 'n toename slegs in die vyfde desimaal (van 0,75622 tot 0,75628) ná die toevoeging van die ASAT-tellings in 'n meervoudige regressie. Daarteenoor het die GMS-GKMP-korrelasie gestyg van 0,756 tot 0,765 ná die korreksie volgens die beperkte metode en tot 0,767 ná die kor-

reksie volgens die onbeperkte metode (op matriekvakke). Die onderhawige gegewens suggereer dus dat dit voordeliger mag wees om matriekpunte vir variasie in puntoekenningstandaarde te korrigeer as om 'n toets soos die ASAT te laat afneem en as bykomende voorspeller te gebruik.

In hierdie ondersoek was slegs die matrieksimboolpunte (A, B, ens.) beskikbaar. Daar kan verwag word dat die effek van die korreksieprosedures groter sou wees indien die spesifieke persentasiepunte beskikbaar was. Ná afloop van matriekksamens oorweeg die Suid-Afrikaanse Sertifiseringsraad die punte in verskillende vakke en hierdie liggaam kan dit aanpas om dit onderling meer vergelykbaar te maak (Prof. D.J. de Waal, lid van die genoemde raad, persoonlike kommunikasie, 30 Maart 1999). Die onderhawige metode sou klaarblyklik met groot vrug vir hierdie doel oorweeg kon word. Deur byvoorbeeld die prestasies in twee vakke gelyk te stel deur hul gemiddeldes oor slegs die onderskeie groepe wat hulle gevolg het, gelyk te stel, word die gemiddeldes van dieselfde subgroep wat albei gevolg het nie noodwendig gelykgestel nie. Gevolglik word daar moontlik nie voldoende gekompenseer vir die laer gemiddelde wat die subgroep leerlinge behaal het in 'n vak met relatief streng puntoekenningstandaarde nie.

## SYNOPSIS

### THE EFFECT OF MARKING STANDARDS ON THE CORRELATION BETWEEN MATRICULATION AND UNIVERSITY PERFORMANCE

Grading standards differ between university courses in the sense that the same group of students shows different course means in different courses. Goldman and Widawski (1976) proposed a within-persons procedure to adjust for such differences. For every pair of courses, this approach identifies the group of students who has registered for both and obtains the difference between this group's mean (course) grades in them. The mean of these differences between a particular course and all other courses represents an index of the grading leniency or strictness of that course. A positive index would indicate a course which, on average, is marked more leniently than the other courses; conversely, a negative value would signify relatively stricter grading standards. Students' grades may be adjusted by subtracting these indices from their grades in courses showing positive adjustment indices and adding their absolute values in the case of courses with negative indices. Elliott and Strenta (1988) extended this procedure to allow for meaningful comparisons between the marking standards of courses, within departments, with small enrollment numbers (e.g., high-level courses). This adjustment entails that the mean performance of the group of students who has registered for any particular course is compared with the mean of these students' means over their respective sets of remaining courses in that department. (Because this method is not restricted to situations where there is a large number of students for every pair of courses, it will be referred to as the unrestricted method.) By combining the original Goldman-Widawski procedure to compare marking standards between departments and their own adjustment procedure to compare grading standards within departments, Elliott and Strenta demonstrated an improvement in the validity of high school rank, Scholastic Aptitude Test scores and achievement tests scores as predictors of university performance.

Due to their decentralized high school system with a different examination authority in almost every city or county, it comes as no surprise that no study could be found in which the above procedure has been applied to high school courses or subjects in the United States of America. However, the application of this procedure is more feasible in South Africa with its much smaller number of matriculation authorities. In the present study both the original and the unrestricted method were applied to the matriculation marks of the following three groups of first-year students from historically white high schools who registered at the University of the Orange Free

State in 1996: the 248 students who wrote the matriculation examination of the Orange Free State Province the preceding year and whose matriculation subjects were all among the set of 13 subjects with the highest degree of pairwise registrations; all 550 students who wrote this matriculation examination, irrespective of subjects taken; and all 976 first-year students, irrespective of their matriculation authority. The unrestricted method was also applied to the mean of the percentage marks (MCPM) obtained in the first-semester curriculum.

The correlation between the matriculation symbol point total (MSPT) and MCPM was highest (namely, 0,729) for the group of OFS matriculation students whose subjects were among the set of 13 subjects, as opposed to a correlation of 0,699 for the total group of students. The difference between these correlations suggests that the presence of different marking standards between different matriculation authorities affects the present correlation negatively. When the MSPT of the total group of students was adjusted by either the original or the unrestricted method, this correlation increased to 0,728 and when MCPM was also adjusted, the correlation rose to 0,739. The highest correlation of 0,750 was obtained for the smaller group of OFS matriculants after their MSPT and MCPM were adjusted by the restricted and unrestricted methods respectively. For a subgroup of students for whom scores on the General Scholastic Ability Test (GSAT) were available, a greater degree of predictive effectiveness was obtained by adjusting MSPT by the above procedures than by including GSAT scores as an additional predictor.

#### ERKENNING

Finansiële bystand gelewer deur die Sentrum vir Wetenskap-ontwikkeling (RGN, Suid-Afrika) vir hierdie navorsing word hiermee erken. Menings uitgespreek en gevolgtrekkings waartoe geraak is, is dié van die outeur, en moet nie noodwendig aan die Sentrum vir Wetenskapontwikkeling toegeskryf word nie.

#### VERWYSINGS

Claassen, N.C.W., De Beer, M., Hugo, H.L.E. & Meyer, H.M. (1991). *Handleiding vir die Algemene Skolastiese Aanlegtoets*

(ASAT). Pretoria: Raad vir Geesteswetenskaplike Navorsing.

- Crouse, C.F., Erens, G. & Steffens, F.E. (1982). *Ondersoek na die stel van die minimum matrikulasietoelatingsvereistes* (Verslag SANSO-127). Pretoria: Departement van Nasionale Opvoeding.
- Elliott, R. & Strenta, A.C. (1988). Effects of improving the reliability of the GPA on prediction generally and on comparative predictions for gender and race particularly. *Journal of Educational Measurement*, 25, 333-347.
- Ferguson, G.A. (1966). *Statistical analysis in psychology and education* (2nd ed.). New York: McGraw-Hill.
- Goldman, R.D. & Slaughter, R.E. (1976). Why college grade point average is difficult to predict. *Journal of Educational Psychology*, 68(1), 9-14.
- Goldman, R.D. & Widawski, M.H. (1976). A within-subjects technique for comparing college grading standards: Implications in the validity of the evaluation of college achievement. *Educational and Psychological Measurement*, 36, 381-390.
- Gulliksen, H. (1950). *Theory of mental tests*. New York: Wiley.
- Linn, R.L. (1966). Grade adjustments for prediction of academic performance: A review. *Journal of Educational Measurement*, 3, 313-329.
- Lord, F.M. & Novick, M.R. (1968). *Statistical theories of mental test scores*. Reading, Massachusetts: Addison-Wesley.
- Ramist, L., Lewis, C. & McCamley-Jenkins, L. (1994). *Student group differences in predicting college grades: Sex, language and ethnic groups* (College Board Report No. 93-1/ETS R.R. No. 94-27). New York: College Entrance Examination Board.
- Stoker, D.J., Engelbrecht, C.S., Crowther, N.A.S., Du Toit, S.H.C. & Herbst, A. (1985). *Ondersoek na differensiële toelatingsvereistes tot tersiêre onderwysinrigtings*. Pretoria: Raad vir Geesteswetenskaplike Navorsing.
- Strenta, A.C. & Elliott, R. (1987). Differential grading standards revisited. *Journal of Educational Measurement*, 24, 281-291.
- Stricker, L.J., Rock, D.A., Burton, N.W., Muraki, E. & Jirele, T.J. (1994). Adjusting college grade point average criteria for variations in grading standards: A comparison of methods. *Journal of Applied Psychology*, 79(2), 178-183.
- Young, J.W. (1993). Grade adjustment methods. *Review of Educational Research*, 63(2), 151-165.