

WISKUNDEPRESTASIE, VERBALE EN NIE-VERBALE  
REDENERINGSVERMOË BY SWART STANDERD  
5-LEERLINGE: 'N PSIGOMETRIESE ONDERSOEK

deur

HENDRIK FREDERIK CORNELIUS MALAN  
B.Sc (HONS)

Voorgelê luidens die vereistes vir die graad

MAGISTER SCIENTIAE

in die vak

PSIGOLOGIE

aan die

POTCHEFSTROOMSE UNIVERSITEIT VIR CHRISTELIKE  
HOËR ONDERWYS

STUDIELEIER: DR. S. OOSTHUIZEN

JANUARIE 1981

## OPMERKING

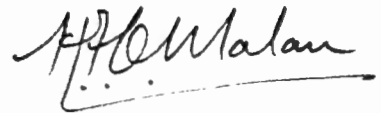
Volgens instruksies mag geen RGN-toetse, handleidings, nasienmaskers en dies meer of dele daarvan, 'n deel van enige verhandeling vorm nie. Dit was derhalwe nie moontlik om items of die aanwysings vir die toepassing van die meetinstrumente in hierdie verhandeling op te neem nie.

Die toetse, handleidings en so meer is egter onderhewig aan die goedkeuring van die President en is by die kantoor van die RGN ter insae, sowel as die oorspronklike statistiese resultate.

Die menings wat in hierdie verhandeling uitgespreek word, is die van die skrywer en is nie noodwendig ook dié van die RGN nie.

Hiermee getuig ek dat hierdie my eie, oorspronklike werk is.

PRETORIA  
JANUARIE 1981



.....  
H.F.C. MALAN

## DANKBETUIGING

Hiermee my opregte dank en waardering aan al die persone en instansies wat op die een of ander wyse behulpsaam was met die voorbereiding en afhandeling van hierdie ondersoek.

In die besonder bedank ek graag

- \* Dr Stanley Oosthuizen wat as studieleier opgetree het, vir sy bekwame leiding en deskundige hulp
- \* Die Raad vir Geesteswetenskaplike Navorsing vir die beskikbaarstelling van die navorsingsgegevens wat in hierdie verhandeling gebruik is
- \* Die Afdeling Sielkundige Dienste van die Departement van Onderwys en Opleiding
- \* Dr J.D. van Staden vir sy belangstelling en gewaardeerde kritiek
- \* My kollegas vir hulle daadwerklike hulp en raad
- \* Die tiksters Truida, Mona en Rona wat hierdie verhandeling in 'n baie kort tydjie so netjies en noukeurig getik het
- \* Meneer Abel Roux vir raad en hulp waar taalprobleme soms opgeduik het.

H.F.C. MALAN

PRETORIA

JANUARIE 1981

# INHOUDSOPGAWE

Bladsy

## HOOFSTUK 1

1	AGTERGROND, WOORDOMSKRYWING, PROBLEEM- EN DOELSTELLING	1
1.1	AGTERGROND	1
1.2	OMSKRYWING VAN BEGRIPPE	3
1.2.1	Prestasietoetse	3
1.2.2	Gestandaardiseerde toetse	3
1.2.3	Aanvangsevaluasie	4
1.2.4	Intelligensie	4
1.2.5	Verbale Redeneringsvermoë	4
1.2.6	Nie-verbale Redeneringsvermoë	4
1.2.7	Voorspellers	5
1.3	PROBLEEMSTELLING	5
1.4	DOELSTELLINGS	6
1.4.1	Wiskundeprestasietoets	6
1.4.2	Hipotesetoetsing	6

## HOOFSTUK 2

2	LITERATUURSTUDIE	8
2.1	INLEIDING	8
2.2	WISKUNDE IN DIE ALGEMEEN	8
2.3	INTELLIGENSIE, VERBALE REDENERING EN NIE-VERBALE REDENERING	11
2.3.1	Inleiding	11
2.3.2	Intelligensie	11
2.3.3	Verbale aspekte	13
2.3.4	Nie-verbale aspekte	13
2.4	AANLEG EN PRESTASIE	14

# INHOUDSOPGAWE

	<u>Bladsy</u>
2.4.1 Aanleg	14
2.4.2 Prestasie	15
2.5 TOETSKONSTRUKSIE	17
2.5.1 Inleiding	17
2.5.2 Beplanning	18
2.5.3 Skryf van toetsitems	19
2.5.4 Itemontleding en -seleksie	21
2.5.4.1 Die doel met itemontleding	22
2.5.4.2 Voordele van itemontleding	22
2.5.4.3 Beperkinge van itemontleding	23
2.5.4.4 Die informasie wat deur itemontleding ver= skaf word	23
2.5.4.5 Itemseleksie	26
2.5.5 Norms	28
2.6 GELDIGHEID	29
2.6.1 Inhoudsgeldigheid	31
2.6.2 Voorspellingsgeldigheid	32
2.6.3 Gelyktydige geldigheid	33
2.7 BETROUBAARHEID	34
2.7.1 Inleiding	34
2.7.2 Omskrywing van Betroubaarheid	36
2.7.3 Bepaling van toetsbetroubaarheid	37
2.7.3.1 Algemeen	37
2.7.3.2 Die bepaling van toetsbetroubaarheid met behulp van item- en toetsstatistieke	38
2.7.4 Interpretasie van Betroubaarheidskoëffisiënte	38

# INHOUDSOPGAWE

Bladsy

## HOOFSTUK 3

3.	STANDAARDISERING VAN 'N WISKUNDEMEETMIDDEL	41
3.1	AGTERGROND	41
3.1.1	Inleiding	41
3.1.2	Doel met die toets	41
3.1.3	Rasionaal	42
3.2	BESKRYWING VAN DIE TOETS	42
3.2.1	Algemeen	42
3.2.2	Toets 1: Parate Kennis	44
3.2.3	Toets 2: Begrip	45
3.2.4	Toets 3: Woordsomme	46
3.2.5	Oefenvoorbeelde	47
3.2.6	Antwoordblad	47
3.3	STANDAARDISERING	48
3.3.1	Ondersoekopset by die eksperimentele toepassing	48
3.3.2	Beskrywing van steekproef	49
3.3.3	Toepassing van die toets	50
3.3.4	Bespreking van itemontleding	52
3.3.5	Itemseleksie	56
3.3.6	Samestelling van die toets vir die normbepaling	58
3.3.6.1	Ondersoekopset by normbepaling	58
3.3.6.2	Vorms A en B	58
3.3.6.3	Antwoordblad	61
3.3.7	Beskrywing van normgroep	63
3.3.8	Toepassing vir normbepaling	66
3.3.8.1	Toetsmateriaal	66
3.3.8.2	Inleidende woorde	66

## INHOUDSOPGAWE

	<u>Bladsy</u>	
3.3.8.3	Uitdeel van toetsmateriaal	66
3.3.8.4	Invul van die leerlinge se persoonlike besonderhede	66
3.3.8.5	Oefenvoorbeelde	66
3.3.8.6	Aanwysings vir toetse	66
3.3.8.7	Toetstye	67
3.3.9	Bespreking van resultate	67
3.3.9.1	Inleiding	67
3.3.9.2	Bespreking	68
3.3.9.3	Bepaling van norms	71
3.3.9.4	Interpretasie van die normpunte en die gebruik van normtabelle	76
3.4	STATISTIESE EIENSKAPPE VAN DIE TOETS	81
3.4.1	Gemiddelde en Standaardafwyking	81
3.4.1.1	Inleiding	81
3.4.1.2	Steekproefgrootte	82
3.4.1.3	Gemiddeldes	83
3.4.1.3	Standaardafwyking	84
3.4.1.5	Skeefheid	85
3.4.1.6	Kurtose	86
3.4.2	Verskille in die gemiddelde wiskundeprestasie van verskillende groepe	87
3.4.2.1	Inleiding	87
3.4.2.2	Verskille in die gemiddelde wiskundeprestasie ten opsigte van geslag	89
3.4.2.3	Verskille in die gemiddelde wiskundeprestasie ten opsigte van gebied	90
3.4.2.4	Verskille in die gemiddelde wiskundeprestasie ten opsigte van taal	91

## INHOUDSOPGAWE

	<u>Bladsy</u>
3.4.3 Geldigheid	95
3.4.3.1 Formule om geldigheid te bereken	95
3.4.3.2 Voorspellingsgeldigheid	96
3.4.3.3 Gelyktydige geldigheid	98
3.4.3.4 Inhoudsgeldigheid	98
3.4.4 Betroubaarheid	99
3.4.4.1 Kuder-Richardsonformule 8	99
3.4.4.2 Kuder-Richardsonformule 21	100
3.4.5 Standaardmetingsfoute	100
3.5 GEBRUIK VAN DIE TOETS	101
3.5.1 Praktiese gebruik van die toets	101
3.5.2 Verdere navorsing	102

## HOOFSTUK 4

4 METODE BY DIE TWEDE GEDEELTE VAN HIERDIE ONDER= SOEK	103
4.1 INLEIDING	103
4.2 ONDERSOEKGROEP	103
4.3 STEEKPROEF	103
4.4 MEETINSTRUMENTE	104
4.4.1 Wiskundeprestasietoets	104
4.4.2 Skolastiese Aanlegtoets vir Swart leerlinge (SATB)	105
4.4.2.1 Inleiding	105
4.4.2.2 Patroonvoltooiing	105
4.4.2.3 Verbale Redenering	106

## INHOUDSOPGAWE

	<u>Bladsy</u>
4.4.4 Betroubaarheid	107
4.5 DIE TOEPASSINGSPROSEDURE	107
4.6 DIE ANTWOORDBLAD	108

### HOOFSTUK 5

5 RESULTATE, BESPREKING EN GEVOLGTREKKING TEN OPSIGTE VAN DIE TWEEDE GEDEELTE VAN HIERDIE ONDERSOEK	109
5.1 INLEIDING	109
5.2 ONDERSOEK	109
5.2.1 Korrelasies tussen Wiskunde, Verbale Redenering en Nie-Verbale Redenering	109
5.2.2 Gemiddelde Wiskundeprestasie ten opsigte van geslag	112
5.2.3 Intelligensie as voorspeller van wiskundeprestasie	115

### HOOFSTUK 6

6 SAMEVATTING, BETEKENIS VAN DIE ONDERSOEK EN AANBEVELINGS	118
6.1 SAMEVATTING	118
6.2 BETEKENIS VAN DIE ONDERSOEK	119
6.3 AANBEVELINGS	119
BYLAES	121
BRONNELYS	132
SUMMARY	139

\*\*\*

## LYS VAN TABELLE EN FIGURE

<u>Tabel nr.</u>		<u>Bladsy</u>
3.1	SPESIFIKASITABEL VAN TOETS 1: PARATE KENNIS	43
3.2	SPESIFIKASITABEL VAN TOETS 2: BEGRIP	43
3.3	SPESIFIKASITABEL VAN TOETS 3: WOORD= SOMME	43
3.4	TOEPASSINGSDOEL, MEETINSTRUMENTE, TOETSE EN TOETSGROEP	49
3.5	ITEMSTATISTIEK VAN VORM A VIR ITEMS 4 EN 5	53
3.6	INTERKORRELASIES VAN TOETSE	59
3.7	GESKATTE STATISTIESE GEGEWENS VAN VORM A EN VORM B	60
3.8	TOEPASSINGSDOEL, MEETINSTRUMENTE, TOETSE EN TOETSGROEP	60
3.9	VORMS A en B: SPESIFIKASITABEL VAN TOETS 1: PARATE KENNIS	62
3.10	VORMS A EN B: SPESIFIKASITABEL VAN TOETS 2: BEGRIP	62
3.11	STEEKPROEF VOLGENS TAALGROEPE	64
3.12	STEEKPROEF VOLGENS STREKE	65
3.13	WERKLIKE GETAL LEERLINGE GETOETS VOLGENS GESLAG	65
3.14	TYD VIR NORMBEPALINGTOEPASSING	67
3.15	VORM A: OMSETTING VAN ROUTELLINGS NA STANDAARDTELLINGS	74
3.16	VORM B: OMSETTING VAN ROUTELLINGS NA STANDAARDTELLINGS	75
3.17	INTERPRETASIE VAN STANEGES	76
3.18	STANDAARDFOUT VAN DIE VERSKIL TUSSEN TWEË T-TELLINGS	80
3.19	STATISTIESE EIENSKAPPE VAN DIE TOETS	82
3.20	BETEKENISVOLHEID VAN VERSKILLE TUSSEN DIE GEMIDDELDE PRESTASIE TEN OPSIGTE VAN GESLAG	89

3.21	BETEKENISVOLHEID VAN VERSKILLE TUSSEN DIE GEMIDDELDE PRESTASIE TEN OPSIGTE VAN GEBIED	90
3.22	BETEKENISVOLHEID VAN VERSKILLE TUSSEN DIE GEMIDDELDE PRESTASIE TEN OPSIGTE VAN TAAL	91
3.23	BETEKENISVOLHEID VAN DIE VERSKILLE TUSSEN DIE GEMIDDELDE PRESTASIES VAN DIE VERSKILLENDE TAAFGROEPE	94
3.24	INTERKORRELASIES VAN TOETSE (BEGIN 1978) EN EKSAMENPUNTE (EINDE 1978) VAN AL DIE LEERLINGE (N = 430)	96
3.25	INTERKORRELASIES VAN TOETSE (BEGIN 1979) EN EKSAMENPUNTE (EINDE 1978) VAN AL DIE LEERLINGE (N = 324)	97
4.1	TEORETIESE STEEKPROEF VOLGENS SKOOL EN GESLAG	104
4.2	BETROUBAARHEID VAN DIE TOETSE	107
5.1	KORRELASIES TUSSEN WISKUNDE, VERBALE REDENERING EN NIE-VERBALE REDENERING (UNIVERSUM)	110
5.2	GEMIDDELDE WISKUNDEPRESTASIE VAN SEUNS (N = 242) EN MEISIES (N = 238)	113
5.3	KORRELASIE TUSSEN WISKUNDEPRESTASIE EN INTELLIGENSIE	115

## FIGUUR

1	GRAFIEK OM ROUTELLINGS NA PERSENTIEL= RANGE EN T-TELLINGS TE HERLEI	72
---	--	----

# BYLAES

<u>Bylae</u>		<u>Bladsy</u>
A-1	NAME VAN SKOLE IN EKSPERIMENTELE TOEPAS= SING	121
A-2	NAME VAN SKOLE IN NORMBEPALINGTOEPASSING	122
B-1	VORM a: STATISTIESE GEGEWENS VAN DIE EKSPERIMENTELE TOEPASSING	123
B-2	VORM b: STATISTIESE GEGEWENS VAN DIE EKSPERIMENTELE TOEPASSING	124
B-3	VORM c: STATISTIESE GEGEWENS VAN DIE EKSPERIMENTELE TOEPASSING	125
B-4	VORM d: STATISTIESE GEGEWENS VAN DIE EKSPERIMENTELE TOEPASSING	126
B-5	VORM A: STATISTIESE GEGEWENS VAN DIE GESELEKTEERDE ITEMS	127
B-6	VORM B: STATISTIESE GEGEWENS VAN DIE GESELEKTEERDE ITEMS	128
B-7	VORM A: STATISTIESE GEGEWENS VAN DIE GESELEKTEERDE ITEMS VAN DIE FINALE TOETS	129
B-8	VORM B: STATISTIESE GEGEWENS VAN DIE GESELEKTEERDE ITEMS VAN DIE FINALE TOETS	130
B-9	VERDELING VAN PUNTE DEUR DIE LEERLINGE BEHAAL VIR VORM A (TOETSE 1 EN 2) (N = 495)	131

\*\*\*

## HOOFSUK 1

### 1 AGTERGROND, WOORDOMSKRYWING, PROBLEEM- EN DOELSTELLING

#### 1.1 AGTERGROND

Die volgende bewering is deur die Departement van Inligting (-, p. 3) gemaak: *With intelligence and perseverance Black men and women can today climb the ladder of success to important and responsible positions ..... A new purpose, a new drive, a new sense of adventure marks the lives of today's generation of Blacks.*

Hiertoe word die Swartman op vele maniere bygestaan en gehelp deur onderwys. Die Departement van Onderwys en Opleiding doen alles in sy vermoë om, ten spyte van die geweldige toename in die leerlingtal in skole en daarmee gepaardgaande probleme, die opvoeding en onderrig van die Swartkind as baie belangrik te beskou en uit te bou tot 'n hoë standaard wat goed sal vergelyk met dié van ander bevolkingsgroepe.

Die volgende syfers dui aan dat daar tereg van 'n onderwysontploffing gepraat kan word. In 1976 was daar, volgens die Departement van Onderwys en Opleiding se Jaarverslag (1977, p. 50) 2 898 754 leerlinge in die primêre skole in die Republiek van Suid-Afrika ingeskryf wat 'n vermeerdering van 4,1 persent op die totaal van 1975 was. Hierdie getal behoort, behalwe deur die normale aanwas, verder toe te neem as gevolg van die nuwe maatreëls wat in Januarie 1977 in werking getree het, naamlik dat daar van ouers wat hulle kinders in substanderd A in skole in stedelike Swart woongebiede in die Blankegebied ingeskryf het, verwag word om skriftelik te onderneem om hulle kinders vir minstens 4 jaar op skool te hou. Dit sal ook meebring dat daar met verloop van 'n aantal jare 'n groter toename sal wees in die getal leerlinge wat tot standerd 6 toegelaat sal word. Volgens die 1976-jaarverslag (1977, p. 58) was die totale getal Swart leerlinge aan sekondêre skole 389 066, wat 'n vermeerdering van 54,1 persent op die getalle van 1975 was.

Daar kan dus gesien word dat die onderwysontploffing waarna Swanepoel (1975, p. 1) en Chamberlain (1977, p. 1) verwys, reeds verder toegeneem het en nog verder sal toeneem. Tese hiermee ontstaan probleme ten opsigte van voorligting aan Swart leerlinge oor vakkeuses, beroepskeuses en die plek wat hulle in die samelewingsverband moet inneem.

Daar bestaan noue samewerking tussen die Departement van Onderwys en Opleiding en die Suid-Afrikaanse Instituut vir Psigologiese en Psigometriese Navorsing (SAIPPN) by die Raad vir Geesteswetenskaplike Navorsing (RGN). Baie van die voorligting waarna in die vorige paragraaf verwys is, word gedoen met behulp van toetse wat deur die RGN vir die Departement van Onderwys en Opleiding gestandaardiseer is. Die volgende tipes toetse is reeds gestandaardiseer (RGN-katalogus, 1979).

- (a) Aanlegtoetse
- (b) Bekwaamheidstoetse
- (c) Diagnostiese toetse
- (d) Persoonlikheidsvraelyste
- (e) Beroepsbelangstellingsvraelyste

Die Veld wat nog braak lê, is die standaardisering van prestasietoetse. Volgens Van Staden (1975, p. viii) bestaan daar 'n dringende behoefte aan sielkundige en skolastiese prestasietoetse vir die Swart volksgroepe. Die Departement van Onderwys en Opleiding voel ook hierdie behoefte aan en het die RGN genader met 'n spesifieke versoek, naamlik om Wiskundeprestasietoetse te standaardiseer wat vir aanvangsevaluasie in die verskillende standerds gebruik kan word.

Wanneer die algemene peil van 'n leerling se taalvermoë en sy begripsvermoë ten opsigte van bewerkings met getalle verhoog kan word, sal die standaard van opvoeding ook verhoog.

Die gebruik van gestandaardiseerde Wiskundetoetse sal meehelp om hierdie doel te bereik.

Samevattend blyk dit dus dat daar 'n groter gretigheid by die Swartkind bestaan om te wil leer en om sodoende in die snel ontwikkelende Suid-Afrika ook sy plek te kan volstaan. Hiertoe word hy baie gehelp deur die Departement van Onderwys en Opleiding, wat met behulp van die SAIPPN allerlei meetmiddels tot die beskikking van die onderwyser, inspektoraat en inspekteur van skole (sielkundige dienste) stel waarmee die standaard van opvoeding en onderwys verhoog kan word.

Daar word verwag dat hierdie studie en die gebruik van die beoogde Wiskundeprestasietoets, in die praktyk sal meehelp om die standaard van opvoeding en onderrig in Wiskunde in die skole van die Departement Onderwys en Opleiding te verhoog.

## 1.2 OMSKRYWING VAN BEGRIPPE

Duidelikheidshalwe word sommige van die begrippe wat in hierdie ondersoek gebruik gaan word, verder omskryf.

### 1.2.1 Prestasietoetse

Drever (1974, p. 9) omskryf prestasietoetse as volg:

*Tests constructed and standardized to measure proficiency in school subjects.*

Vir die doeleindes van prestasiemeting, waar die prestasiepeil bepaal moet word, kan prestasie gedefinieer word as die graad van sukses of welslae wat bepaal word op 'n sekere kognitiewe terrein.

Prestasie in Wiskunde is afhanklik van 'n leerling se

- (i) kennis en begrip van wiskundige feite, terminologie en beginsels;
- (ii) vaardigheid in wiskundige bewerkings en
- (iii) vermoë om aangeleerde wiskundige kennis en vaardigheid te kan toepas om probleme in nuwe situasies op te los.

### 1.2.2 Gestandaardiseerde toetse

Drever (1974, p. 281) omskryf gestandaardiseerde toetse as volg:

*Tests which have been carefully selected, with reference to their objective, and the group for which they are intended, tried out, and arranged appropriately on the basis of the try-out, for which the method of administration has been exactly and unambiguously detailed, and careful instructions, given with respect to the method of scoring, for which norms of performance have been made available, and of which the consistency and validity have been determined, and are known to be high.* Dit is dus 'n toets wat deur deskundiges gekonstrueer is met standaardinstruksies en dit word toegepas onder standaardomstandighede op 'n verteenwoordigende steekproef van die populasie vir die verkryging van norms.

### 1.2.3 Aanvangsevaluasie

Met 'n aanvangsevaluasietoets word verstaan dat dit 'n meetmiddel is wat daarop ingestel is om 'n breë oorsig te gee van leerlinge se vermoëns in die belangrikste aspekte van die vorige jaar se werk. So 'n toets word uitsluitlik aan die begin van die jaar toegepas.

### 1.2.4 Intelligensie

Dit is nie die bedoeling om hier 'n volledige relaas van die begrip intelligensie te gee nie; trouens daar bestaan uit die literatuur volgens Sonnekus (1958, p. 45) groot meningsverskil oor die begrip en wese van intelligensie. Die begrip intelligensie waarna daar in paragraaf 1.4.2(c) verwys word, is die uitkoms soos gemeet deur die Verbale en Nie-verbale Redeneringstoets.

### 1.2.5 Verbale Redeneringsvermoë

'n Verbale Redeneringsvermoëtoets kan gesien word as 'n toets wat verband hou met die meting van verstandsvermoë om begrippe wat in woorde uitgedruk is, te verstaan, en in verband daarmee verhoudings in te sien, te abstraheer en te veralgemeen.

Van der Walt (1970, p. 318) stel dit kortliks soos volg:  
*Dit meet die vermoë om konstruktief deur middel van taal te dink.*

#### 1.2.6 Nie-verbale Redeneringsvermoë

Volgens Van der Walt (1970, p. 316) word onder Nie-verbale Redeneringsvermoë die volgende verstaan: *Die meting van die abstrakte redenasievermoë, dit is die verstandsvermoë om verhoudings in te sien, reëls te veralgemeen en toe te pas.*

#### 1.2.7 Voorspellers

Volgens Von Mollendorf (1978, p. 72) is voorspelling in die psigometriese opset daarop gebaseer dat waarskynlike uitkomstes op 'n gegewe veranderlike, op grond van bestaande inligting oor 'n ander veranderlike, voorspel word. In hierdie ondersoek gaan die toetse in Verbale Redeneringsvermoë en Nie-verbale Redeneringsvermoë se voorspellingswaarde ondersoek word.

Inligting ten opsigte van die voorspellingsmoontlikhede van 'n (psigologiese) meetinstrument vir spesifieke situasies, word dus by wyse van psigometriese ondersoek verkry. Hierdie inligting word van praktiese waarde wanneer die (psigologiese) voorligter dit in sy voorligtingswerk kan aanwend. Kennis van die sterkte van verband wat daar tussen die prestasie op die onafhanklike veranderlike (voorspeller) en die afhanklike veranderlike (kriterium) bestaan, verleen gevolglik stewigheid aan voorligtingswerk.

### 1.3 PROBLEEMSTELLING

By die bestudering van reeds gestandaardiseerde toetse (soos hierbo genoem) en bestudering van die literatuur oor prestasietoetse het die volgende vrae onwillekeurig ontstaan:

(a) In watter mate is intelligensie 'n voorspeller van wis-

kundeprestasie?

- (b) Is Verbale Redeneringsvermoë 'n beter voorspeller van wiskundeprestasie as Nie-verbale Redeneringsvermoë, of omgekeerd?
- (c) Watter rol speel geslag by wiskundeprestasie?

Daar bestaan toetse wat Verbale en Nie-verbale Redeneringsvermoë meet, maar ongelukkig bestaan daar geen toetse wat wiskundevermoë van Swart leerlinge meet nie.

Daar kon geen literatuur gekry word wat bogenoemde vrae ten opsigte van plaaslike Swartes beantwoord nie. Hierdie toedrag van sake het gelei tot die formulering van hipoteses om hierdie probleme te probeer oplos.

#### 1.4 DOELSTELLINGS

##### 1.4.1 Wiskundeprestasietoets

Hierdie ondersoek bestaan uit twee afdelings. Die een afdeling, wat ook die grootste deel van hierdie ondersoek uitmaak, bestaan uit die saamstel van 'n Wiskundeprestasietoets wat sal dien as 'n meetmiddel wat in die tweede deel van hierdie ondersoek gebruik sal word. Dit impliseer dus dat in hierdie gedeelte daar nie verskillende hipoteses gestel kan word nie.

##### 1.4.2 Hipotesetoetsing

In die tweede gedeelte van hierdie ondersoek gaan na die spesifieke verbande tussen Verbale en Nie-verbale Redeneringsvermoë en Wiskundeprestasie by Swart seuns en Swart meisies gekyk word. Die volgende hipoteses gaan ondersoek word:

- a.  $H_0: \rho_1 = \rho_2$  Die nul-hipotese wat gestel word, is dat daar nie 'n verskil in die voorspellingswaarde van Verbale en Nie-verbale Redeneringsvermoë ten opsigte van wiskundeprestasie is nie.

$H_1: \rho_1 \neq \rho_2$  As alternatiewe hipotese word gestel dat daar wel 'n verskil is in die voorspellingswaarde van Verbale en Nie-verbale Redeneringsvermoë ten opsigte van wiskundeprestasie.

- b.  $H_0: \mu_1 = \mu_2$  'n Tweede nul-hipotese wat gestel word, is dat daar geen verskil tussen die gemiddelde wiskundeprestasie van seuns en meisies op standerd 5-vlak is nie.

$H_1: \mu_1 \neq \mu_2$  As alternatiewe hipotese word gestel dat daar wel 'n verskil tussen die gemiddelde wiskundeprestasie van seuns en meisies op standerd 5-vlak is.

- c.  $H_0: \rho = 0$  'n Derde nul-hipotese wat gestel word, is dat intelligensie nie 'n voorspeller vir wiskundeprestasie is nie.

$H_1: \rho \neq 0$  As alternatiewe hipotese word gestel dat intelligensie 'n voorspeller vir wiskundeprestasie is.

## HOOFSTUK 2

### 2 LITERATUURSTUDIE

#### 2.1 INLEIDING

Soos dit blyk uit hoofstuk 1 bestaan 'n groot deel van hierdie ondersoek uit die standaardisering van 'n Wiskunde-prestasietoets. Toetskonstruksie vervul dus 'n belangrike rol in dié studie en sal 'n groot deel van hierdie hoofstuk uitmaak. Betroubaarheid en Geldigheid is twee belangrike eienskappe van 'n goeie toets en alhoewel dit 'n onderafdeling van toetskonstruksie is, sal dié eienskappe apart bespreek word.

Aangesien die studie die standaardisering van 'n Wiskunde-prestasietoets behels, sal begrippe soos wiskunde, prestasie en aanleg in die algemeen bespreek word.

Deurdadig daar by die tweede gedeelte van hierdie ondersoek na Intelligensie, Verbale- en Nie-verbale Redenering as voorspellers van wiskunde-prestasie gekyk word, gaan hierdie begrippe ook kortliks bespreek word.

#### 2.2 WISKUNDE IN DIE ALGEMEEN

Algemeen gesproke is Wiskunde een van die vakke wat die meeste probleme vir die leerling op skool lewer en as dit nie was dat standerd 10-wiskunde 'n voorvereiste vir toelating tot baie van die studierigtings op universiteit was nie, sou daar baie minder leerlinge gewees het wat Wiskunde as vak neem. Wiskunde is by uitnemendheid dié vak om die leerling se verstandelike vorming en logiese denkvermoë te ontwikkel, dit bied onontbeerlike toerusting vir elke beskaafde mens in die moderne wêreld en dit is 'n haas onmisbare hulpmiddel vir byna elke studierigting of praktiese loopbaan.

Steyn (1960, p. 233) sê in hierdie verband: *Dit is absoluut noodsaaklik dat elke leerling wat enigsins daartoe in staat is, Wiskunde as vak moet kies.*

Dit is egter duidelik dat alle leerlinge nie Wiskunde teen dieselfde tempo kan leer nie, en dat sommige leerlinge vroeër ryp is vir die verwerwing van bepaalde insigte as ander. Van Rooy (1965, p. 243) sê die vraag is blykbaar nie wat die leerlinge kán leer nie, maar wánnêr hulle dit kan leer. Die veiligste weg is blykbaar om te aanvaar dat alle leerlinge (behalwe dié wat duidelike afwykings toon) wiskundige begrippe kan bemeester, die een net gouer as die ander.

Adler (1963, p. 509) waarsku soos volg teen die gebruik om leerstof van sekere leerlinge te weerhou bloot op grond van hulle I.K.'s. *Those who rely on the I.Q. theory assume that most children are not capable of learning much. A natural consequence is that they do not try to teach them much. They offer these children watered-down courses devoid of significant mathematical content. They take pupil failure for granted, and so neither teacher nor child is required to exert the effort that might prevent failure. The major result is cumulative retardation from year to year. The children become less and less prepared to cope with the work of the higher grades. The process of educational decay is selfreinforcing, because it is controlled by a typical feedback loop: We think the children are not capable of learning, so we teach them less. Then they learn less and become less capable of learning. So again we teach them less, and so on ad infinitum. Teachers and pupils become trapped in a vicious circle.* Uit Adler se relaas blyk dit dat die onderwyser in 'n groot mate te blameer is vir die agteruitgang van Wiskunde.

Wat die probleme in Swart skole betref, sê Michau (1978, p. 22) *It is not at high school level, however, that the grounds for the poor achievement are established. The roots of the problem would appear to lie in the cultural background of the child and in his early school years.* Dit gaan vir haar meer om die betekenis van woorde wat verskil in die Swarte se moedertaal wat dan lei tot probleme in die onderrig van Wiskunde.

Vir die skrywer kan die swak wiskunde-uitslae waarna Joan Michau in haar artikel *Problem areas in the acquisition of Mathematical concepts by Black Children in South Africa*, verwys, toegeskryf word aan

- (a) laat identifisering van leerlinge met spesifieke wiskunde probleme en
- (b) die onderrig van Wiskunde.

Die tweede probleem (volgens die navorser) lê daarin dat as gevolg van die groot leerlingtal in Swart Onderwys, daar nie genoeg opgeleide onderwysers (spesifiek in Wiskunde) is, wat Wiskunde as vak kan aanbied nie. Die eerste probleem (volgens die navorser) lê daarin dat die Swart onderwyser nie gestandaardiseerde Wiskundetoetse het om die prestasiepeil van sy leerlinge te bepaal nie en om te diagnoseer met watter aspekte van die sillabus die leerlinge probleme ondervind.

Dit is dan die eerste doel van hierdie studie om 'n gestandaardiseerde Wiskundeprestasietoets daar te stel wat die Swart onderwyser sal help om die swak presteerder uit te wys en ook dié leerling, wat probleme ondervind met sekere aspekte van die sillabus, sodat die onderwyser meer aandag aan hierdie leerlinge kan gee.

## 2.3 INTELLIGENSIE, VERBALE REDENERING EN NIE-VERBALE REDENERING

### 2.3.1 Inleiding

Intelligensiemeting speel 'n groot rol in die skoolsituasie waar dit gebruik word om skoolkinders te klassifiseer en te groepeer. Daar kan dus gesê word dat intelligensiemeting voorspelling van skolastiese prestasie vergemaklik. Grundlingh (1967, p. 8) sê dat hierdie aanduiding van die potensiaal van die kind dit moontlik maak om leiding te gee wat vir die kind die beste geleentehede sal bied vir groei en ontwikkeling.

In hierdie studie gaan dit om te bepaal in watter mate Verbale Redenering, Nie-verbale Redenering en in 'n mindere mate intelligensie voorspellers is van wiskunde-prestasie.

Daar is reeds baie geskryf oor intelligensie en intelligensietoetse en dit is nie die doel van hierdie studie om 'n opsomming daarvan te maak nie. Daar gaan net kortliks gekyk word na die begrippe.

### 2.3.2 Intelligensie

Intelligensie is volgens Lyman (1963, p. 86) 'n abstrakte begrip en ten spyte van baie navorsing oor intelligensie en die baie definisies wat daaraan gekoppel word, is dit moeilik om een algemeen aanvaarbare definisie daar te stel. Guilford (1967, p. 12) beweer dat daar selfs van ongedefinieerde terme in die definisies gebruik gemaak word.

Wechsler (1958, p. 7) definieer intelligensie soos volg: *Intelligence is the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment.*

Drever (1974, p. 141) definieer intelligensie as

- (i) *the capacity to meet novel situations, or to learn to do so, by new adaptive responses, and*
- (ii) *the ability to perform tests or tasks, involving the grasping of relationships, the degree of intelligence being proportional to the complexity, or the abstractness, or both, of the relationships.*

Van die definisies is laasgenoemde die aanvaarbaarste.

Goddard (1920, p. 573) beskou intelligensie as relatief en meen dat daar slegs graadverskille van persoon tot persoon is.

Die teoretiese aspekte van die aard van intelligensie kan funksioneel en struktureel benader word. By die funksionele benadering gaan dit om verklarings van intelligensie wat gegrond is op die waarneming van menslike gedrag soos sy aanpassing by die omgewing, sy vermoë om abstrak te dink. By die strukturele benadering gaan dit om die bestaan van spesifieke groepsfaktore. Voorbeelde is, die twee-faktor-teorie van Spearman, die veelfaktor-teorie van Thorndike en die groepfaktor-teorie waarvan Thurstone 'n voorstander is (Morgan, 1961, pp. 438 - 455).

Die funksionele benadering, naamlik intelligensie as vermoë om nuwe kennis op te neem en om abstrak te dink, word by die opstel van die meeste intelligensietoetse as uitgangspunt geneem.

Die meeste groepintelligensietoetse, soos byvoorbeeld die NSAG, bestaan uit 'n verbale en 'n nie-verbale afdeling.

Ongelukkig bestaan daar tot op hede nie intelligensietoetse vir alle bevolkingsgroepe in Suid-Afrika nie. So bestaan daar byvoorbeeld nie 'n intelligensietoets vir Swart standerd vyf-leerlinge nie. Vir die doel van

hierdie ondersoek word die Verbale en die Nie-verbale re= deneringstoetse uit die Skolastiese Aanlegtoetsbattery vir Swartes vir standerd ses-leerlinge gebruik om hulle intelligensie te meet. Gronlund (1965, p. 234) sê in hierdie verband *the obvious advantage of using both verbal and nonverbal scores is in identifying learning potential at two different levels.*

### 2.3.3 Verbale aspekte

Drever (1974, p. 310) definieer verbale aspekte soos volg: *Characterized by being expressed in words, either orally or graphically.* Dit gaan dus vir hom uitsluitlik om die geskrewe en gesproke woord.

Volgens Morgan (1961, p. 437) word toetse waarvan die instruksies en die beantwoording van die vrae deur middel van die geskrewe woord geskied, verbale toetse genoem.

Volgens Gronlund (1965, p. 234) speel verbale vermoë 'n prominente rol in die aanleer van skooltake en gevolglik sal die verbale telling sukses op skool die beste voorspel. Gronlund sê dat 'n hoë verbale telling dui op die vermoë wat nodig is vir sukses in skoolwerk.

### 2.3.4 Nie-verbale aspekte

Die take wat in 'n nie-verbale toets ingesluit word is gebaseer op *objects on drawings or figures, and require assembly, classification, arrangement, selection, manipulation or some other response. Sometimes even the instructions involve no words, but are given in pantomime* (Ebel, 1972, p. 504).

'n Toets wat uit nie-verbale aspekte bestaan word ook 'n handelingstoets genoem. By sulke toetse is die leeselement heeltemal uitgeskakel. Voorbeelde van sulke toetse is prentvoltooing, prentrangskikking en voorwerpe bymekaar voeg.

Die telling in 'n Nie-verbaletoets dui op 'n bevredigende skatting van 'n leerling se onbekende potensiaal. Gron=

lund (1965, p. 235) sê in hierdie verband *a pupil with above average mental ability who is deficient in reading is apt to have a low verbal score but an above average nonverbal score. Although his chances of immediate success in school work are poor because of his reading deficiency, his chances of future success are good if he is given appropriate remedial work.*

Dit kom dus daarop neer dat die nie-verbale telling as 'n kontrole dien van die leerpotensiaal van die persone met lae verbale tellings.

## 2.4 AANLEG EN PRESTASIE

Vir die doel van hierdie ondersoek sal daar meer aandag gegee word aan prestasie as aan aanleg.

### 2.4.1 Aanleg

Gouws en ander (1979, p. 1) definieer aanleg soos volg: *Die hoofsaaklik aangebore vermoë om in die toekoms in 'n bepaalde taak of handeling te presteer, of 'n vaardigheid te bemeester.*

Alberts (1967, p. 10) onderskryf hierdie definisie deurdat hy aanleg sien as die verstandelike vermoëns waaroor 'n individu moet beskik om 'n sukses te maak van een of ander beroep of studierigting. Remmers en Gage (1955, p. 218) sien aanleg as huidige vermoëns wat dien as voorspellers van toekomstige prestasie. Ahmann (1962, p. 80) sluit ook by hierdie groep aan deurdat hy aanleg sien as *the individual's potential for learning a given skill when he is provided with appropriate instruction.*

Uit die uitsprake van genoemde navorsers blyk dit dat

aanleg te doen het met prestasie in die toekoms. Lötzt (1967, p. 76) vat dit mooi saam deur te sê *aanleg is 'n persoon se kapasiteit om te leer; dit is die basiese vermoë of kombinasie van basiese vermoëns wat aangewend kan word om sukses en prestasie in 'n bepaalde veld aan te dui.*

Al hierdie definisies gee 'n goeie beskrywing van aanleg, maar die omvattendste definisie word gegee deur Owen (1971, p. 6). Hy beskryf aanleg as *die potensiële verstandelike vermoëns, aangebore sowel as verworwe, waaroor die individu op 'n sekere stadium beskik en wat hom in staat stel om sekere bekwaamhede te ontwikkel. Aanlegte, asook ander persoonlikheidseienskappe soos belangstelling, motivering en houding, tesame met opleiding en onderrig, sal bepaal watter peil van bekwaamheid bereik word.*

Gericke (1972, p. 20) vereenselwig hom ook met hierdie definisie omdat dit volgens hom al die relevante aspekte van aanleg insluit.

Vir die doel van hierdie ondersoek beskou die skrywer aanleg as die spesifieke wiskundige vermoë, aangebore en verworwe, waaroor 'n individu moet beskik en wat sal dien as voorspeller van toekomstige wiskundeprestasie.

Aanleg dui dus op spesifieke take, terwyl intelligensie op globale fasette van die individu betrekking het.

#### 2.4.2 Prestasie

Die Psigologie-Woordeboek van Gouws en ander (1979, p. 238) omskryf prestasie soos volg:

- (i) *Die bereiking van die doel van 'n handeling of 'n handelingsreeks, dit wil sê die voltooiing van 'n taak, byvoorbeeld die slaag van 'n eksamen.*

- (ii) *Die produk van 'n doelgerigte handeling of handelingsreeks, byvoorbeeld 'n kunswerk wat voltooi is.*
- (iii) *Die standaard van sukses wat met 'n taak behaal is, byvoorbeeld die mate waarin 'n leerling die leertake van sy standerd bemeester, of die produksie van 'n werknemer.*

Prestasie kan dus gesien word as die suksesvolle bereiking van 'n doelwit wat gebaseer is op verworwe kennis. Prestasie het dus met enigiets wat geleer is of geleer word te make.

Aiken (1971, p. 81) definieer prestasie as *the degree of success or accomplishment in an area of endeavour*. Alhoewel die terrein van prestasie nog nie afgebaken is nie, beskou Gericke (1972, p. 29) hierdie definisie as operasioneel. Hy sê dat met hierdie definisie as basis, die presiese terrein van prestasie afgebaken kan word, byvoorbeeld in 'n skoolvak. Vir die doel van hierdie studie sal Wiskunde die afbakeningsveld wees.

Drever (1974, p. 9) definieer prestasie as *Performance in a standardized series of tests, usually educational*. Hierdie definisie is van direkte toepassing op hierdie ondersoek, want die eerste deel van hierdie ondersoek handel oor die daarstel van 'n gestandaardiseerde Wiskundeprestasietoets en in die tweede gedeelte word die Wiskundeprestasietoets gebruik om die wiskundeprestasie van leerlinge te meet.

Dit laat die vraag ontstaan of prestasie 'n begrip is wat gemeet kan word. Volgens Ebel (1972, p. 37) is dit maklik om saam te stem dat ons nie weet hoe om alle opvoedkundige resultate te meet nie. *But in principle, all important outcomes of education are measurable.*

Hy sê verder dat dit miskien nie meetbaar is met die hui-

dige toetse nie. *But if they are known to be important, they must be measurable.*

As daar merkbare verskille bestaan ten opsigte van sekere vermoëns in persone, *then the achievement is measurable* (p. 38). Die gevolgtrekking kan dus gemaak word dat prestasie meetbaar is. Wiskundeprestasie kan dus met behulp van 'n gestandaardiseerde Wiskundeprestasietoets gemeet word.

Samevattend beskou Grondlund (1965, p. 230) die verskil tussen 'n aanlegtoets en 'n prestasietoets soos volg: *A common distinction is that achievement tests measure what a pupil has learned and scholastic aptitude tests measure his ability to learn new tasks.*

## 2.5 TOETSKONSTRUKSIE

### 2.5.1 Inleiding

Volgens Nunnally (1970, p. 196) was daar tot onlangs nog geen eenstemmigheid wat betref die metodes van toetskonstruksie nie. Volgens hom is dit die gevolg van onvoldoende beginsels ten opsigte van skaling, geldigheid en betroubaarheid. Hy gaan verder deur te sê dat toetskonstruksiemetodes vir meetmiddels wat poog om voorspellingsgeldigheid of konstrukgeldigheid te hê, dieselfde behoort te wees, terwyl verskillende metodes by inhoudsgeldigheid ter sprake is. Die metode van toetskonstruksie wat in die eerste deel van hierdie ondersoek (dit wil sê die standaardisering van Wiskundeprestasietoets) gevolg sal word, sal meer op inhoudsgeldigheid betrekking hê. Dit sluit aan by Nunnally (1967, p. 237) wat sê: *The achievement test is the most obvious example of a measure that requires content validity.*

## 2.5.2 Beplanning

Volgens Heinichen (1972, p. 126) moet die toetsopsteller duidelikheid hê oor dit wat hy wil meet. Dit is nie voldoende om 'n toets op te stel met 'n vaag opgestelde mening dat dit op iets goeds sal uitloop nie.

Helmstadter (1964, p. 157) sê: *The first step in the construction of a test or measuring device is one of the most important yet most often neglected. It requires careful delimitation and breakdown of the area or trait involved. If the measure is one of Achievement, the process is the straightforward development of a table of specifications.*

Nunnally (1970, p. 198) sê dat beplanning die volgende insluit;

- (a) Die tipe item wat geskryf moet word,
- (b) Die aantal items wat by elke toepassing in elke afdeling van die toets gebruik gaan word en voorbeelde van die items,
- (c) Die tyd wat dit sal neem om die toets toe te pas,
- (d) Die manier waarop dit toegepas sal word,
- (e) Hoe dit nagesien sal word,
- (f) Die soorte norms wat gebruik sal word.

In aansluiting hiermee kan genoem word dat die rasionaal en doel van die toets vooraf baie duidelik uiteengesit moet word.

Wanneer die beplanning van die toets opgestel is, moet dit aan vakkeners, opvoedkundiges en sielkundiges voorgelê word vir kommentaar. Nadat die kommentare in die beplanning ingewerk en weer goedgekeur is, kan met die volgende stap, naamlik die skryf van die items van die toets begin word,

Samevattend word die hoofdoel van toetsbeplanning gesien as die uiteensetting van die inhoud van die instrument wat opgestel moet word. Deur 'n deeglike en volledige beplanning daar te stel, word een van die belangrikste eienskappe van 'n prestasietoets, naamlik inhoudsgeldigheid, verseker.

### 2.5.3 Skryf van toetsitems

'n Goeie toetsbeplanning dra by tot goeie toetsitems, maar *unless items are skillfully written, the plan never materializes* (Nunnally, 1970, p. 199). Dit is 'n kuns wat min mense bemeester, want daar word volgens Chamberlain (1975, p. 14) groot eise aan die toetsopsteller gestel wat sy oorspronklikheid, aanleg, kennis, verwante ervaring en belangstelling betref. Hy moet goed bekend wees met die vereistes vir 'n goeie item,

'n Toetsitem verwys na 'n enkele eenheid van 'n toets. Dit is 'n enkele vraag in 'n toets wat uit baie vrae bestaan. Wat die veelkeusige items betref, sluit die term item ook die moontlike antwoorde in.

In dié ondersoek word gebruik gemaak van die veelkeusige item omdat dit ekonomies (ten opsigte van nasientyd) en maklik is om toe te pas en tellings daaraan toe te ken. Die metode kan volgens Nunnally (1964, p. 239) feitlik alles akkuraat meet en dit korreleer hoog met 'n opsteltoets oor dieselfde onderwerp.

Gronlund (1965, p. 140) sê die volgende in verband met die veelkeusige item: *The multiple-choice item can measure at both the knowledge and understanding levels ..... It is also free of many of the limitations of other forms of objective items.*

'n Veelkeusige item bestaan uit 'n stam (dit wil sê, die gedeelte waarin die probleem gestel word) en 'n paar keuses. Die keuses verwys na die moontlike antwoorde in 'n toetsitem. Die verkeerde antwoorde word die afleiers van die item genoem. 'n Vyfkeusige item bestaan dus uit vier afleiers en die antwoord. Verskillende skrywers soos byvoorbeeld Gronlund (1965, pp. 140 - 159), Nunnally (1970, pp. 199 - 201), Ebel (1972, pp. 191 - 224) en Helmstadter (1964, pp. 161 - 162) gee goeie wenke vir itemskryf,

Die belangrikste hiervan is die volgende:

- (a) Die vraag of aanwysing moet so duidelik as moontlik geformuleer wees,
- (b) Ingewikkelde en lang bewoording behoort vermy te word,
- (c) Dit moet beantwoord aan al die vereistes om 'n redelike basis te voorsien vir die keuse van response.
- (d) Bevoegde kritici moet eenstemmig wees oor die respons wat veronderstel is om reg te wees.
- (e) Al die afleiers moet aanvaarbaar wees vir persone wat nie kennis dra van die feite wat ter sprake is nie,
- (f) Afleiers wat mekaar oorvleuel of 'n herhaling van dieselfde gegewens bevat, behoort vermy te word.
- (g) Al die keuses behoort grammatikaal te pas by die stam.
- (h) Die afleiers behoort nie langer of korter as die korrekte antwoord in 'n spesifieke item te wees nie,
- (i) Stereotipe of handboekgerigte formulering behoort vermy te word behalwe in gevalle waar memorisering getoets word.

- (j) Byvoeglike naamwoorde soos "altyd" en "nooit" behoort nie in verkeerde keuses gebruik te word nie.
- (k) Korrekte antwoorde se posisies moet verskil en moet ook nie sekere patrone vorm nie,
- (l) Strikvrae (doelbewuste misleidende vrae) wat 'n item onnodig moeilik maak, behoort vermy te word,
- (m) Lang en ingewikkelde bewerings vol kwalifiserende frases behoort nie as items gebruik te word nie,
- (n) Woorde wat in elke afleier herhaal word, behoort, indien moontlik, deel van die stam uit te maak,
- (o) Afleiers behoort in toepaslike gevalle in logiese orde geplaas te word,
- (p) Die dubbele ontkennde vorm behoort nie in die stam gebruik te word nie,

Met in agneming van bogenoemde punte, tesame met 'n deeglike bestudering van die sillabusse, behoort dit vir die toetsopsteller duidelik te wees aan watter vereistes goeie items behoort te beantwoord,

#### 2.5.4 Itemontleding en -seleksie

Itemontleding is een van die prosesse in die standaardisering van 'n toets en behels die insameling en berekening van soveel gegewens as moontlik oor die items van 'n toets. Swart (1978, p. 1) wys daarop dat itemseleksie nie 'n spesifieke proses is wat noodwendig op 'n bepaalde tydstip in die konstruksie van 'n toets uitgevoer word nie. Dit kan volgens hom as 'n deurlopende proses beskou word wat reeds begin met die opstel van 'n rasionaal vir die toets en sy hoogtepunt bereik met die statistiese verwerking van die informasie wat met itemontleding ingesamel is en die seleksie wat daarop volg.

#### 2.5.4.1 Die doel met itemontleding

Heinichen (1972, p. 1) vat die doelstellings met itemontleding soos volg saam:

- (a) Om objektiewe informasie in verband met die item wat geskryf is te verkry waardeur die subjektiewe oordeel van die toetsopsteller in verband met die keuse van so 'n item gekontroleer kan word,
- (b) Om sodanige informasie in verband met items in te win dat die toets wat uiteindelik daaruit saangestel sal word, bepaalde kenmerke ten opsigte van moeilikheidsgraad, betroubaarheid en geldigheid sal hê.
- (c) Om gebreke in items te ontdek wat reggestel kan word sodat van swak items weer redelik goeie items gemaak kan word,
- (d) Om die geskikste items vir 'n toets te kan bepaal.

#### 2.5.4.2 Voordele van itemontleding

'n Toets kan op doeltreffende wyse verbeter word deur die ontleding van 'n leerling se toetsresponse in 'n objektiewe toets. *Item analysis indicates which items may be too easy or too difficult and which may fail for other reasons to discriminate clearly between the better and the poorer examinees* (Ebel, 1972, p. 383). Hierdie items wat te maklik of te moeilik is, kan dan geëlimineer word. Met itemontleding word die gebreke in 'n item ondersoek en reggestel as dit reggestel kan word. Al die items wat te moeilik of te maklik is, hoef dan nie geëlimineer te word nie.

Guilford (1966, p. 417) sê itemontleding *provides the opportunity to check upon the test writer's subjective judgement in selecting the items to compose the*

*test*. Subjektiewe selektering van items word dus gekontroleer.

#### 2.5.4.3 Beperkinge van itemontleding

Anstey (1948, p. 167) beweer dat dit nie die rekenaar is wat 'n item goed of sleg maak nie, want die item kom van die skrywer af. Geen tegniek kan van 'n swak item 'n goeie een maak nie.

Davis (1963, p. 266) sluit hierby aan deur te sê dat die gebruik van die rekenaar tot gevolg het dat die skryf en redigering van items, wat eerder 'n kuns as 'n tegniek is, verwaarloos word. Guilford (1966, p. 418) beweer dat die regte tipe items vir 'n toets nie deur die rekenaar bepaal word nie, maar slegs verkry word deur deeglike beplanning en vooraf besinning.

Volgens Swart (1978, p. 3) is itemontleding bloot 'n gids of hulpmiddel en kan nie die oorspronklikheid, moeite en vaardigheid asook die subjektiewe oordeel van die toetsopsteller vervang nie - dis slegs aanvullend.

Ten slotte sê Ahmand en Glock (1967, p. 199) dat die resultate wat met itemontleding behaal word relatief is en onder andere beïnvloed word deur die toets waarvan die item 'n deel uitmaak, die wyse waarop dit toegepas word en die groep waarop dit toegepas word.

#### 2.5.4.4 Die informasie wat deur itemontleding verskaf word

Volgens Helmstadter (1964, p. 163) is daar ten minste drie en twintig verskillende metodes van itemontleding. Dit dien geen doel om van die metodes hier te bespreek nie en slegs die verkreë informasie uit die metode wat gevolg is, gaan in hierdie ondersoek bespreek word.

(a) Die moeilikheidswaarde

Helmstadter (1964, p. 163) definieer die moeilikheidswaarde as *the proportion of persons answering each item correctly*. Hierdie metode is die eenvoudigste en word volgens Guilford (1954, p. 418) gebruik wanneer die itemtelling 0 of 1 kan wees, dit wil sê of verkeerd of reg.

Die formule vir die bepaling van die proporsie toetslinge wat 'n item korrek het, is die volgende:

$$p_i = \frac{N_r}{N_a}$$

waar

$p_i$  = die moeilikheidswaarde van die i-de item

$N_r$  = die getal toetslinge wat die item reg het

$N_a$  = die getal toetslinge wat die item aangedurf het,

Conrad (1945, p. 1) druk die verhouding as 'n persentasie uit en noem dit dan die maklikheidswaarde van die item, dit wil sê, hoe hoër die indeks, hoe makliker is die item.

Die moeilikheidswaardes wat in hierdie ondersoek verkry is, word in hoofstuk 3 bespreek.

(b) Die diskriminasiewaarde

Die diskriminasiewaarde dui die mate aan waarin 'n item tussen goeie en swak toetslinge diskrimineer. Die goeie toetsling kry die item reg en dit dra by tot 'n hoër totaalstelling as in die

geval van die swak toetsling wat die item verkeerd het en 'n lae telling kry (Swart, 1978, p. 10).

Gronlund (1965, p. 211) sê 'n item diskrimineer in 'n positiewe rigting as meer leerlinge in die boonste groep die item reg het. 'n Positiewe diskriminasie dui daarop dat die item in dieselfde rigting diskrimineer as die totale toetstelling. Hy skryf verder: *The discriminating power of an achievement test item refers to the degree to which it discriminates between pupils with high and low achievement.*

Die formule om die punt-twee-reekskorrelasie tussen die item en die totaalstelling van die toets te bereken, is die volgende:

$$r_{it} = \frac{\bar{X}_r - \bar{X}_t}{S_t} \cdot \sqrt{\frac{p}{q}}$$

waar

$r_{it}$  = die korrelasie van die item met die totaalstelling

$\bar{X}_r$  = die gemiddelde van almal wat die item reg het

$\bar{X}_t$  = die gemiddelde van die hele groep

$S_t$  = die standaardafwyking in die totaalstellings van die hele groep

$p$  = die moeilikheidswaarde van die item en

$q$  =  $1 - p$

Vergelyk ook paragraaf 3.3.9.2(b) vir 'n bespreking

van die diskriminasiewaardes wat in hierdie ondersoek verkry is.

(c) Algemeen

Ander informasie wat gebruik word, is

- (i) die toetslinge wat die item oorgeslaan het;
- (ii) die toetslinge wat nie 'n item bereik het nie;
- (iii) informasie in verband met die toets as geheel, soos die betroubaarheid en die metingsfout, asook die gemiddelde en die standaardafwyking van elke kriterium wat vir die berekening van die diskriminasiewaardes gebruik is (vergelyk hoofstuk 3 in hierdie verband).

2.5.4.5 Itemseleksie

Volgens Nunnally (1967, p. 243) moet by itemseleksie meer klem gelê word op die korrelasie van 'n item met die totale tellings as op die moeilikheidswaarde van items. Vir itemseleksie moet 'n groot aantal items op 'n groot aantal persone, wat verteenwoordigend is vir die individue op wie die finale toets toegepas sal word, geneem word. Nunnally (1967, p. 232) sê daar moet minstens twee keer soveel items wees as wat in die finale toets sal voorkom en items moet toegepas word op minstens driehonderd persone - verkieslik eenduisend of meer. Tensy daar minstens vyf keer soveel persone is as items, mag die resultate misleidend wees. Dit is in ooreenstemming met Guilford (1954, p. 418).

Na die items toegepas is, word die diskriminasiewaardes en moeilikheidswaardes tesame met die ander informasie wat in die vorige paragraaf beskryf is, bereken.

Heinichen (1972, p. 198) gee die volgende stappe van itemseleksie wat ook in hierdie ondersoek gevolg is:

- (a) Alle items word eers gekeur op grond van die gemiddeldes van die betrokke groep wat die verskillende afleiers gekies het. As daar êrens 'n tekortkoming in die item te bespeur is, word dit voorlopig verwerp.
- (b) Die oorblywende items word nou op grond van die diskriminasiewaardes gekeur. Vir hierdie ondersoek is  $r_{it}$ -waardes bokant 0,2 geneem (vergelyk hoofstuk 3).
- (c) Die moeilikheidswaardes van die items word dan in aanmerking geneem. Uit 'n statistiese oogpunt sal die ideale item om te gebruik, die een met 'n moeilikheidswaarde van 50 persent wees. Prakties is dit natuurlik onmoontlik om slegs items met moeilikheidswaardes van 50 persent te selekteer. Vir die doel van hierdie ondersoek is daar gepoog om die gemiddelde moeilikheidswaarde van die items so na as moontlik aan 50 persent te kry.
- (d) Indien daar nie genoeg geskikte items oorbly om die finale toets te kan saamstel nie, kan teruggegaan word na daardie items wat voorlopig verwerp is. Hulle mag dalk so verbeter word dat hulle wel vir gebruik geskik is.
- (e) Die volgende stap, volgens Nunnally (1967, p. 243), is om 'n frekwensieverdeling op te stel. As dit "bevredigend" (normaal) is, is selektering van items voldoende. As verspreiding skeef is aan die hoë kant, is die toets te moeilik. Van die makliker items wat goed gekorreleer het met die totaaltellings, kan nou van die moeiliker items vervang.

As verspreiding skeef is aan die laer kant, moet moeiliker items die maklikes vervang.

Daar moet altyd gemik word na die normale. Nunnally noem dit die "cut-and-try"-metodes om verspreiding simmetries te maak.

Ten slotte word die items van maklik na moeilik gerangskik om die toetsling te bemoedig.

#### 2.5.5 Norms

Nunnally (1967, pp. 244 - 245) sê die volgende in hierdie verband: In die standaardisering van prestasietoetse is een van die belangrikste stappe die bepaling van norms. Dit word byvoorbeeld verkry deur middel van tellings van 'n verteenwoordigende steekproef toetslinge van oor die hele land. Die telling van 'n spesifieke toetsling kan vergelyk word met die totaal tellings van toetslinge oor die hele land. Berekening van norms behels feitlik net soveel werk as die konstruksie van 'n toets. Om verteenwoordigend te wees moet honderde getoets word. Statistiese analises word dan aangewend om finale norms te bekom. Dit word gewoonlik uitgedruk in die vorm van standaard tellings en persentiele. 'n Persentiel dui die persentasie persone in die normgroep aan wat minder as 'n spesifieke telling behaal het. Persentiele is makliker om te interpreteer as omgeskakelde standaard tellings.

Ten slotte kan dus gesê word dat norms 'n transformasie is van routellings na 'n meer betekenisvolle skaal wat verkry word uit die resultate van 'n verteenwoordigende steekproef en dat dit 'n aanduiding gee van die individuele relatiewe posisie binne 'n normatiewe steekproef.

Dit dui ook die vergelyking van 'n individuele tellings in verskillende toetse aan asook van die tellings van verskillende individue in dieselfde toets. Die bereke-

ning van norms in hierdie ondersoek word in hoofstuk 3 bespreek.

## 2.6

### GELDIGHEID

#### Inleiding

By die standaardisering van 'n meetmiddel is een van die belangrikste aspekte die geldigheid daarvan. Die begrip geldigheid is nie los te maak van 'n toets nie. Huysamen (1978, p. 71) definieer die geldigheid van 'n meetinstrument as *die mate waarin dit aan sy doel beantwoord*. Hierin word hy ondersteun deur Lyman (1963, p. 25) wat geldigheid sien as *the ability of the test to do the job we want it to do*.

Fourie (1967, p. 148) sluit by hulle aan deur daarop te wys dat die begrip geldigheid met die doel van 'n toets in verband gebring moet word. Volgens Gericke (1972, p. 108) kom dit dus daarop neer dat 'n toets in een geval 'n hoë mate van geldigheid besit terwyl dit in 'n ander geval nie 'n hoë mate van geldigheid besit nie. Dit hang net af van die doel waarvoor die toets gebruik moet word.

Brown (1970, p. 133) beskryf geldigheid op funksionele wyse soos volg: *The validity of a test is defined either by (i) the extent to which the test measures the hypothesized underlying trait, construct or factor or (ii) the relationship between test scores and some extra-test interior measure*.

Hy koppel sy definisie aan tipes geldigheid, byvoorbeeld inhoudsgeldigheid, kriteriumverwante geldigheid en konstruktgeldigheid. In aansluiting by Brown word daar by die bestudering van die literatuur gevind dat daar

in wye verskeidenheid indelings van geldigheid bestaan, onder andere:

(a) Komitee van die American Psychological Association (1966)

- Kriteriumverwante geldigheid
- Inhoudsgeldigheid
- Konstruktgeldigheid

(b) Anastasi (1964, pp. 135 - 152)

- Inhoudsgeldigheid
- Voorspellingsgeldigheid
- Gelyktydige geldigheid
- Konstruktgeldigheid

(c) Helmstadter (1964, pp. 87 - 156)

- Inhoudsgeldigheid
- Empiriese geldigheid
- Konstruktgeldigheid

(d) Stilson (1966, pp. 126 - 129)

- Eksterne geldigheid
- Interne geldigheid

Bestudering van bostaande indeling toon dat daar op enkele uitsonderings na, vier algemene erkende tipes geldigheid, naamlik inhoudsgeldigheid, gelyktydige geldigheid, voorspellingsgeldigheid en konstruktgeldigheid bestaan.

Vir die doel van hierdie ondersoek word die volgende drie geldigheidstipes, naamlik inhoudsgeldigheid, gelyktydige geldigheid en voorspellingsgeldigheid bespreek.

## 2.6.1 Inhoudsgeldigheid

### Omskrywing

Volgens Anastasi (1971, p. 100) behels inhoudsgeldigheid die sistematiese ondersoek van die toetsinhoud om te bepaal of dit 'n verteenwoordigende steekproef is van die gedragsgebied wat ondersoek word.

Volgens Brown (1970, p. 134) behels inhoudsvalidering die bepaling van die toerekenbaarheid van die steekproefneming van items vir die universum van potensiële items.

Uit bogenoemde blyk dat inhoudsgeldigheid te make het met 'n evaluering van die inhoud van 'n toets of vraelys. Die evaluering is 'n subjektiewe evaluering en word in die reël nie in die vorm van 'n syfer of simbool uitgedruk nie.

Madge (1976, p. 4) skryf oor die evaluering van inhoudsgeldigheid die volgende: *Hoewel die evaluering van inhoudsgeldigheid die verantwoordelikheid van die toetsopsteller is, is dit noodsaaklik dat hy die items aan 'n paneel van bekwame beoordelaars voorlê. Uiteindelik vermag hy dat sy instrument deur ander verstaan en aanvaar word en dus is dit essensieel dat hy reeds in 'n vroeë stadium van toetsontwikkeling die opinies van ander spesialiste op die gebied kry. Dit beteken dat 'n aantal bevoegde persone elke item moet beoordeel met betrekking tot die relevansie van die item tot die eienskap wat gemeet word. Vir hierdie doel moet die inhoudsuniversum so duidelik as moontlik omskryf word.*

Vir die eerste doel van hierdie ondersoek, naamlik die standaardisering van 'n Wiskundeprestasietoets, het Cronbach (1971, pp. 458 - 459) die volgende raad vir

die toetsopsteller: *The test constructor who takes content validity seriously will begin, not by jotting down items as they occur to him or by soberly preparing one item regarding each page of the textbook but by translating each of his educational objectives into task specifications. He usually will be most successful if he subdivides his gross objective rather finely. Within each subdivision he will need to go beyond listing topics that might be tested to state the forms in which behavior relevant to each topic appears .....*

Hy skryf verder:

*If the original test or battery is a composite covering various types of content or various objectives, it implicitly weighs those elements, either by the number of items allocated to each or by the way the score is calculated. Such a weighing cannot satisfy decision makers who hold values unlike those of the test developers. Consequently, an ideally suitable battery for evaluation purposes will include separate measures for all outcomes the users of the information consider important ..... Reporting separate scores allows for the application of various systems of values.*

In die standaardiseringsproses van die Wiskundeprestasietoets, is hierdie raad gevolg (vergelyk paragrawe 3.2.1 en 3.4.3.4).

#### 2.6.2 Voorspellingsgeldigheid

Waar inhoudsgeldigheid nie in 'n bepaalde syfer uitgedruk word nie, word voorspellingsgeldigheid uitgedruk in terme van 'n syfer tussen -1 en +1. Seibel (in Whitla, 1968, p. 287) beweer dat voorspellingsgeldigheid die verband, uitgedruk as 'n korrelasiekoëffisiënt, tussen die toetstellings en 'n maatstaf van daaropvolgende gedrag is.

Volgens Anastasi (1964, p. 138) dui voorspellingsgeldigheid die doeltreffendheid van 'n toets aan in die voorspelling van 'n toekomstige gebeurtenis.

Volgens Oosthuizen (1973, p. 8) word voorspellingsgeldigheid hoofsaaklik gebruik vir keuring- en klassifikasiedoeleindes. Voorspellingsgeldigheid het dus te make met die bepaling van die verband tussen veranderlikes met die oog op voorspelling vir die toekoms.

Die verband tussen voorspeller en kriterium is volgens Madge (1976, p. 8) van beslissende betekenis en kan dien as 'n operasionele definisie van voorspellingsgeldigheid. Hierdie verband word aangedui deur 'n korrelasiekoëffisiënt en Ghisseli (1964, p. 338) noem dit 'n geldigheidskoëffisiënt.

'n Noodsaaklike eienskap van die kriterium waarna Madge verwys, moet volgens Helmstadter (1970, p. 298) vry wees van 'n toevallige foutevariëansie. Relatiewe vryheid van foutevariëansie is egter noodsaaklik maar nie voldoende vir 'n goeie kriterium nie. Indien die kriterium egter 'n konstante meting is, soos wat dit in hierdie ondersoek die geval is, moet dit ook betroubaar wees.

Die bespreking van die geldigheidskoëffisiënt waarna Ghisseli hierbo verwys, word later in paragraaf 3.4.3 en hoofstuk 5 bespreek.

### 2.6.3 Gelyktydige geldigheid

By voorspellingsgeldigheid is aanvaar dat toetsgegevens eers ingewin is en in 'n latere opvolgstudie is die kriteriumgegevens ingesamel. Wanneer die kriteriumgegevens tesame met die toetsgegevens ingesamel word en daar word korrelasies bereken, word 'n geldigheidskoëffisiënt verkry en word gepraat van gelyktydige geldigheid.

Seibel (in Whitla, 1968, p. 287) omskryf gelyktydige geldigheid as *the relationship, expressed in terms of a correlation coefficient, between the results of the test in question and those yielded by another observation, already accepted as valid of the same behavior.*

Anastasi (1964, p. 141) omskryf gelyktydige geldigheid as die verband tussen toetstellings en indekse van die kriterium wat ongeveer op dieselfde tydstip verkry is.

Dit blyk dus dat gelyktydige geldigheid te make het met 'n gelyktydige toepassing van twee of meer instrumente op 'n groep persone. Volgens Oosthuizen (1973, p. 8) is dit belangrik dat een van die twee instrumente as geldig aanvaar of bewys moet wees.

Samevattend kan daarop gelet word dat die hoofverskil tussen voorspellingsgeldigheid en gelyktydige geldigheid die tydsverloop tussen toepassings is. Anastasi (1964, p. 141) sê in hierdie verband: *In a number of instances, concurrent validity is found merely as a substitute for predictive validity. It is frequently impracticable to extend validations procedures over the time required for predictive validity or to obtain a suitable preselection sample for testing purposes.*

Die algemene metode om gelyktydige geldigheid te ondersoek is deur die berekening van 'n korrelasie tussen toetsdata en kriteriumdata. Hoe hoër die positiewe korrelasiekoëffisiënt wat verkry word, hoe beter is die kriteriumgeldigheid van die toets. Die geldigheid van 'n toets word in hoofstuk 3 verder bespreek.

## 2.7 BETROUBAARHEID

### 2.7.1 Inleiding

Die kenmerke van 'n goeie toets is dat dit geldig en

betroubaar is en dat dit oor norms beskik. In hierdie geval verwys die betroubaarheid na die konsekwentheid waarmee die toets as meetinstrument meet. Daar is egter faktore wat 'n rol speel in die konsekwentheid van die resultate van die meting. Die vernaamste van hierdie faktore is die volgende:

- (a) Die invloed van die meetinstrument self: By die meting van die wiskundeprestasie van 'n individu sal twee meetinstrumente, wat albei gebaseer is op die inhoud van die leergang, nie noodwendig dieselfde punt oplewer nie.
- (b) Die invloed wat uitgaan van die individu: Sy gesondheid, sy emosionele toestand en sy motiveering sal sy punt in 'n toets beïnvloed.
- (c) Die invloed van die subjektiewe oordeel van die nasiener van die toets.

Dit is volgens Willmott en Nuttall (1975, p. 1) sommige van die faktore wat bydra tot onakkuraatheid van of gebrek aan betroubaarheid in geestemeting. Dit sal nooit moontlik wees om hulle almal uit te skakel nie, maar dit is noodsaaklik dat van hulle teenwoordigheid kennis gedra word en 'n poging aangewend word om hulle invloed en effek te skat. Dit sal maniere daarstel om die onakkuraatheid van meetmiddels tot 'n minimum te beperk. Tensy die meetmiddel wat gebruik word, redelik akkuraat is, dit wil sê tensy van 'n kandidaat verwag kan word om dieselfde telling te behaal as hy die toets op 'n ander tydstip aflê of die toets deur 'n tweede nasiener laat nasien, is die meetmiddel van min waarde aangesien dit geen betekenisvolle inligting van die individu gee nie. Die betroubaarheid van die tellings of punte in 'n toets word dus 'n belangrike onderwerp vir ondersoek en maak verslaggewing van die fout inherent

in die metingsprosedure, op dieselfde wyse as in die fisiese wetenskappe moontlik.

### 2.7.2 Omskrywing van Betroubaarheid

Betroubaarheid verwys in die algemeen na die mate waarin tellings wat in verskillende toepassings van dieselfde instrument (of 'n ekwivalente vorm daarvan) verkry sou kon word, veralgemeen kan word.

Hoe minder die twee metings van mekaar verskil, hoe betroubaarder is die toets, of konstanter die resultate wat daarmee verkry kan word.

Kerlinger (1964, p. 430) verstaan onder betroubaarheid die konstantheid en akkuraatheid van metings. Onder konstantheid word volgens Adams (1964, p. 68) verstaan dat as dieselfde toets twee keer op dieselfde groep persone toegepas word, die leerlinge wat hoë tellings in die eerste toepassing behaal het ook hoë tellings in die tweede toepassing behoort te behaal en omgekeerd. Akkuraatheid van metings kan aan die hand van 'n horlosie verduidelik word. Indien 'n horlosie drie minute in 'n uur wen, beteken dit dat die horlosie konstant is, maar nie akkuraat nie. Volgens Ahmand (1967, p. 311) moet die inligting verkry deur 'n toets, die korrekte inligting wees en moet by elke toepassing van die toets dieselfde inligting verkry word.

Willmott en Nuttall (1975, p.2) definieer betroubaarheid soos volg: *Reliability may be defined as the accuracy - or the consistency - with which a test measures whatever it is measuring.*

Hierdie definisie dui op nog 'n belangrike aspek in geestemeting, naamlik dat die meetinstrument meet wat dit veronderstel is om te meet en dit dui weer op die geldigheid van die meetinstrument.

Dit blyk dus dat daar 'n verband bestaan tussen betroubaardheid en geldigheid. Ghiselli (1964, p. 353) beweer dat dit duidelik is dat namate die betroubaardheid van of die voorspeller, of die kriterium verlaag word, die geldigheidskoëffisiënt ook verlaag word. Ook Helmstadter (1964, p. 84) dui die verband aan deurdat hy beweer dat ..... *the maximum possible validity is the square root of the reliability.*

Daar kan ten slotte opgemerk word dat daar twee benaderingswyses ten opsigte van die begrip betroubaardheid bestaan, naamlik die metingsbenadering wat die konstantheid van herhaalde metings benadruk en die statistiese benadering wat die verhouding tussen ware variansie en foutvariensie as uitgangspunt neem.

### 2.7.3 Bepaling van toetsbetroubaardheid

#### 2.7.3.1 Algemeen

Die betroubaardheid van meetinstrumente is volgens Verwey (1975, p. 8) aanvanklik op een van die volgende maniere bepaal:

- (a) Toepassing van twee ekwivalente vorms van 'n toets en die berekening van die korrelasië tussen die twee stelle tellings.
- (b) Halvering van 'n enkele toets in twee groepe items wat veronderstel is om ekwivalent te wees en berekening van die korrelasië tussen die twee stelle tellings.
- (c) Herhaalde toepassing van dieselfde toets en die berekening van die korrelasië tussen die twee stelle tellings.

Aangesien daar in hierdie ondersoek nie van hierdie

metodes gebruik gemaak gaan word nie, gaan dit nie hier bespreek word nie. Dit word egter volledig deur Thorndike (1963, pp. 560 - 620) bespreek.

Die metode wat in hierdie ondersoek gebruik word, is dié van itemhomogeniteit.

#### 2.7.3.2 Die bepaling van toetsbetroubaarheid met behulp van item- en toetsstatistieke

Indien die vorige metodes nie gevolg kan word nie, kan 'n skatting van betroubaarheid verkry word uit inligting verskaf deur die tellings wat deur individue in die items van die meetinstrument behaal word.

Volgens Swart (1980, p. 10) kan betroubaarheid dus bepaal word uit inligting wat deur 'n enkele toepassing van 'n meetinstrument verkry word. Een van die vorige metodes was om 'n toets te halveer en dan die korrelasie tussen die twee stelde gegewens te bereken. Volgens Guilford (1956, p. 380) het Kuder en Richardson nog verder gegaan en die toets in soveel dele verdeel as wat daar items in die toets is en dit dan beskou as soveel toetse met een item elk. Om die betroubaarheid te bereken het Kuder en Richardson (1937, pp. 151 - 160) formules ontwikkel waarvan een bekend staan as die KR-20 formule. Dit word gebruik om die betroubaarheid van 'n toets te bepaal.

Die verskillende formules wat in hierdie ondersoek gebruik is, word volledig in paragraaf 3.4.4 bespreek en gaan dus nie hier bespreek word nie.

#### 2.7.4 Interpretasie van Betroubaarheidskoëffisiënte

Dit spreek vanself dat hoe hoër die koëffisiënt is, hoe betroubaarder sal die toets wees. Dit is egter moeilik om 'n reël neer te lê waarvolgens bepaal kan word of die

verkreë koëffisiënt impliseer dat die toets betroubaar is al dan nie. Ahmand en ander (1967, p. 45) beweer dat daar nie 'n vasgestelde minimumwaarde vir die betroubaardheidskoëffisiënt van 'n toets neergelê is nie. Helmstadter (1964, p. 83) skryf in hierdie verband die volgende: *One final word about reliability concerns the question as to how reliable a test should be before it is useful.*

Hy gee byvoorbeeld die volgende koëffisiënte en dui aan dat dit aan die minimum vereiste voldoen:

- (a) Om die groepsbekwaamheidspeil te evalueer: 0,50.
- (b) Om die verskille in groepsbekwaamheidspeil in twee of meer uitvoerings te evalueer: 0,90.
- (c) Om die individuele bekwaamheidspeil te evalueer: 0,94.
- (d) Om die verskille in individuele bekwaamheidspeil in twee of meer uitvoerings te evalueer: 0,98.

Dit blyk dus dat 'n baie belangrike faktor by die interpretasie van 'n betroubaarheidskoëffisiënt die doel is waarvoor die toets gebruik moet word. Nunnally (1967, p. 226) vat dit mooi saam in die volgende voorbeelde: *If, for example, in a particular school system children with IQ's below 70 are placed in special classes, it makes a great deal of difference whether the child has an IQ of 69 or 70 on a particular test.* Hy skryf verder: *If a college is able to admit only one-third of the students who apply, whether or not a student is in the upper third may depend on only a few score points on an aptitude test. In such instances it is frightening to think that any measurement error is permitted. Even with a reliability of 0,90, the*

*standard error of measurement is almost one-third deviation of test scores.* Hy sê waar sulke besluite na aanleiding van 'n spesifieke toets gemaak moet word, die aanvaarbare betroubaarheidskoëffisiënt 0,95 behoort te wees.

Samevattend word Helmstadter (1964, p. 85) aangehaal:  
*In summary, then, a reliability coefficient must be interpreted in terms of the procedure used and the homogeneity of the analysis group and is to be evaluated in the light of the purpose for which the test is to be used, the content area measured, and the success with which the similar instruments have met in eliminating variable errors.*

## HOOFSTUK 3

### STANDAARDISERING VAN 'N WISKUNDEMEETMIDDEL

#### 3.1 AGTERGROND

##### 3.1.1 Inleiding

Die behoefte aan die daarstelling van objektiewe meetmiddels in Wiskunde vir Swart leerlinge in die primêre en sekondêre skole bestaan reeds lank in die Republiek van Suid-Afrika. Michau (1978, p. 21) skryf in die verband: *There is a need for research into the causes for the low achievement level in mathematics amongst Black high school pupils in the Republic.*

Prestasie in Wiskunde in die sekondêre skool kan verbeter word as 'n stewige grondslag daarvoor in die primêre skool gelê word. Om 'n stewige grondslag te kan lê, moet die onderwyser kennis dra van die prestasiepeil van sy leerlinge. Die jong en onervare onderwyser kan slegs met sekerheid sê wat sy leerlinge se prestasiepeil is as hy gebruik maak van gestandaardiseerde meetmiddels om te bepaal wat hulle prestasiepeil is, relatief tot dié van die normpopulasie waartoe hulle behoort. In die lig van bogenoemde het die Departement van Onderwys en Opleiding die RGN genader om gestandaardiseerde Wiskundeprestasie- en Diagnostiese toetse daar te stel wat deur onderwysers in Primêre Swart skole gebruik kan word om die standaard van Wiskunde te verbeter.

Die RGN het gevolglik onderneem om sulke toetse te standaardiseer. Hierdie hoofstuk handel oor die daarstelling van 'n Wiskundeprestasietoets vir standerd 5.

##### 3.1.2 Doel met die toets

Met hierdie toets word beoog om 'n objektiewe, betroubare en geldige meetmiddel daar te stel om 'n standerd 5-leerling se kennis, begrip van en insig in Wiskunde te bepaal.

Die toets het ook ten doel om die wetslae waarmee die leerlinge die leerstof vir standerd 4 geassimileer het, te bepaal en om te diagnoseer met watter breë aspekte van die sillabus die leerlinge probleme ondervind.

### 3.1.3 Rasionaal

Die rasionaal van die toets is gebaseer op die veronderstelling dat 'n standerd 5-leerling se vermoë om getalle te manipuleer en om wiskundige probleme op te los, 'n geldige kriterium is vir sy prestasiepeil in Wiskunde.

## 3.2 BESKRYWING VAN DIE TOETS

### 3.2.1 Algemeen

Die toets, wat vir itemontledings- en itemseleksiedoeleindes opgestel is, bestaan uit vier vorms met drie subtoetse elk. Elke subtoets bestaan uit 20 items met 'n totaal van 60 items per vorm.

Om die toets te kon opstel is die sillabusinhoud baie deeglik bestudeer sodat die toetsinhoud alle aspekte van die sillabus dek.

Die sillabusinhoud is ontleed en in geskikte onderafdelings ingedeel waaraan gewigte toegeken is. Laasgenoemde is bepaal deur die belangrikheid van die onderskeie afdelings en die tyd wat nodig is om die leerstof af te handel.

Die gewigte wat aan die onderafdelings van die sillabus toegeken is, word in tabelle 3.1 tot 3.3 aangetoon.

Hierdie tabelle is van toepassing op al vier vorms en sal nie by die onderstaande besprekings herhaal word nie.

TABEL 3.1  
SPESIFIKASITABEL VAN TOETS 1: PARATE KENNIS

Beskrywing	Aantal items					
	+	-	x	+	Gemeng	Totaal
Basiese eienskappe van getalle			1		1	2
Notasie en plekwaarde					2	2
Gewone breuke					3	3
Desimale breuke	1	1	1	1	1	5
Persentasies					2	2
Benoemde hoeveelhede		1			1	2
Gemiddeldes					1	1
Verhouding en eweredigheid					1	1
Meetkundige begrippe					2	2
<b>Aantal items</b>	<b>1</b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>14</b>	<b>20</b>

TABEL 3.2  
SPESIFIKASITABEL VAN TOETS 2: BEGRIP

Beskrywing	Aantal
Basiese eienskappe van getalle	2
Notasie en plekwaarde	2
Gewone breuke	2
Desimale breuke	2
Persentasies	2
Benoemde hoeveelhede	2
Gemiddeldes	1
Verhouding en eweredigheid	2
Meetkundige begrippe	2
Versamelings	1
Getallesinne	1
Grafieke	1
<b>Aantal items</b>	<b>20</b>

TABEL 3.3  
SPESIFIKASIE-TABEL VAN TOETS 3: WOORDSOMME

Beskrywing	Natuurlike getalle	Gewone breuke	Desimale breuke	Persentasie breuke	Totaal
Geld	1			1	2
Lengte	1				1
Massa		1	1		2
Inhoud	1				1
Tyd	1				1
Meetkundige begrippe	2				2
Gemiddeldes	1				1
Verhouding en eweredigheid	4	1			5
Ander benoemde hoeveelhede	1		1	3	5
<b>Totaal</b>	<b>12</b>	<b>2</b>	<b>2</b>	<b>4</b>	<b>20</b>

### 3.2.2 Toets 1: Parate Kennis

Daar is vier vorms, naamlik a, b, c en d van hierdie toets opgestel en elke vorm bevat 20 items wat oor parate kennis handel.

Parate kennis sluit daardie items in wat basies feitekenenis soos byvoorbeeld die rangorde van getalle, basiese kennis van geld, tyd en mate toets. Die leerling se parate kennis van die behandelde leerstof en sy vaardigheid in die toepassing van hierdie parate kennis om meganiese bewerkings te doen, word getoets.

Onder hierdie kategorie val daardie items wat die vaardigheid van die leerling ten opsigte van die volgende toets:

(a) Die vier hoofbewerkings met rasionale getalle.

Voorbeelde: (i)  $298 + 512 - 316 = \square$

(ii) Vereenvoudig:  $\frac{7,4 + 3,6}{1,1}$

- (b) Alle ander bewerkings wat volgens 'n vaste patroon (algoritme) of formule wat baie goed aan die leerlinge bekend is, gedoen word.

- Voorbeelde:
- (i) Bepaal die KGV van 4; 8; 12.
  - (ii) Bereken die gemiddelde van 12 ℓ, 15 ℓ en 20 ℓ.

In die samestelling van toets 1 is daar gepoog om in vorm a makliker items as in vorms b, c en d in te sluit. Dit is gedoen om makliker items te hê en sodoende itemseleksie te vergemaklik.

### 3.2.3 Toets 2: Begrip

Die vernaamste doelstelling in die nuwe kernleerplanne van die Departement van Onderwys en Opleiding word soos volg geformuleer: *Om by die leerlinge 'n insig in en 'n begrip van rekenkundige beginsels en prosedures te ontwikkel ... (1975, p. 20).* Om die rede is 'n afsonderlike toets vir begrip beplan, sodat die leerling se begrip van die leerstof en van die daarmee gepaardgaande meganiese bewerkings getoets kan word.

Die items van hierdie toets is dan ook daarop gemik om vas te stel in watter mate die onderrig van Wiskunde in Swart skole daarin slaag om begrip in Wiskunde by die leerlinge te ontwikkel en uit te bou.

#### Voorbeelde:

- (i) Die gemiddelde van drie getalle is 30. Wat is die som van die drie getalle?
- (ii) Hoeveel sentimeter is 1 meter langer as 0,34 meter?
- (iii) Hoeveel keer kan die waarde van die syfer 5 in die getal 2 754 afgetrek word van die waarde van die syfer 2?

Die laaste voorbeeld toets nie net die leerling se vaardigheid om die meganiese bewerking  $2\ 000 \div 50$  te verrig nie, maar ook sy begrip van die plekwaarde van 'n syfer in 'n getal met 10 as basis en sy begrip van die bewerking van deling: Byvoorbeeld, hoeveel keer kan die een hoeveelheid van die ander hoeveelheid afgetrek word?

In die samestelling van hierdie toets is daar gepoog om in vorm d meer maklike items as in vorms a, b en c in te sluit. Dit is gedoen om itemseleksie te vergemaklik.

#### 3.2.4 Toets 3: Woordsomme

Om woordsomme te kan doen moet die leerling in staat wees om te verbaliseer en moet hy *begrip* kan toon. Laasgenoemde impliseer dat hierdie toets met toets 2 (Begrip) oorevleuel. By die finale samestelling van die toets is toets 2 (Begrip) en toets 3 (Woordsomme), as een toets geneem, sodat elke vorm uit net 2 toetse bestaan.

Die items van die derde toets word in teëstelling met items van die vorige twee kategorieë, altyd verbaal gestel en is hoofsaaklik daarop gerig om vas te stel in watter mate aan die doelstellings, soos in die sillabusse van die Departement van Onderwys en Opleiding (1975, p. 20) vermeld, beantwoord is om 'n *logiese denkwys*e by die leerlinge te ontwikkel en om probleme wat betrekking het op hoeveelhede en getalle, op te los.

Hier word getoets of die leerling sy verworwe kennis kan toepas om eenvoudige vraagstukke op te los wanneer hy te doen het met konkrete of benoemde getalle in situasies uit die alledaagse ervaringslewe van die mens.

#### Voorbeelde:

- (i) Jan kry 'n eenrandstuk om drie brode wat 17 sent elk kos, te gaan koop. Hoeveel kleingeld moet Jan terugbring?

- (ii) Petrus ry met sy fiets teen 'n snelheid van 14 km per uur. Hoe lank sal dit hom neem om 49 km af te lê?

Omdat daar in die toets wat vir normbepaling toegepas is, nie 'n aparte toets vir woordsomme was nie en daar nie so baie items oor woordsomme nodig was nie, is daar in toets 3 (Woordsomme) van vorm c meer moeilike items as in vorms a, b en d ingesluit.

### 3.2.5 Oefenvoorbeelde

In al vier vorms is dieselfde oefenvoorbeelde gebruik. Soortgelyke voorbeelde met moontlike antwoorde is:

(i)  $18 + 5 - 12 = \square$

A	B	C	D	E
6	23	11	1	25

(ii) Trek af: R62  
R28  
 R□

A	B	C	D	E
44	34	46	32	90

(iii) Tel op: 519 mans  
 405 mans  
 78 mans  
101 mans  
 □ mans

A	B	C	D	E
1 093	1 003	1 202	1 103	1 113

### 3.2.6 Antwoordblad

Die leerlinge moes die vrae op die aparte antwoordblad

beantwoord. Elke vraag is voorsien van vyf moontlike antwoorde, genommer A, B, C, D en E. Op die antwoordblad is langs elke vraagnommer die ooreenstemmende letters in 'n ovaal gedruk. Die leerling merk sy antwoord deur een van die ovaal spasies by die ooreenstemmende nommer op die antwoordblad in te kleur.

Die antwoordblad is so beplan en ontwerp dat dit masjinaal met 'n optiese merkleser gelees kan word. Die optiese leser plaas die data op 'n skyf of band sodat verdere statistiese verwerkings met behulp van die rekenaar gedoen kan word.

Op die antwoordblad word voorsiening gemaak vir die invul van die volgende identifiserende gegewens:

- Naam
- Vak
- Toetsafnemer
- Datum
- Skool
- Leerlingnommer
- Vorm van toets
- Standaard
- Huistaal
- Geslag
- Ouderdom
- Addisionele gegewens

### 3.3 STANDAARDISERING

#### 3.3.1 Ondersoekopset by die eksperimentele toepassing

In tabel 3.4 word 'n kort samevatting gegee van die toepassingsdoel, die meetinstrumente, die toetse en die toetsgroep.

TABEL 3.4

## TOEPASSINGSDOEL, MEETINSTRUMENTE, TOETSE EN TOETSGROEP

Toepassingsdoel	Meetinstrumente	Toetse in elke vorm	Toetsgroep
Toepassing vir itemontleding en itemseleksie	Wiskundeprasietoets: Vorm a Vorm b Vorm c Vorm d	1 Parate Kennis 2 Begrip 3 Woordsomme	'n Groep st. 5-leerlinge uit Pretoria-omgewing: S* = 240; M = 240; T = 480

\* S = Seuns; M = Meisies; T = Totaal

### 3.3.2 Beskrywing van steekproef

Dit was prakties nie moontlik om by die eksperimentele toepassing 'n verteenwoordigende landswye steekproef te trek nie.

Die Departement van Onderwys en Opleiding het 10 skole uit die Pretoria-Oos-kring op 'n verteenwoordigende wyse geselekteer. 'n Naamlys van die skole verskyn in tabel 1 van bylae A.

Die navorser het uit elke skool 24 seuns en 24 meisies op 'n sistematiese wyse geselekteer. Die vanne van die seuns is alfabeties neergeskryf en getel. Die getal is deur 24 gedeel en die antwoord is tot die volgende heelgetal benader. Veronderstel die getal wat dan verkry is, is  $n$ . Om die eerste seun te selekteer is enige seun op die alfabetiese lys ewekansig gekies, daarna is elke  $n$ de seun geselekteer totdat die verlangde aantal geselekteer is, met ander woorde daar is van sistematiese steekproeftrekking gebruik gemaak. (As die nodige aantal seuns nog nie aan die einde van die lys geselekteer is nie, is daar weer van voor begin en aangehou tel totdat die vereiste getal bereik is).

Voorbeeld:

Gestel daar was 54 seuns op die alfabetiese lys, genommer van 1 tot 54. Die berekening sou dan wees:-

$$\frac{54}{24} = 2\frac{6}{24}$$

Die getal is benader tot die volgende heelgetal, naamlik 3. Enige seun, sê die 23ste, moes ewekansig gekies word as die eerste leerling. Daarna moes elke 3de seun geselekteer word. Die nommers van die seuns wat geselekteer is, sou wees:

23, 26, 29, 32, ... ensovoorts

Die 11de seun wat geselekteer is, sou nommer 53 op die lys wees. Die 12de seun wat geselekteer is, sou nommer 2 op die lys wees (nommer 54 was die laaste seun op die lys en toe is weer van vooraf begin tel). Dieselfde prosedure is by die meisies gevolg.

3.3.3 Toepassing van die toets

a. Toetsmateriaal

Die volgende toetsmateriaal is aan elke leerling gegee:

- (a) Een HB-potlood
- (b) Een sagte uitveër
- (c) Een antwoordblad
- (d) Een toetsboekie (Vorm a of b of c of d)
- (e) Uitwerkpapier

b. Inleidende woorde

Die toetsafnemer het die leerlinge op hulle gemak gestel met 'n algemene praatjie waarom hulle die toets aflê en gevra dat hulle hul bes moes doen.

c. Uitdeel van toetsmateriaal

Die toetsmateriaal, soos beskryf in 3.3.3 (a), is aan elke leerling uitgedeel met die opdrag dat hulle nie in die toetsboekie moes rondblaaï nie.

d. Invul van die leerlinge se persoonlike besonderhede

Alvorens die toetsafnemer die leerlinge hulle persoonlike besonderhede laat invul het, is hulle vertrouð gemaak met die antwoordblad en hoe die gegewens daarop gekodeer word.

Hierna het hy die leerlinge die gegewens, soos genoem onder 3.2.6, op die antwoordblad laat invul en kodeer.

e. Oefenvoorbeelde

Van die drie oefenvoorbeelde het die leerlinge die eerste oefenvoorbeeld saam met die toetsafnemer gedoen en is aan hulle gewys hoe om die korrekte antwoord te kodeer.

Oefenvoorbeelde 2 en 3 is sonder hulp deur die leerlinge self gedoen. Nadat hulle dit gedoen het, is die antwoorde gekontroleer. Die toetsafnemer is nie toegelaat om van die toetsaanwysings af te wyk deur byvoorbeeld onderrig te gee of inhoudelike aspekte te behandel nie.

f. Aanwysings vir toetse 1, 2 en 3

Die toetsafnemer moes vir die leerlinge sê dat die toetse op dieselfde manier as die oefenvoorbeelde beantwoord moes word en beklemtoon dat die leerlinge nie in die boekie mag rondblaaï nie.

g. Toetstye

Sodra 90 persent van die leerlinge 'n toets voltooi het, is die tydsduur aangeteken. Na elke toets is 'n pouse van 10 minute gegee.

### 3.3.4 Bespreking van itemontleding

Nadat die antwoordblaaie met die optiese leser gelees en op band gelaai is, is die gegewens met behulp van die rekenaar verwerk. Die berekeningsmetode wat gebruik is, word deur Anstey (1966, p. 167 - 177) beskryf. Aan elke korrekte antwoord is een punt toegeken. Die verwerkte itemstatistiek is deur die rekenaar uitgedruk en word in tabel 3.5 aangetoon.

Om verduideliking van die itemstatistiek te vergemaklik is elke kolom bo-aan genommer (kyk tabel 3.5).

#### Kolom 1

Die itemnommer.

#### Kolom 2

Die sterretjie dui die korrekte antwoord aan.

#### Kolom 3

NG is die items nie gedoen nie of oorgeslaan, dit wil sê geen responsie is by die betrokke item gegee nie maar wel op (n) daaropvolgende item(s).

1-5: Die vier afleiers en die korrekte antwoorde.

NB dui op die items wat nie bereik is nie, dit wil sê daar was geen responsie op verdere items deur die leerlinge nie.

VERK dui die leerlinge aan wat nie daarin geslaag het om 'n korrekte responsie te gee nie, dit wil sê die totaal van NG, die verkeerde responsies en NB.

#### Kolom 4

Die aantal toetslinge wat die verskillende afleiers gekies het, die item oorgeslaan of nie bereik het nie, word hier aangedui.

TABEL 3.5  
ITEMSTATISTIEK VAN VORM A VIR ITEMS 4 EN 5

1	2	3	4	5	6	7	8	9	10	11	12
Item= nommer	Korrekte afleier	Nommer van afleier	Aantal	Persentasie van aantal gedoen	Persentasie van totale aantal	Gemiddeld vir toets	Verskil (Gemiddeld - Totale gemiddeld)	Kriterium= nommer	Item= kriterium= korrelasie	Item= standaard= afwyking	Produk
4	*.....	NG	5	4,2	4,0	4,4	-4,8	1	0,561	0,499	0,280
		1	65	55,1	52,4	11,3	2,1				
		2	15	12,7	12,1	6,2	-3,0				
		3	21	17,8	16,9	7,0	-2,2				
		4	5	4,2	4,0	7,6	-1,6				
		5	12	10,2	9,7	8,6	-0,6				
		NB	1	0,8	0,8	0,0	-9,2				
		VERK	59	50,0	47,6	6,8	-2,3				
5	*.....	NG	4	3,4	3,2	3,5	-5,7	1	0,491	0,500	0,246
		1	4	3,4	3,2	7,3	-1,9				
		2	8	6,7	6,5	6,0	-3,2				
		3	64	53,8	51,6	11,0	1,9				
		4	28	23,5	22,6	7,9	-1,3				
		5	15	12,6	12,1	7,8	-1,4				
		NB	1	0,8	0,8	0,0	-9,2				
		VERK	60	50,4	48,4	7,1	-2,0				

KRITERIUM

1

2

AANTAL

124

124

GEMIDDELDE

9,2

23,7

STANDAARDAFWYKING

3,947

9,361

K-R 14-betroubaarheid = 0,770

Metingsfout = 1,893

K-R 8-betroubaarheid = 0,837

#### Kolom 5

Die aantal leerlinge wat die verskillende afleiers gekies het, word uitgedruk as 'n persentasie van die aantal wat die betrokke item gedoen het.

#### Kolom 6

Soos vir kolom 5 maar nou gebaseer op die totale aantal leerlinge wat die toets gedoen het.

#### Kolom 7

In hierdie kolom word die gemiddelde prestasie van die groep wat die verskillende afleiers gekies het, die item nie gedoen of nie bereik het nie, aangetoon.

#### Kolom 8

In hierdie kolom word aangedui hoeveel elkeen van die genoemde groepe bo of onder die moeilikheidswaarde van die item presteer het.

#### Kolom 9

In hierdie kolom word die kodes vir die kriteria wat gebruik word vir berekening van die diskriminasiewaarde van die item aangedui. (Gewoonlik is dit die totaalstelling van die toets.)

#### Kolom 10

Die diskriminasiewaarde van elke item, bereken volgens die verskillende kriteria, word hier aangetoon.

#### Kolom 11

Hier word die itemstandaardafwyking getoon.

## Kolom 12

Die itembetroubaarheidsindeks, wat die produk van die item=standaardafwyking en die diskriminasiewaarde is, word hier gegee.

Aan die *einde van elke toets* is die volgende informasie deur die rekenaar uitgedruk:

- (i) Die aantal leerlinge wat die toets gedoen het
- (ii) Die gemiddelde van elke toets
- (iii) Die standaardafwyking van elke toets
- (iv) K-R 14-betroubaarheid
- (v) Metingsfout
- (vi) K-R 8-betroubaarheid
- (vii) Kriterium 1 en 2 is onderskeidelik toets 1 en toetse 1 + 2

In tabelle 1 tot 4 in bylae B word die volgende inligting vir toetse 1, 2 en 3 van vorms a, b, c en d gegee:

- (a) Die itemnommer
- (b) Die moeilikheidswaarde van elke item
- (c) Die diskriminasiewaarde van elke item
- (d) Die persentasie leerlinge wat die item nie bereik het nie
- (e) Die gemiddelde moeilikheidswaarde van elke toets
- (f) Die aantal leerlinge wat elke toets beantwoord het
- (g) Die betroubaarheidskoëffisiënt soos bereken met K-R 8
- (h) Die metingsfout

Volgens die tabelle in bylae B blyk dit uit die gemiddelde moeilikheidswaardes van 40,7%, 34,6%, 32,0% en 38,5% vir vorms a, b, c en d onderskeidelik, dat die toetse vir die leerlinge moeilik was.

Die feit dat vorms a en d die hoogste gemiddelde het, kan daaraan toegeskryf word dat elkeen van hierdie vorms een toets het (vorm a, toets 1 en vorm d, toets 2) wat meer

maklike items bevat as die ooreenstemmende toetse van die ander vorms.

Die feit dat vorm c die laagste gemiddelde het, kan daaraan toegeskryf word dat vorm c, toets 3, meer moeilike items as die ooreenstemmende toetse van die ander vorms het.

Aan die hand van bogenoemde inligting is items geselekteer vir die twee vorms A en B wat by normbepaling toegepas is. In afdeling 3.3.5 word die prosedure hoe items geselekteer is, bespreek.

### 3.3.5 Itemseleksie

By die seleksie van 'n item is hoofsaaklik na die diskriminasie- en moeilikheidswaardes gekyk.

Ebel (1965, p.364) gee die volgende indeling vir die diskriminasiewaardes:

*0,4 en hoër : baie goed*  
*0,30 - 0,39 : redelik*  
*0,20 - 0,29 : grensgevälle*  
*onder 0,19 : swak en moet verwerp word*

In hierdie ondersoek is items met diskriminasiewaardes bo-  
kant 0,20 geselekteer.

Nunnally (1964, p.164) meen dat items waarvan die moeilikheidswaardes naby die grense (20% en 80%) lê, nie ingesluit moet word nie, terwyl Alberts (1967, p.66) aanvaar dat waardes tussen 16,2% en 84,0% in baie van die toetse wel mag voorkom.

In hierdie ondersoek is items met moeilikheidswaardes van 20% tot 80% (op enkele uitsonderings na) geselekteer.

In tabelle 1 tot 4 van bylae B is die items wat buite die bogenoemde grense val, onderstreep.

Omdat die gemiddelde moeilikheidswaardes van die verskillende toetse gewissel het van 24,9% tot 47,5% en dus 'n aanduiding is dat die toets vir die leerlinge moeilik was, is vier makliker items in die toets wat vir normbepaling toegepas is, ingesluit.

Die moeilikheidswaardes van hierdie items (by itemseleksie) was onderskeidelik 80,8%, 83,5%, 83,9% en 84,3%, wat bokant die grens van 80% val wat vir hierdie ondersoek gestel is. Die onderste grens van 20% is egter gehandhaaf en diskriminasiewaardes van items vir die toetse het gewissel van 0,24 tot 0,43. Die items wat geselekteer is, se diskriminasiewaardes is met die uitsondering van vier items, bokant 0,30. Dié vier items se diskriminasiewaardes is onderskeidelik 0,235; 0,254; 0,276 en 0,283, wat nog bokant die minimum vereiste van 0,20 is wat deur Garrett en Woodworth (1964, p.364) gestel word.

Tabelle 5 en 6 in bylae B bevat die statistiese informasie van vorms A en B van die geselekteerde items wat vir normbepaling toegepas is. Die volgende inligting is daarin saamgevat:

- (a) Die moeilikheidswaardes van hoog na laag gerangskik
- (b) Die diskriminasiewaardes van die items
- (c) Die oorspronklike vorm met die itemnommer waarin die items voorgekom het, byvoorbeeld item 3, toets 1 van vorm c (T1c3)
- (d) Die itemnommer in die nuwe toets
- (e) Die persentasie leerlinge wat die item nie bereik het nie
- (f) Die gemiddelde moeilikheidswaarde van elke toets

#### Opsomming

Die geskatte statistiese gegewens van vorms A en B wat uit vorms a, b, c en d geselekteer is, word in tabel 3.7 saamgevat.

Aangesien die geselekteerde items deur vier verskillende groepe leerlinge beantwoord is (die vier groepe leerlinge het onderskeidelik vorms a, b, c en d beantwoord), moet die gegewens in tabel 3.7 slegs as geskatte waardes beskou word, wat dien as 'n voorspelling van die finale toets se statistieke.

Die korrelasies tussen die toetse van elke vorm word in tabel 3.6 gegee en omdat daar 'n hoë korrelasie tussen toets 2 en toets 3 van elke vorm is, kan daaruit afgelei word dat daar regverdiging bestaan vir die saamstel van die twee vorms waarvan die geskatte waardes in tabel 3.7 aangetoon word.

Statistieke soos betroubaarheid, standaardmetingsfout en standaardafwyking word later in paragraaf 4 bespreek. Hier kan gemeld word dat 'n toets met 'n gemiddelde moeilikheidswaarde van 49,9% en 'n betroubaarheidskoëffisiënt van 0,890 as baie bevredigend beskou kan word.

### 3.3.6 Samestelling van die toets vir die normbepaling

#### 3.3.6.1 Ondersoekopset by normbepaling

In tabel 3.8 word 'n kort samevatting gegee van die toepassingdoel, die meetinstrumente, die toetse en die toetsgroep wat betrekking het op die toepassing van die toets.

#### 3.3.6.2 Vorms A en B

##### a. Vorm A

By die toepassing vir itemseleksie het vorms a, b, c en d elk uit drie toetse, naamlik Parate Kennis, Begrip en Woordsomme bestaan. Omdat woordsomme ook in 'n baie groot mate insig en begrip toets (vergelyk 3.2.4) is beplan dat van die items van toets 3 van vorms a, b, c en d in toets 2 van vorms A en B opgeneem word. Volgens 'n gemiddelde korrelasie van 0,62 tussen toetse 2 en 3 van vorms a, b, c en d kan ook aanvaar word dat die twee toetse in hoë mate

TABEL 3.6  
INTERKORRELASIES VAN TOETSE

Vorm a (N = 124)

	Toets 1	Toets 2	Toets 3	Totaal
Toets 1	1			
Toets 2	0,629	1		
Toets 3	0,477	0,565	1	
Totaal	0,842	0,851	0,823	1

Vorm b (N = 119)

	Toets 1	Toets 2	Toets 3	Totaal
Toets 1	1			
Toets 2	0,635	1		
Toets 3	0,588	0,623	1	
Totaal	0,870	0,852	0,863	1

Vorm c (N = 118)

	Toets 1	Toets 2	Toets 3	Totaal
Toets 1	1			
Toets 2	0,670	1		
Toets 3	0,600	0,698	1	
Totaal	0,858	0,621	0,855	1

Vorm d (N = 119)

	Toets 1	Toets 2	Toets 3	Totaal
Toets 1	1			
Toets 2	0,588	1		
Toets 3	0,615	0,611	1	
Totaal	0,865	0,841	0,867	1

TABEL 3.7

GESKATTE STATISTIESE GEGEWENS VAN VORM A EN VORM B

	Vorm A	Vorm B
Getal items	40	40
Gemiddelde moeilikheidswaarde	49,9	49,9
Mediaanwaarde	0,432	0,428
Betroubaarheid (K-R 21)	0,89	0,89
Standaardmetingsfout	5,42	5,40
Standaardafwyking	16,7	16,6

TABEL 3.8

TOEPASSINGSDOEL, MEETINSTRUMENTE, TOETSE EN TOETSGROEP

Toepassingsdoel	Meetinstrumente	Toetse	Toetsgroep
Normbepaling	Wiskundeprasietoets Vorm A Vorm B	Parate kennis; Begrip Parate kennis; Begrip	Sistematiese gestratifiseerde steekproef uit st. 5-leerlinge van 8 taalgroepe *Totaal = 1 000

\* Tabelle 3.11 en 3.12 dui onderskeidelik die aantal leerlinge aan wat uit elke taalgroep en streek geselekteer is.

dieselfde konstruk meet.

Toetse 1, 2 en 3 van vorms a, b, c en d wat toegepas is vir itemseleksie toets dus dieselfde inhoudelike aspekte as toetse 1 en 2 van vorms A en B wat vir normbepaling toegepas is.

Die spesifikasietabelle 3.9 en 3.10 vir vorms A en B gee 'n aanduiding van die gewigte wat aan die onderafdelings van die sillabus toegeken is.

b. Vorm B

Net soos by vorm A, is van die items van toets 3 van vorms a, b, c en d in toets 2 van vorm B bygevoeg.

By die samestelling van vorms A en B is daar gepoog om die items so te selekteer dat die ooreenstemmende items van die twee vorms dieselfde inhoudelike aspekte toets en dat die mediaanwaardes en gemiddelde moeilikheidswaardes vir vorms A en B min of meer dieselfde is (kyk tabel 3.7). Met die uitsondering van enkele klein verskille is deurgaans in die doel geslaag.

Omdat vorms A en B, wat vir normbepaling toegepas is slegs 'n Parate kennistoets en 'n toets vir Begrip ingesluit het, is daar nuwe spesifikasietabelle opgestel.

Die spesifikasietabelle 3.9 en 3.10 wat die gewigte aantoon wat aan die onderafdelings toegeken is, verskyn op die volgende bladsy.

3.3.6.3 Antwoordblad

Die antwoordblad is dieselfde as wat vir die eerste toepassing gebruik is, en dieselfde prosedure is by die nasien daarvan gevolg.

TABEL 3.9

VORMS A EN B

SPESIFIKASIETABEL VAN TOETS 1: PARATE KENNIS

Beskrywing	Getalle	Geld	Mate massa	Breuke	Gemeng	Totaal
Optel	2	1	1	1		5
Aftrek	3		1	1		5
Vermenig= vuldig	1		1	1	1	4
Deling	1	1		1		3
Verhouding					1	1
Benoemde hoeveelhede					1	1
Meetkundige Begrippe					1	1
Totaal	7	2	3	4	4	20

TABEL 3.10

VORMS A EN B

SPESIFIKASIETABEL VAN TOETS 2: BEGRIP

BESKRYWING	Aantal
Versamelings	1
Basiese eienskappe van getalle	2
Notasie en plekwaarde	2
Breuke	3
Benoemde hoeveelhede	1
Verhouding en eweredigheid	1
Getallesinne	1
Grafieke	2
Meetkundige begrippe	2
Probleme	5
Totaal	20

### 3.3.7 Beskrywing van normgroep

Die ondersoekgroep is geselekteer uit die totale standaard 5-populasie van die volgende streke:

Noord-Transvaal  
Hoëveld  
Oranje-Vaal  
Johannesburg  
Kaapland  
Oranje-Vrystaat  
Natal

Hierdie streke het die volgende taalgroepe verteenwoordig:

Noord-Sotho  
Suid-Sotho  
Swazi  
Tsonga  
Tswana  
Venda  
Xhosa  
Zoeloe

Inspekteurs van skole van die Afdeling Sielkundige Dienste en Inspekteurs van onderwys van die Departement van Onderwys en Opleiding het die toetse toegepas. Elke toetsafnemer het twee skole ewekansig uit die gebied wat hy verteenwoordig, gekies. In tabel 2 in bylae A verskyn die name van die 50 skole wat in hierdie steekproef betrek is.

By elke skool is daar twintig leerlinge (tien seuns en tien meisies) op 'n sistematiese wyse geselekteer. Dieselfde prosedure soos beskryf in paragraaf 3.3.2 is ook hier gevolg.

In die steekproef is gepoog om die verhouding van die taalgroepe in ooreenstemming met die universum te bring.

Tabel 3.11 gee 'n uiteensetting van hoe die verskillende

taalgroepe in die steekproef verteenwoordig is.

TABEL 3.11

STEEKPROEF VOLGENS TAALGROEPE

Taalgroep	Aantal skole	Aantal leerlinge	% Leerlinge	% Inskrywingstal
Xhosa	8	160	16	15,2
Suid-Sotho	8	160	16	9,2
Noord-Sotho	10	200	20	16,2
Zoeloe	10	200	20	30,8
Tswana	8	160	16	15,6
Tsonga	2	40	4	5,2
Venda	2	40	4	4,0
Swazi	2	40	4	3,7
Totaal	50	1 000	100	99,9

Uit die tabel is dit duidelik dat die persentasie leerlinge van elke bevolkingsgroep wat getoets is, redelik ooreenstem met die persentasie van hulle inskrywingstal soos weergegee deur die Jaarverslag van die Departement van Onderwys en Opleiding (1978, p.279). Die afwykings wat daar wel was (kyk Suid-Sotho- en Zoeloe-taalgroepe), is te wyte aan die feit dat daar een Suid-Sotho-toetsafnemer te veel of anders gestel, een Zoeloe-toetsafnemer te min was.

Die steekproef is met behulp van die Chi-kwadraat-toets vergelyk met die universum. 'n Nie-betekenisvolle Chi-kwadraat van 11,64 is gevind (die kritieke waarde vir 7 grade van vryheid is 14,007 op die 5%-peil van betekenisvolheid), met ander woorde daar is nie rede om te glo dat die steekproef nie verteenwoordigend was nie.

Die verspreiding van leerlinge in die verskillende streke word in tabel 3.12 weergegee.

TABEL 3.12

## STEEKPROEF VOLGENS STREKE

Streek	Skole	Leerlinge
Noord-Transvaal	6	120
Oranje-Vaal	12	240
Hoëveld	2	40
Johannesburg	12	240
Kaapprovinsie	6	120
Oranje-Vrystaat	8	160
Natal	4	80
Totaal	50	1 000

Volgens die steekproef behoort daar ewe veel seuns en meisies te wees. By sekere skole was daar nie 'n totaal van twintig seuns of twintig meisies nie. Waar daar te min leerlinge van een geslag was, moes die toetsafnemer die getal met die teenoorgestelde geslag aanvul. As gevolg hiervan is daar nie presies ewe veel seuns en meisies in die steekproef betrek nie.

Tabel 3.13 toon die aantal seuns en meisies wat elke vorm beantwoord het.

TABEL 3.13

## WERKLIKE GETAL LEERLINGE GETOETS VOLGENS GESLAG

	Vorm A		Vorm B		Totaal	
	Seuns	Meisies	Seuns	Meisies	Seuns	Meisies
Getal	247	240	238	242	485	482

### 3.3.8 Toepassing vir normbepaling

#### 3.3.8.1 Toetsmateriaal

Die volgende toetsmateriaal is aan elke leerling verskaf:

- (a) Een HB-potlood
- (b) Een sagte uitveër
- (c) Een antwoordblad
- (d) Een toetsboekie (Vorm A of B)
- (e) Uitwerkpapier

#### 3.3.8.2 Inleidende woorde

Die toetsafnemer moes 'n inleidende praatjie lewer om die leerlinge op hulle gemak te stel en aan hulle te verduidelik hoekom hulle die toets skryf.

#### 3.3.8.3 Uitdeel van toetsmateriaal

Die toetsmateriaal soos beskryf in paragraaf 3.3.8.1 is uitgedeel met die opdrag aan die leerlinge dat hulle nie in die toetsboekies moes rondblaai nie.

#### 3.3.8.4 Invul van die leerlinge se persoonlike besonderhede

Die vertrouwdmaking met en invul van die antwoordblaaie is net soos vir die eksperimentele toepassing gedoen (kyk paragraaf 3.3.3(d)).

#### 3.3.8.5 Oefenvoorbeelde

Dieselfde oefenvoorbeelde as by die vorige toepassing is gebruik. Die eerste oefenvoorbeeld is saam met die toetsafnemer gedoen en oefenvoorbeelde 2 en 3 is deur die leerlinge self gedoen en deur die toetsafnemer gekontroleer.

#### 3.3.8.6 Aanwysings vir toetse

Die toetse moes op dieselfde manier as die oefenvoorbeelde beantwoord word. Die aanwysings vir die toepassing wat in

die handleiding uiteengesit is, moes noukeurig deur die toetsafnemer gevolg word.

### 3.3.8.7 Toetstye

By die eerste toepassing is die tye aangeteken wat 90% van die leerlinge nodig gehad het om die toetse te voltooi. Die tyd wat dit geduur het om die toetsmateriaal uit te deel en om die algemene aanwysings te gee en die oefenvoorbeelde te doen, is ook aangeteken.

Hierdie tye is as norm gebruik vir die berekening van die tye vir vorms A en B soos in tabel 3.14 uiteengesit.

TABEL 3.14

#### TYD VIR NORMBEPALINGTOEPASSING

	Tyd
Uitdeel van toetsmateriaal	±7 minute
Algemene aanwysings	±30 minute
Oefenvoorbeelde	±8 minute
Toets 1: Parate Kennis	50 minute
<u>Pouse</u>	10 minute
Toets 2: Begrip	42 minute
Totaal	±147 minute

### 3.3.9 Bespreking van resultate

#### 3.3.9.1 Inleiding

Die gegewens op die antwoordblaaie is, soos by die eerste toepassing, deur klerke gekontroleer en deur 'n optiese leser gelees en op band gelaai. Die itemstatistiek is daarna bereken.

Die volgende informasie oor die geselekteerde items van die finale toets word in tabelle 7 en 8 in bylae B aangetoon:

- (i) Moeilikhedswaarde van die items (M-waarde)
- (ii) Diskriminasiewaarde van die items ( $r_{it}$ )
- (iii) Persentasie leerlinge wat die item nie bereik het nie (% NB)
- (iv) Gemiddelde M-waarde van toetse
- (v) Gemiddelde M-waarde van toets 1 en toets 2

Die Frekwensieverdeling van die tellings soos behaal deur die leerlinge word in tabel 9 in bylae B aangetoon. Die tabel bevat die volgende gegewens:

- (a) Die frekwensieverdeling van die tellings soos deur die leerlinge behaal
- (b) Die kumulatiewe frekwensie
- (c) Persentasie frekwensie
- (d) Die kumulatiewe frekwensie persentasie
- (e) Die aantal leerlinge
- (f) Die gemiddelde
- (g) Standaardafwyking
- (h) Skeefheid
- (i) Kurtose

### 3.3.9.2 Bespreking

#### a. Moeilikhedswaardes

Soos genoem in paragraaf 3.3.5 is items met moeilikhedswaardes van 20% tot 80% geselekteer. Daar is egter vier items geselekteer waarvan die moeilikhedswaardes onderskeidelik 80,8%, 83,5%, 83,9% en 84,3% (vergelyk paragraaf 3.3.5) by die eerste toepassing was. Hierdie items se moeilikhedswaardes by die finale toepassing is nou onderskeidelik 64,6%, 70,1%, 53,2% en 61,6% wat binne die daargestelde grense val.

Dit is in ooreenstemming met Alberts (1967, p.66) se bevindinge.

ding dat items met moeilikheidswaardes tussen 16,2% en 84,0% in toetse wel mag voorkom.

Drie items wat moeilikheidswaardes van onderskeidelik 66,7%; 69,8% en 45,5% by die eerste toepassing gehad het, het by die tweede toepassing moeilikheidswaardes van respektiewelik 82,4%; 18,4% en 18,9% gehad.

Hierdie drie items se diskriminasiewaardes by die tweede toepassing was onderskeidelik 0,399; 0,294 en 0,271 en is dus meer as die minimum vereiste van 0,20. Omdat hulle dus goed diskrimineer en omdat hulle sekere inhoudelike aspekte van die sillabus dek, is hulle in die finale toets behou.

Die gemiddelde moeilikheidswaardes soos verkry uit die normbepalingtoepassing van vorms A en B, is onderskeidelik 47,0% en 43,1% (kyk tabelle 7 en 8 in bylae B).

Die gemiddelde moeilikheidswaardes vir vorms A en B is onderskeidelik 2,9% en 6,7% laer as die geskatte waardes van die twee vorms wat uit die vier vorms van die eerste toepassing geselekteer is.

b. Diskriminasiewaardes

Item 19 van toets 2 (vorm B) waarvan die diskriminasiewaarde 0,235 was (vergeelyk tabel 6 in bylae B) het met die finale toepassing baie beter gediskrimineer met 'n diskriminasiewaarde van 0,488 (vergeelyk tabel 8 in bylae B).

Waar die item aanvanklik nie as 'n goeie item beskou is nie, dui die itemstatistiek daarop dat dit nou 'n baie goeie item is.

Die omgekeerde is egter ook waar. Die itemstatistiek van die normbepaling toon dat daar een item (vorm A, toets 1, nommer 19) met 'n diskriminasiewaarde van 0,16 is. Die

item se diskriminasiewaarde was in die eerste toepassing 0,32 wat impliseer dat dit toe goed gediskrimineer het.

'n Moontlike rede waarom dit by die tweede toepassing nie goed gediskrimineer het nie, kan wees dat die term *oppervlakte* wat in die item gebruik word nie by die leerlinge bekend is nie. In standerd 4 het hulle met die term te doen gekry in hulle moedertaal. Die toets is egter aan die begin van standerd 5 toegepas in welke geval die medium van onderrig Engels of Afrikaans is.

'n Tweede moontlike rede is dat (by die eerste toepassing) die steekproef beperk was tot 'n klein area (Mamelodi-woongebied) terwyl die steekproef by die tweede toepassing al die streke betrek het. Omdat dit die enigste item met 'n diskriminasiewaarde van onder 0,20 was en omdat dit in die eerste toepassing goed gediskrimineer het, is dit om inhoudsgeldigheidsrede behou.

Die diskriminasiewaardes van vorms A en B soos verkry uit die normbepalingtoepassing het gewissel van 0,157 tot 0,614 (kyk tabelle 7 en 8 in bylae B).

'n Moontlike rede waarom die moeilikheids- en diskriminasiewaardes uit die normbepalingtoepassing laer as dié van die geselekteerde items uit die eerste toepassing is, kan wees dat die eerste steekproef meer homogeen (komende uit een woonbuurt) as die tweede steekproef (komende uit die RSA) was. 'n Tweede rede is dat by die eerste steekproef net stedelike leerlinge betrek is en by die tweede groep stedelike sowel as plattelandse leerlinge. Volgens die gemiddeldes in tabel 3.21 het die leerlinge in die stedelike gebied effens beter gedoen as dié op die platteland. Daar kon egter volgens die Z-toets geen betekenisvolle verskil tussen die gemiddeldes gevind word nie (vergelyk paragraaf 3.4.2.3).

c. Die persentasie leerlinge wat 'n item nie bereik het nie

In tabelle 7 en 8 in bylae B word die persentasie leerlinge

wat elke item nie bereik het nie, aangetoon. Onderskeidelik 26% en 23% het nie die laaste item van toets 1 en toets 2 van vorm A bereik nie. By vorm B is dit onderskeidelik 23% en 19%.

### 3.3.9.3 Bepaling van norms

#### a. Inleiding

Die telling wat 'n leerling in 'n toets behaal, word 'n onverwerkte of routelling genoem. Hierdie routellings verskaf nie direk interpreteerbare inligting ten opsigte van 'n leerling se prestasiepeil nie. Wanneer routellings met die frekwensieverdeling van die tellings van die normpopulasie waartoe die leerling behoort, vergelyk word, kan betekenisvolle afleidings gemaak word betreffende 'n leerling se prestasiepeil in die onderskeie toetse of in al die toetse saam.

'n Waarde (of reeks waardes) wat die verwagte, normale of gemiddelde prestasie van 'n groep persone aandui en waarmee die prestasie van 'n individu vergelyk kan word, word 'n norm genoem. Norms word gewoonlik uitgedruk in terme van persentielrange of standaardtellings.

Vir hierdie toetse is norms bepaal uit die verkreë punte-tellings van die normgroep, dit wil sê die groep leerlinge in standerd 5 op wie die toetse vir normbepaling toegepas is.

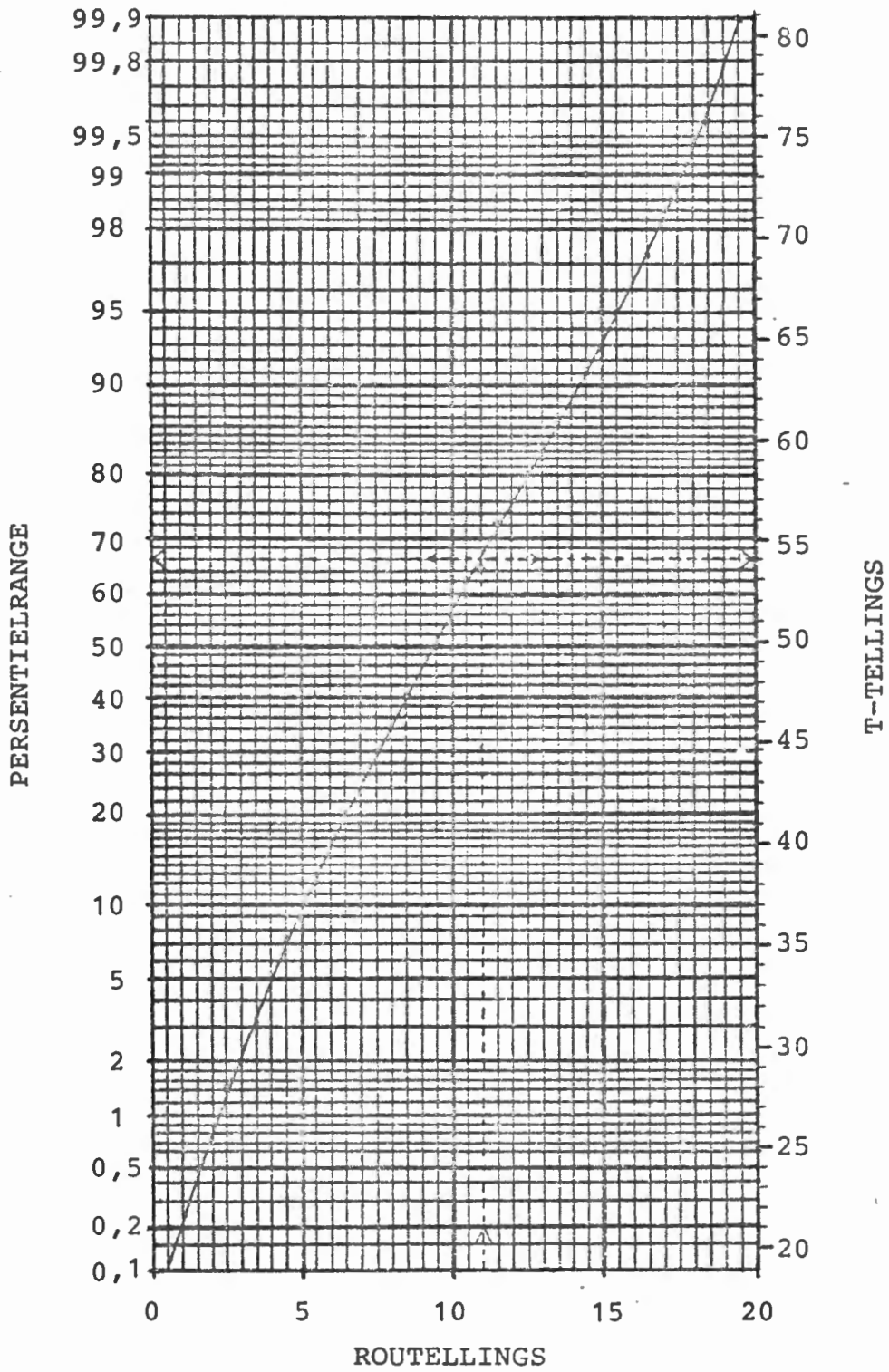
#### b. Normbepaling

Vir die bepaling van normpunte word normaalwaarskynlikheidspapier as hulpmiddel gebruik (kyk figuur 1).

Op die linkerkantse vertikale lyn verskyn 'n skaal wat kumulatiewe frekwensies aandui. Aan hierdie kant word die persentielrange afgelees.

Op die regterkantse vertikale lyn verskyn 'n indeling wat 'n standaardskaal aandui. Aan hierdie kant word die T-

FIGUUR 1  
 GRAFIEK OM ROUTELLINGS NA PERSENTIELRANGE EN T-TELLINGS  
 TE HERLEI



tellings afgelees (h verduideliking van T-tellings volg later).

Op die onderste horisontale lyn word die indeling vir routellings gemaak. Vir elke toets word h afsonderlike bladsy geneem en die routellings word op die horisontale lyn neergeskryf.

Die kumulatiewe frekwensie persentasie verkry uit die rekenaarprogram is op die waarskynlikheidspapier teenoor die routellings uitgestippel. Die waardes van die routellings met hulle ooreenstemmende persentasies kumulatiewe frekwensies word in tabel 9 in bylae B aangetoon. As die verdeling van tellings normaal is, sal die punte volgens Guilford (1968, p.498) min of meer in h reguit lyn lê. Die beste lyn word deur hierdie punte getrek sodat die kurwe kontinu en glad is. Ewe veel punte behoort aan weerskante van die lyn te lê nadat dit klaar getrek is. Enige routelling kan nou met behulp van die grafiek omgesit word in h standaardpunt.

As van die routelling vertikaal tot by die grafiek en dan weer horisontaal na links beweeg word, word die persentielrang afgelees. Deur van die lyn horisontaal na regs te beweeg, word die T-telling afgelees. Uit die grafiek (figuur 1) kan gesien word dat h routelling van 11 uit 20 h persentielrang van 66 en h T-telling van 54 gee.

Die ooreenstemmende persentielrang en T-telling vir elke routelling word in tabelle 3.15 en 3.16 aangetoon.

h Derde normskaal, naamlik staneges, word ook gebruik by die interpretasie van routellings. Tabel 3.17 gee h interpretasie van die stanegeskaal en hoe die nege punte met die persentielrange vergelyk.

TABEL 3.15

VORM A

## OMSETTING VAN ROUTELLINGS NA STANDAARDTELLINGS

Routellings			Standaardtellings			Beskrywing
Toets 1	Toets 2	Totaal	T-telling	Persen= tielrang	Stanege	
19-20 18 17	19-20 18 17	36-40 34-35 33 32	77 74 71 69	100 99 98 97	9	Baie goed
16 15 14	16 15	31 30 29 28	68 66 65 63	96 94 93 91	8	
13 12	14 13	27 26 25 24	62 61 59 58	89 86 82 80	7	Goed
11 10	12 11	23 22 21 20	57 56 55 54	76 73 68 66	6	Bogemiddeld
9 8	10 9	19 18 17 16	53 52 51 50 49 48	60 57 53 50 46 42	5	Gemiddeld
7 6	8	15 14	47 46 45	40 35 30	4	Onder= gemiddeld
5 4	7 6	13 12 11	43 42 41 40 39 38	22 20 18 16 14 12	3	Swak
3	5 4	10 9 8 7	37 35 34 33	10 7 6 5	2	Baie swak
2 1 0	3 2 0-1	6 5 0-4	32 30 26 21	4 2 1 0	1	

TABEL 3.16

VORM B

## OMSETTING VAN ROUTELLINGS NA STANDAARDELLINGS

Routellings			Standaardtellings			Beskrywing
Toets 1	Toets 2	Totaal	T-telling	Persen= tielrang	Stanege	
20	19-20	36-40	78	100	9	Baie goed
19	18	33-35	75	99		
18	17	32	71	98		
17		31	69	97		
16	16	30	68	96	8	
	15	29	67	95		
15		28	65	93		
14	14	27	64	92		
		26	63	91		
13	13	25	62	87	7	Goed
		24	61	85		
		23	60	83		
12	12	22	59	82		
		22	58	80		
11	11	21	57	76	6	Bogemiddeld
		20	56	73		
10	10	19	55	70		
9		18	54	68		
		18	53	62		
8	9	17	52	59	5	Gemiddeld
	8	16	51	55		
7		15	50	48		
		15	49	45		
6	7	14	48	40	4	Onder= gemiddeld
		13	47	36		
5	6	12	46	34		
		12	44	26		
4	5	11	42	22	3	Swak
		10	40	17		
		9	39	13		
3	4	8	37	10	2	
	3	7	36	9		
		7	34	5		
2		6	33	4	1	Baie swak
	2	5	32	3		
1	1	4	29	2		
0	0	0-3	26	1		
		0-3	21	0		

TABEL 3.17

## INTERPRETASIE VAN STANEGES

Stanege	Beskrywing	Persentielrang
1	Baie swak (onderste 11%)	0-4
2		5-11
3	Swak (volgende 12%)	12-23
4	Ondergemiddeld (volgende 17%)	24-40
5	Gemiddeld (middelste 20%)	41-60
6	Bogemiddeld (volgende 17%)	61-77
7	Goed (volgende 12%)	78-89
8	Baie goed (boonste 11%)	90-96
9		97-100

#### 3.3.9.4 Interpretasie van die normpunte en die gebruik van normtabelle

Drie soorte norms is bereken, naamlik persentielrange, T-tellings en staneges. Die normskale word kortliks hieronder bespreek en daar word verduidelik hoe elkeen gebruik kan word om die routellings wat die leerlinge in die toetse behaal het, te interpreteer.

##### a. Persentielrange

'n Persentielrang dui 'n leerlinge se relatiewe posisie in 'n bepaalde normpopulasie aan in terme van *die persentasie leerlinge wat 'n laer telling as hy/sy behaal het.*

Die persentielrang van 'n bepaalde toetstelling is dus die persentasie leerlinge in die normpopulasie wat 'n laer toetstelling as daardie bepaalde toetstelling sal behaal.

### Voorbeeld

As 'n standerd 5-leerling 'n puntetelling van 10 uit 20 in toets 2 van vorm A behaal, is die ooreenstemmende persentielrang 57 (kyk tabel 3.15). Dit beteken dat 57% van die leerlinge in die normpopulasie 'n laer routelling as 10 sal behaal. Dus is die betrokke leerling se telling beter as die tellings wat 57% van die normpopulasie sal behaal.

As dieselfde leerling ook 'n puntetelling van 10 uit 20 in toets 1 van vorm A sou behaal, is die ooreenstemmende persentielrang 68 (kyk tabel 3.15). Dit beteken dat 68% van die leerlinge in die normpopulasie 'n laer routelling as 10 uit 20 sal behaal. Dus is die betrokke leerling se telling in toets 1 beter as die tellings wat 68% van die normpopulasie sal behaal.

### b. Staneges

Die stanegeskaal is 'n genormaliseerde standaardskaal waarvolgens routellings in nege dele verdeel word. Die hoogste stanege is 9, die laagste is 1 en die gemiddelde van die verdeling is 5 met 'n standaardafwyking van 1,96.

Die prestasie van 'n individuele leerling in vergelyking met die normpopulasie word gewoonlik bepaal met behulp van persentielrange, terwyl die bepaling van die gemiddelde prestasiepeil van 'n besondere klas of groep leerlinge maklik kan geskied aan die hand van staneges.

Dit is van belang vir 'n onderwyser om te weet of sy leerlinge 'n goeie, gemiddelde of swak groep is. Gewoonlik kan hy die werklike standaard van sy groep leerlinge nie doeltreffend bepaal volgens sy eie subjektiewe oordeel

nie. Aan die hand van norms kan hy die standaard van 'n klas of groep betreklik akkuraat bepaal en in die vorm van 'n eensyfergetal met behulp van staneges uitdruk.

### Voorbeeld

'n Standerd 5-klas behaal 'n gemiddelde totaalstelling van 25 uit 40 vir vorm A. Hierdie telling val binne 'n staneg van 7 (kyk tabel 3.15). Die gemiddelde en bogemiddelde groep leerlinge in die normpopulasie se standaardtellings word verteenwoordig deur staneges 5 en 6. 'n Staneg van 7 val in die gedeelte bokant dié van die bogemiddelde groep en daar kan aanvaar word dat die betrokke klas se prestasie beter is as dié van die bogemiddelde groep in die normpopulasie.

Net so kan met behulp van 'n individuele leerling se punte bepaal word tot watter groep hy behoort.

### c. T-tellings

Routellings dui slegs die getal korrekte antwoorde in 'n toets aan. Verder kan routellings in twee verskillende toetse ook nie met mekaar vergelyk word nie, aangesien geen twee toetse identies is wat inhoud en moeilikheidsgraad betref nie.

Een van die groot voordele van T-tellings is dat hulle 'n betekenisvolle vergelyking tussen die tellings van verskillende leerlinge moontlik maak. Daarbenewens maak hulle ook 'n vergelyking van 'n leerling se tellings in verskillende toetse moontlik.

Die T-skaal het 'n gemiddelde van 50 en 'n standaardafwyking van 10 en die verdeling van die routellings van die normgroep is sodanig dat ongeveer 68% van die normgroep T-tellings tussen 40 en 60, 98% T-tellings tussen 30 en 70 en 99% T-tellings tussen 20 en 80 behaal.

Wanneer die T-tellings van twee leerlinge in dieselfde toets of die T-tellings van een leerling in twee verskillende toetse vergelyk word, is dit belangrik om in gedagte te hou dat die verkreë T-tellings moontlik deur toevallige faktore beïnvloed kon gewees het en nie noodwendig werklike verskille weerspieël nie.

Die *standaardfout van die verskil tussen twee tellings* is 'n statistiek wat gebruik kan word om 'n redelike aanduiding te gee van die kleinste verskille tussen twee tellings wat as sinvol beskou kan word, dit wil sê dit is moontlik om met 'n aansienlike mate van sekerheid te aanvaar dat, as die verskil tussen twee tellings groter is as die standaardfout van die verskil tussen twee tellings, die tellings werklik verskil en dat die verskil nie aan toevallige faktore toegeskryf kan word nie. Die standaardfout van die verskil tussen twee T-tellings word in tabel 3.18 aangetoon.

Die syfers in die blokkies word gebruik om die T-tellings van twee leerlinge in dieselfde toets met mekaar te vergelyk ten einde te bepaal of die tellings betekenisvol verskil of nie. Die ander syfers word gebruik om die T-tellings van dieselfde leerling in twee verskillende toetse met mekaar te vergelyk.

Die volgende formule word deur Anastasi (1971, p.98) aanbeveel om die standaardfout van die verskil tussen twee tellings te bereken

$$\sigma_{\text{verskil}} = \sqrt{\sigma_{\text{verskil}_1}^2 + \sigma_{\text{verskil}_2}^2}$$

waar  $\sigma_{\text{verskil}}$  = die standaardfout van die verskil tussen twee tellings

$\sigma_{\text{verskil}_1}$  = die standaardmetingsfout van die eerste toets

$\sigma_{\text{verskil}_2}$  = die standaardmetingsfout van die tweede toets

Hierdie formule word in terme van die standaardafwyking

TABEL 3.18

STANDAARDFOUT VAN DIE VERSKIL TUSSEN TWEE T-TELLINGS

VORM A

	Toets 1	Toets 2	Totaal
Toets 1	5,0	6,0	6,0
Toets 2	-	6,0	6,0
Totaal	-	-	6,0

VORM B

	Toets 1	Toets 2	Totaal
Toets 1	5,0	6,0	5,0
Toets 2	-	6,0	6,0
Totaal	-	-	6,0

en die betroubaarheidskoëffisiënt soos volg deur Nunnally omskryf

$$\sigma_{\text{verskil}} = SD\sqrt{2-r_{11}-r_{22}}$$

waar SD = standaardafwyking (vir T-tellings is dit 10 en vir staneges is dit 1,96)  
 $r_{11}$  = Betroubaarheidskoëffisiënt van toets 1  
 $r_{22}$  = Betroubaarheidskoëffisiënt van toets 2

### Voorbeeld

- (a) Jack se T-telling in toets 1 van vorm A is 57 en Jill s'n is 50 (in dieselfde toets). Die verskil tussen die twee tellings is  $57 - 50 = 7$ . Dit is meer as die standaardfout van die verskil tussen twee tellings, naamlik 5,0 (kyk tabel 3.18 van Vorm A). Daarom kan aanvaar word dat die verskil tussen die twee tellings op 'n werklike verskil dui en dat die verskil nie aan toevallige faktore toegeskryf kan word nie. Daar kan dus met 'n redelike mate van sekerheid aanvaar word dat Jack in werklikheid beter as Jill in die toets presteer het.
- (b) Jack se telling in toets 1 van vorm A is 57 en in toets 2 is dit 54. Die verskil tussen die tellings is  $57 - 54 = 3$ . Dit is minder as die standaardfout van die verskil tussen twee tellings, naamlik 6,0 (kyk kolom 2 en ry 1 van vorm A in tabel 3.18). Daar kan dus nie beweer word dat Jack wel verskil-lend in die twee toetse presteer het nie.

## 3.4 STATISTIESE EIENSKAPPE VAN DIE TOETS

### 3.4.1 Gemiddelde en Standaardafwyking

#### 3.4.1.1 Inleiding

Die meeste statistiese eienskappe van die toets, soos

dit voorgekom het by itemseleksie en normbepaling, is reeds genoem. 'n Uitbreiding daarop en 'n samevatting daarvan word hier gegee.

In tabel 3.19 word vir vorms A en B die volgende statistiese eienskappe, soos verkry uit die normbepalingtoepassing, gegee:

- (a) Steekproefgrootte
- (b) Aantal items
- (c) Gemiddelde van toetstellings
- (d) Standaardafwyking van toetse
- (e) Betroubaarheidskoëffisiënte
- (f) Standaardmetingsfoute
- (g) Skeefheid
- (h) Kurtose

TABEL 3.19

STATISTIESE EIENSKAPPE VAN DIE TOETS

Vorm	Toets	Steek= proef= grootte N	Punte= totaal	Gemid= delde	Stan= daard= afwy= king	Betrou= baar= heid K-R 8	Stan= daard= me= tings= fout	Skeef= heid	Kur= tose
A	1	495	20	8,5	3,9	0,83	1,60	0,438	2,72
A	2	495	20	9,7	3,5	0,76	1,71	0,121	2,68
A	1+2	495	40	18,2	6,7	0,80*	2,99	0,380	2,81
B	1	495	20	8,7	4,1	0,86	1,53	0,455	2,63
B	2	495	20	8,9	3,5	0,76	1,71	0,340	2,91
B	1+2	495	40	17,6	6,7	0,81*	2,92	0,493	2,96

\* K-R 21

3.4.1.2 Steekproefgrootte

Tabel 3.19 toon aan dat daar 495 leerlinge met behulp van vorm A en 495 leerlinge met behulp van vorm B getoets is.

Volgens die steekproef, soos beskryf in paragraaf 3.3.7 en tabel 3.12 moes daar 1 000 leerlinge vir die twee vorms gewees het. Tien van die antwoordblaaie (vyf per vorm) het ðf ontbrekende gegewens bevat, ðf is sodanig beskadig dat die inhoud daarvan nie op band gelaai kon word nie.

In tabel 3.20 word die aantal leerlinge wat vorms A en B beantwoord het onderskeidelik aangegee as 493 en 480. In tabel 3.21 is die aantal leerlinge onderskeidelik 487 en 480 in tabel 3.22 onderskeidelik 487 en 467. Die rede waarom hierdie getalle nie ooreenstem met die getalle in tabel 3.19 nie, is omdat dié leerlinge nie op die antwoordblaaie hulle geslag en/of die gebied waar hulle bly en/of die taalgroep waartoe hulle behoort, aangetoon het nie.

#### 3.4.1.3 Gemiddeldes

In tabel 3.19 word die rekenkundige gemiddeldes ( $\bar{X}$ ) van elke toets van elke vorm aangetoon. Hierdie gemiddeldes is bereken met behulp van die onderstaande formule:

$$\bar{X} = \frac{\Sigma X}{N}$$

waar  $\Sigma X$  = die routellings van al die leerlinge bymekaar-  
getel en

$N$  = die aantal leerlinge

Die gemiddeldes van die toetse wissel van 42,5% tot 48,5%. Die gemiddeldes vir vorm A en vorm B is 45,% en 44,0% re-  
spektiewelik. Dit is onderskeidelik 4,4% en 5,9% laer as die aanvanklike geskatte waardes soos aangetoon in tabel 3.7. Die verskil kan toegeskryf word aan die feit dat by die eerste seleksieproses die groep 'n meer homogene groep was omdat die leerlinge binne 'n gebied met 'n straal van 5 kilometer geselekteer is. By die tweede steekproef is die leerlinge geselekteer uit al vier die provinsies en is dus meer heterogeen.

Die gemiddeldes 45,5% en 44,0% dui aan dat die toetse eer-

der aan die moeilike as maklike kant is. Dit is egter verkieslik omdat die Departement van Onderwys en Opleiding die toetse ook vir opgraderingsdoeleindes wil gebruik. As die standaard opgegradeer is, sal die toetse waarskynlik nog gebruik kan word deur net ander norms te bereken. So word voorsien dat die toetse 'n langer leeftyd sal hê.

#### 3.4.1.4 Standaardafwyking

Die standaardafwykings van die verskillende toetse word in tabel 3.19 aangetoon. Die standaardafwyking is die vierkantswortel van die variansie. Die variansie word gedefinieer as die gemiddelde van die kwadrate van die afwykings van al die toetstellings vanaf die rekenkundige gemiddelde en word bereken met behulp van die volgende formule:

$$\sigma^2 = \frac{\sum x^2}{N-1}$$

waar  $\sigma^2$  = variansie

$$\sum x^2 = \sum_{i=1}^n (X_i - \bar{X})^2 = \text{die som van die kwadrate van die verskil van elke toets-telling met die rekenkundige gemiddelde}$$

N = die aantal leerlinge

Die standaardafwyking ( $\sigma$ ) is dus:

$$\sigma = \sqrt{\frac{\sum x^2}{N-1}}$$

In 'n normaalverdeling lê 99,74% van die gevalle tussen  $-3\sigma$  en  $+3\sigma$ , dit wil sê binne ses standaardafwykings, naamlik drie bokant en drie onderkant die rekenkundige gemiddelde van die verdeling. Die ideale standaardafwyking vir 'n toets wat uit 20 tel is  $\frac{20}{6} = 3,33$  en  $\frac{40}{6} = 6,67$  vir 'n toets wat uit 40 tel. Die standaardafwyking vir toetse 1 en 2

van vorms A en B wissel van 3,5 tot 4,1, dit wil sê 0,2 onder en 0,4 bokant die ideale. Vir vorms A en B as geheel is die standaardafwyking 6,7.

### 3.4.1.5 Skeefheid

Skeefheid word deur Glass en Stanley (1970, p.88) soos volg beskryf: *The degree to which the frequency distributions of a group of scores is asymmetrical is its skewness.*

'n Waarde vir skeefheid kan met behulp van die volgende formule bereken word:

$$\text{Skeefheid} = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^3}{S_x^3}$$

waar

$\sum (X_i - \bar{X})^3$  = som van die derdemagte van die verskil tussen elke toetstelling en die rekenkundige gemiddelde

$n$  = getal proefpersone

$S_x$  = standaardafwyking

Omdat  $z_x = \frac{X_i - \bar{X}_n}{S_x}$  kan die eerste formule geskryf word as

$$\text{Skeefheid} = \frac{1}{n} \sum_{i=1}^n z_x^3 = \bar{z}^3$$

waar

$z_x$  = 'n standaardtelling wat bekend staan as 'n z-telling

$\bar{z}^3$  = die gemiddelde van die z-tellings wat tot die derdemag verhef is.

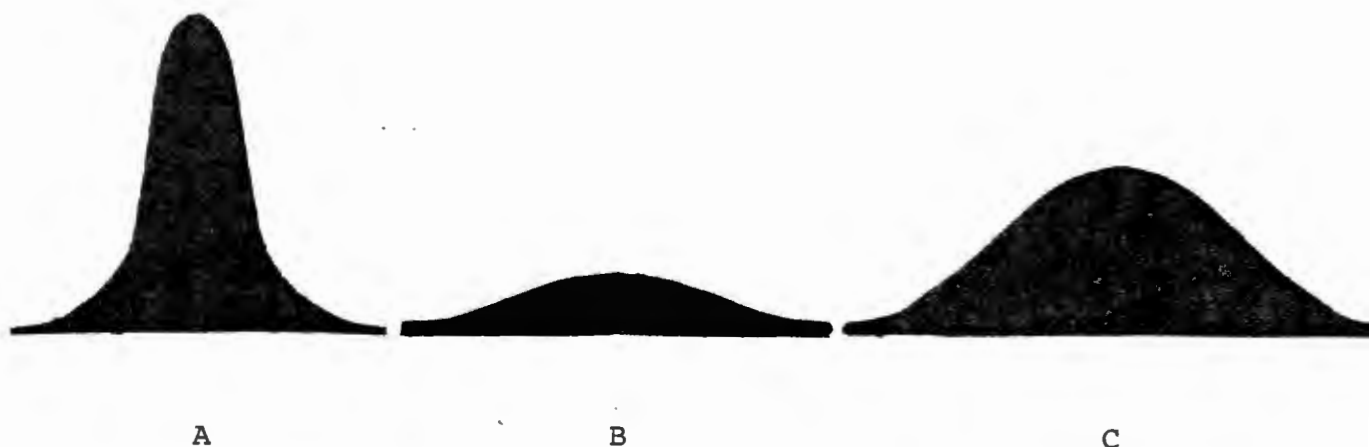
Die waardes verkry uit die formule sal gewoonlik in die gebied -3 tot +3 varieer. As die waarde nul is, is die verspreiding van tellings simmetries om die rekenkundige gemiddelde. As die waarde van  $\bar{z}^3$  negatief is, is die kurwe

negatief skeef, dit wil sê die verdeling is na links afgeplat. Die omgekeerde geld as die waarde positief is. Wanneer die waarde groot positief is, word die toets as moeilik beskou omdat daar meer tellings onderkant die gemiddelde is. As die omgekeerde die geval is, word die toets as maklik beskou.

Die skeefheidskoeffisiënte in tabel 3.19 lê tussen 0,1 en 0,5, wat klein positief is en beteken dat daar 'n redelike mate van normale verspreiding van punttellings in al die toetse aanwesig is.

#### 3.4.1.6 Kurtose

Kurtose word deur Glass en Stanley (1970, p.91) soos volg beskryf: *Kurtosis is a Greek word and refers to the quality of "peakedness" of a curve.*



Die mate van platheid of spitsheid in vergelyking met die normale verdeling, word in bostaande figure aangetoon.

Kurwe A is spitser as die normale kurwe (C) en word leptokurties genoem. Kurwe B is weer platter as kurwe C en word platikurties genoem. Die normale kurwe C word mesokurties genoem. Die formule om die kurtose te bepaal is:

$$\text{Kurtose} = \frac{1}{n} \frac{\sum (X_i - \bar{X})^4}{S_x^4}$$

waar alle simbole reeds beskryf is.

$$\text{Omdat } z_x^4 = \left[ \frac{(X_i - \bar{X}_n)}{S_x} \right]^4$$

kan die eerste formule geskryf word as

$$\text{Kurtose} = \frac{1}{n} \sum_{i=1}^n z_{x_i}^4 = \bar{z}^4$$

waar

$z_x$  = 'n standaardtelling wat bekend staan as 'n z-telling

$\bar{z}^4$  = die gemiddelde van die z-tellings wat tot die vierdemag verhef is.

As die waarde van  $\bar{z}^4 = 3$ , dan is die verdeling mesokurties. As dit groter as 3 is, is dit leptokurties. As dit groter as nul en kleiner as 3 is, is dit platikurties.

Die kurtose vir die verskillende toetse word in tabel 3.19 aangetoon en wissel van 2,63 tot 2,96, dit wil sê 3 benaderd. Daar kan dus gesê word dat die punte in 'n redelike mate normaal verdeel is. Die Kolmogorov-Smirnov-tegniek kan ook gebruik word om die normaliteit van die verdeling van die punte te ondersoek.

### 3.4.2 Verskille in die gemiddelde wiskundeprestasie van verskillende groepe

#### 3.4.2.1 Inleiding

*Wanneer in eksperimentele ondersoeke, soos ook in hierdie ondersoek die geval is, vergelykings tussen die prestasies van verskillende groepe getref moet word, dan is die standaardfout van die verskil tussen die gemiddeldes van meer praktiese belang. Die vraag kan ontstaan of die verskil tussen die gemiddelde prestasies van sekere groepe werklike verskille verteenwoordig en of dit maar aan toevallige faktore toegeskryf kan word (Heinichen, 1970, p.154).*

Vir die psigometris wat uitspraak moet gee oor verskille in prestasies van groepe is dit belangrik om met behulp van statistiese tegnieke vas te stel of verskille in prestasie werklike verskille verteenwoordig en of dié verskille aan

toevallige faktore toegeskryf kan word.

Die standaardfout van die verskil tussen die gemiddeldes vir onafhanklike groepe word volgens Guilford (1956, p.184) met die volgende formule bereken:

$$SF_{vg} = \sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}$$

waar

$SF_{vg}$  = die standaardfout van die verskil tussen gemiddeldes

$S_1^2$  en  $S_2^2$  = die variansie van die verskillende toetse

$N_1$  en  $N_2$  = die aantal leerlinge in elke toets.

Om betekenis te kan heg aan  $SF_{vg}$  moet dit omgesit word in 'n standaardvorm, naamlik 'n z-waarde. Dit word gedoen met behulp van die volgende formule:

$$z = \frac{\bar{X}_1 - \bar{X}_2}{SF_{vg}}$$

waar

$\bar{X}_1 - \bar{X}_2$  = die verskil tussen die gemiddeldes van die toetse.

As  $z \geq 1,96$  is die kans 5 uit 'n 100 dat die waargenome verskil aan toevallige faktore toegeskryf kan word.

As  $z \geq 2,58$  is die kans 1 uit 'n 100 dat die waargenome verskil aan toevallige faktore toegeskryf kan word.

Die verskille is dus betekenisvol op die 5%- of die 1%-peil.

As  $z < 1,96$  dan is die verskil nie betekenisvol nie.

3.4.2.2 Verskille in die gemiddelde wiskundeprestasie ten opsigte van geslag

In tabel 3.20 word vir die seuns en die meisies, die aantal leerlinge, die toetsgemiddelde en die standaardafwyking vir elke vorm aangetoon. Die standaardfout van die verskil tussen die gemiddeldes en die ooreenstemmende z-waardes van die groepe in vorm A en vorm B word ook aangetoon.

TABEL 3.20

BETEKENISVOLHEID VAN VERSKILLE TUSSEN DIE GEMIDDELDE PRESTASIE TEN OPSIGTE VAN GESLAG

Statistiese gegewens	Vorm A		Vorm B	
	Seuns	Meisies	Seuns	Meisies
Aantal leerlinge (N)	247	246	238	242
Toetsgemiddelde	18,25	18,8	17,89	17,31
Standaardafwyking	6,7	6,4	6,4	6,4
Standaardfout van die verskil tussen die gemiddeldes	0,59		0,62	
z-waarde	0,93		0,93	

Uit tabel 3.20 blyk dit dat  $z = 0,93 < 1,96$  vir vorm A is en dat  $z = 0,93 < 1,96$  vir vorm B is. Die z-waardes is opgesoek met behulp van Pearson en Hartley se tabelle (Du Toit, 1971, p.46).

Gevolgtrekking

Daar kan dus met 'n redelike mate van sekerheid aanvaar word dat daar nie 'n verskil tussen die gemiddelde wiskundeprestasies van seuns en meisies is nie.

### 3.4.2.3 Verskille in die gemiddelde wiskundeprestasie ten opsigte van gebied

In tabel 3.21 word vir stedelik en platteland die aantal leerlinge, die toetsgemiddelde en die standaardafwyking ten opsigte van elke vorm aangetoon. Die standaardfout van die verskil tussen die gemiddeldes en die ooreenstemmende z-waarde vir die groepe in vorms A en B word ook aangetoon.

TABEL 3.21

BETEKENISVOLHEID VAN VERSKILLE TUSSEN DIE GEMIDDELDE PRESTASIE TEN OPSIGTE VAN GEBIED

Statistiese gegewens	Vorm A		Vorm B	
	Stede=lik	Platte=land	Stede=lik	Platte=land
Aantal leerlinge (N)	339	148	338	142
Toetsgemiddelde	18,30	18,03	17,82	17,38
Standaardafwyking	6,72	6,17	6,94	6,31
Standaardfout van die verskil tussen die gemiddeldes	0,623		0,650	
z-waarde	0,432		0,677	

Uit die tabel is  $z = 0,432 < 1,96$  vir vorm A en is  $z = 0,677 < 1,96$  vir vorm B.

#### Gevolgtrekking

Daar kan dus met 'n redelike mate van sekerheid aanvaar word dat daar nie 'n betekenisvolle verskil tussen die gemiddelde wiskundeprestasies van stedelike en plattelandse leerlinge is nie.

3.4.2.4 Verskille in die gemiddelde wiskundeprestasie ten opsigte van taal

In tabel 3.22 word die aantal leerlinge, die toetsgemiddelde en die standaardafwyking wat op vorms A en B betrekking het ten opsigte van die volgende taalgroepe, gegee: Zoeloe, Xhosa, Venda, Tsonga, Tswana, Swazi, Suid-Sotho en Noord-Sotho.

TABEL 3.22

BETEKENISVOLHEID VAN VERSKILLE TUSSEN DIE GEMIDDELDE PRESTASIE TEN OPSIGTE VAN TAAL

Groep	Vorm A			Vorm B		
	Getal N	Toetsgemiddelde	Standaardafwyking	Getal N	Toetsgemiddelde	Standaardafwyking
Zoeloe	85	17,05	6,68	90	16,15	5,84
Xhosa	67	22,39	6,94	68	20,82	6,50
Venda	19	16,84	5,52	21	18,63	7,13
Tsonga	20	16,20	4,58	22	17,41	7,50
Tswana	74	19,78	6,61	73	17,89	6,81
Swazi	29	15,76	6,02	31	16,85	5,34
Suid-Sotho	97	19,14	6,02	92	17,01	6,70
Noord-Sotho	96	15,62	5,52	70	16,03	7,05

Met behulp van Scheffé se meervoudige vergelykingsmetode word bostaande gemiddeldes met mekaar vergelyk.

Hierdie metode word deur Hays (1965, p.484) om die volgende redes aanbeveel:

- (i) Dit kan maklik toegepas word.
- (ii) Dit word gebruik wanneer die aantal gevalle in die verskillende groepe nie gelyk is nie.
- (iii) Dit is nie so gevoelig vir foute wanneer die ver-

eistes in verband met homogeniteit van variansie en die normale verdeling van tellings nie nagekom word nie, tensy hierdie afwykings uitermate is.

- (iv) Dit kan gebruik word om enige vergelyking tussen groepe te tref.
- (v) Dit maak gebruik van die beskikbare F-tabelle.
- (vi) Dit is 'n meer konserwatiewe toets as enige ander meervoudige vergelykingsmetode ten opsigte van 'n Tipe I-fout, dit wil sê verwerping van die nulhipotese, terwyl dit in werklikheid aanvaar moes gewees het en lei dus tot 'n kleiner aantal betekenisvolle verskille.

Die formule om die F-waarde te bepaal is die volgende:

$$F = \frac{(\bar{X}_1 - \bar{X}_2)^2}{S_b^2 \left( \frac{n_1 + n_2}{n_1 n_2} \right) (k - 1)}$$

waar

$S_b^2$  = die binnegroepe-variensie en

$k$  = die aantal groepe

$\bar{X}_1, \bar{X}_2$  is reeds verklaar.

Die binnegroepe-variensie word eers bereken met behulp van die volgende formule:

$$S_b^2 = \frac{\sum x^2}{N-k} = \frac{\sum_{i=1}^k n_i s_i^2}{N-k} \text{ met } i = 1, 2, \dots, 8$$

waar

$x$  = die afwyking van 'n leerling se telling van die gemiddelde van die groep waarin hy val

$s_i$  = standaardafwyking van elke taalgroep.

Die binnegroepvariansie vir die agt taalgroepe word soos volg bereken:

$$\begin{aligned}
 S_b^2 &= \{85(6,68)^2 + 67(6,94)^2 + 19(5,52)^2 + 20(4,58)^2 + \\
 &\quad 74(6,61)^2 + 29(6,02)^2 + 97(6,02)^2 + 96(5,52)^2\} \div \\
 &\quad \{487 - 8\} \\
 &= \{3\,792,90 + 3\,226,96 + 578,94 + 419,53 + 3\,233,22 + \\
 &\quad 1\,050,97 + 3\,515,32 + 2\,925,16\} \div (479) \\
 &= 18\,743 \div 479 \\
 &= 39,13
 \end{aligned}$$

Die betekenisvolheid van die verskil tussen die gemiddelde prestasie van die Zoeloes (groep 1) en Xhosas (groep 2) word dan soos volg bereken:

$$\begin{aligned}
 F_{1,2} &= \frac{(17,05 - 22,39)^2}{39,13 \left( \frac{85 + 67}{85 \times 67} \right) (8 - 1)} \\
 &= \frac{28,516}{39,13(0,0267)(7)} \\
 &= \frac{28,516}{7,311} \\
 &= 3,90
 \end{aligned}$$

Volgens die F-tabel van Snedecor in Du Toit (1971, p.62) is:

$$0,95 F_{7,479} = 2,03$$

Die verkreeë waarde van F, naamlik 3,9 is groter as die kritiese waarde 2,03 en dus betekenisvol op die 5%-peil. Daar kan dus met 'n redelike mate van sekerheid gesê word dat die verskil tussen gemiddeldes 'n werklike verskil is en nie aan kansfaktore toegeskryf kan word nie, en dat die wiskundeprestasie van die Xhosa-taalgroep beter as dié van die Zoeloe-taalgroep is.

Tabel 3.23 dui die F-waardes tussen die verskillende taalgroepe aan.

TABEL 3.23

BETEKENISVOLHEID VAN DIE VERSKILLE TUSSEN DIE GEMIDDELDE PRESTASIES VAN DIE VERSKILLENDE TAALGROEPE

## VORM A

Taal-groepe	Zoeloe	Xhosa	Venda	Tsonga	Tswana	Swazi	Suid-Sotho	Noord-Sotho
				F-waardes				
Zoeloe	-							
Xhosa	3,90*	-						
Venda	0,00	1,66	-					
Tsonga	0,04	2,15*	0,01	-				
Tswana	1,08	0,88	0,48	0,74	-			
Swazi	0,13	3,25*	0,05	0,01	1,23	-		
Suid-Sotho	0,72	1,53	0,31	0,52	0,06	0,93	-	
Noord-Sotho	0,34	6,60*	0,09	0,02	2,64*	0,01	2,18*	-

## VORM B

Taal-groepe	Zoeloe	Xhosa	Venda	Tsonga	Tswana	Swazi	Suid-Sotho	Noord-Sotho
				F-waardes				
Zoeloe	-							
Xhosa	2,75*	-						
Venda	0,34	0,25	-					
Tsonga	0,09	0,63	0,05	-				
Tswana	0,12	0,98	0,03	0,01	-			
Swazi	0,04	1,09	0,13	0,01	0,08	-		
Suid-Sotho	0,11	1,85	0,15	0,01	0,10	0,00	-	
Noord-Sotho	0,00	2,58*	0,36	0,10	0,40	0,05	0,12	-

$$0,95 F_{7,479} = 2,03$$

$F \leq 2,03$  impliseer dat die verskille in gemiddeldes wat daar wel tussen dié groepe bestaan nie betekenisvol op die 5%-peil van betekenisvolheid is nie. Die een groep is dus nie beter of swakker as die ander groep nie.

\*  $F > 2,03$  impliseer dat die verskille in gemiddeldes wat daar voorkom betekenisvol op die 5%-peil van betekenisvolheid is. Die kans is vyf uit 'n honderd dat die verskille in gemiddeldes tussen die groepe aan kansfaktore toegeskryf kan word.

Volgens tabelle 3.22 en 3.23 het die Xhosas in Vorm A beter as die Zoeloes, Swazi's, Noord-Sotho's en Tsongas op die 5%-peil van betekenisvolheid presteer. Die Suid-Sotho's en Tswanas het beter as die Noord-Sotho's op die 5%-peil van betekenisvolheid presteer.

Volgens tabelle 3.22 en 3.23 het die Xhosas in vorm B beter as die Zoeloes en Noord-Sotho's op die 5%-peil van betekenisvolheid presteer. Uit die tabel blyk dat, behalwe vir die taalgroepe soos hierbo genoem, die ander taalgroepe nie betekenisvol van mekaar verskil het nie. Hoekom sekere taalgroepe oënskynlik beter as ander presteer het, kan nie hier verklaar word nie en indien 'n rede daarvoor gesoek word behoort, dit verder ondersoek te word.

### 3.4.3 Geldigheid

#### 3.4.3.1 Formule om geldigheid te bereken

Die formule wat gebruik is, is die Produk-momentkorrelasie (Guilford, 1965, p.95) wat op die volgende neerkom:

$$r_{xy} = \frac{\sum xy}{N\sigma_x\sigma_y}$$

waar

- $r_{xy}$  = korrelasie tussen  $x$  en  $y$
- $x$  = afwyking van 'n telling  $X$  vanaf die gemiddelde ( $\bar{X}$ ) van die toets
- $y$  = afwyking van die ooreenstemmende  $Y$ -telling vanaf die gemiddelde ( $\bar{Y}$ ) van die eksamen
- $\sum xy$  = som van die produk van die afwykingstellings
- $\sigma_x$  en  $\sigma_y$  = standaardafwykings van die verdelings van die  $X$ - en  $Y$ -tellings
- $N$  = aantal leerlinge getoets

Die interkorrelasies van die toetse van elke vorm is ook met hierdie formule bereken.

In tabel 3.24 word die interkorrelasies van die toetse (begin 1978 toegepas) met die eksamenpunte (einde van 1978) gegee. Die eksamenpunte van die einde van die jaar is die departementele eksterne standaard 5-eindeksamenpunte.

TABEL 3.24

INTERKORRELASIES VAN TOETSE (BEGIN 1978) EN EKSAMENPUNTE (EINDE 1978) VAN AL DIE LEERLINGE (N = 430)

Wiskunde= toets	Eksamenpunte			
	Vorm a N = 124	Vorm b N = 119	Vorm c N = 118	Vorm d N = 119
Toets 1	0,631	0,651	0,616	0,654
Toets 2	0,612	0,576	0,450	0,565
Toets 3	0,511	0,619	0,591	0,660
Totaal	0,698	0,713	0,721	0,731

In tabel 3.25 word die interkorrelasies van die Wiskunde-prestasietoets (begin 1979 toegepas) met die eksamenpunte (einde 1978 toegepas) gegee. Hierdie korrelasiekoëffisiënte gee 'n aanduiding van gelyktydige geldigheid.

Volgens Guilford (1956, p.146) kan die geldigheidskoëffisiënt, dit is die korrelasie tussen die toetstellings en die buitekriterium (eksamenpunt), enigiets tussen 0,00 en 0,60 wees, met die meeste van die korrelasies in die onderste helfte van hierdie gebied. Sommige toetsopstellers neem volgens Heinichen (1972, p.303) vir praktiese doeleindes 0,45 as die minimum waarde vir hierdie koëffisiënt.

#### 3.4.3.2 Voorspellingsgeldigheid

Om die voorspellingsgeldigheid te bepaal is die punte wat die leerlinge behaal het in die Wiskundeprestasietoets (voorspeller) wat vir eksperimentele doeleindes

TABEL 3.25

INTERKORRELASIES VAN TOETSE (BEGIN 1979) EN EKSAMENPUNTE (EINDE 1978) VAN AL DIE LEERLINGE (N = 324)

## VORM A

	Toets 1	Toets 2	Toetse 1+2	Eksamenpunte
Toets 1	1,000			
Toets 2	0,575	1,000		
Toetse 1+2	0,899	0,875	1,000	
Eksamenpunte	0,449	0,425	0,488	1,000

## VORM B

(N = 327)

	Toets 1	Toets 2	Toetse 1+2	Eksamenpunte
Toets 1	1,000			
Toets 2	0,548	1,000		
Toetse 1+2	0,899	0,858	1,000	
Eksamenpunte	0,435	0,418	0,451	1,000

(vorms a, b, c en d) aan die begin van 1978 toegepas is, gekorreleer met die wiskundepunte (kriterium) wat hulle in die departementele eindeksamen van 1978 behaal het (kyk tabel 3.24).

Korrelasiekoëffisiënte van 0,698; 0,713; 0,721 en 0,731 is verkry tussen die leerlinge se totale wiskundepunte in die aanvangsevaluasietoets en hulle eindeksamenpunt in Wiskunde (kyk tabel 3.24).

Dit impliseer dat onderskeidelik 48,7%, 50,8%, 52,0% en 53,4% van die variansie gedeel word deur die wiskunde=toetspunte en die eindeksamenpunte.

Die aanname kan gemaak word dat die toets 'n redelike mate van voorspellingsgeldigheid besit en dat dié korrelasies as bevredigend beskou kan word.

#### 3.4.3.3 Gelyktydige geldigheid

Die gelyktydige geldigheid is bepaal deur die wiskundepunte wat die leerlinge in die departementele eindeksamen van 1978 behaal het, te korreleer met die punte wat die leerlinge in die Wiskundeprestasietoets (vorms A en B), wat in die begin van 1979 vir normbepaling toegepas is, behaal het (kyk tabel 3.25). Daar het dus 'n kort periode (twee maande waarvan een en 'n halfmaand vakansie was) verloop tussen die verkryging van die onderskeie stelle punte.

Die korrelasie tussen die eksamenpunte en die totaal van toetse 1 en 2 van vorms A en B is onderskeidelik 0,49 en 0,45 (kyk tabel 3.25) wat volgens Guilford (1956, p.146) as bevredigend beskou kan word. Die korrelasiekoëffisiënte kon hoër gewees het as die standarde van skool tot skool en van onderwyser tot onderwyser nie so verskil het nie, want die 1978-standerd 4-eindeksamen was 'n interne eksamen waarvoor elke skool sy eie vraestel opgestel het en waarvan die eksamenskrifte deur die klasonderwyser self nagesien is.

#### 3.4.3.4 Inhoudsgeldigheid

*Inhoudsgeldigheid* word in die reël nie deur 'n bepaalde syfer uitgedruk nie, maar moet van die begin af op verskeie wyses deur die toetsopsteller in die toets "ingebou" word.

Om inhoudsgeldigheid te verseker, is 'n ontleding gemaak van die Wiskundesillabus vir standerd 4 van die Departement van Onderwys en Opleiding. Die items van die toetse is aan

vakkeners voorgelê om te verseker dat die sillabus deeglik gedek en die regte gewigte aan die verskillende afdelings daarvan toegeken is.

Die gewigte wat aan die onderafdelings toegeken is, is reeds bespreek (kyk tabelle 3.1, 3.2, 3.3, 3.9 en 3.10).

#### 3.4.4. Betroubaarheid

Die betroubaarheidskoëffisiënte van die toetse word in tabel 3.19 aangetoon. Die afsonderlike toetse se betroubaarheid is met behulp van die Kuder-Richardsonformule 8 en die betroubaarheid van vorms A en B met behulp van Kuder-Richardsonformule 21 bereken.

##### 3.4.4.1 Kuder-Richardsonformule 8

Al die inligting verlang vir die aanwending van hierdie formule word uit die itemontledingsresultate verkry. Hierdie indeks word volgens Nunnally (1970, p.127) sterk aanbeveel as 'n homogeniteitsindeks.

Die formule is die volgende:

$$r_{tt} = \frac{\sigma_t^2 - \Sigma pq}{2\sigma_t^2} \pm \sqrt{\frac{\Sigma r_{it}^2 pq}{\sigma_t^2} + \left(\frac{\sigma_t^2 - \Sigma pq}{2\sigma_t^2}\right)^2}$$

waar :

$r_{tt}$  = betroubaarheidskoëffisiënt

$\sigma_t^2$  = variansie van die toets

p = itemmoeilikeidswaarde

q = 1-p

$r_{it}$  = itemtotaaltellingkorrelasie

#### 3.4.4.2 Kuder-Richardsonformule 21

Hierdie formule vereis dat die matriks van die iteminterkorrelasies 'n rangorde van 1 het; dat korrelasies almal gelyk is en dat alle items dieselfde moeilikheidsgraad besit.

Die voordeel van hierdie formule is dat slegs die aantal items, die gemiddelde van die toets en die variansie benodig word.

Die formule wat deur Guilford (1954, p.381) gegee word, is die volgende:

$$r_{tt} = \frac{n\sigma_t^2 - m_t(n - m_t)}{(n - 1)\sigma_t^2}$$

waar

$m_t$  = gemiddelde toetstelling

$n$  = aantal items

Kuder en Richardson (1937, p.159) voer aan dat die betroubaarkheidskoëffisiënt verkry deur K-R 21 kleiner of gelyk is aan dié verkry deur die ander K-R-formules. Guilford (1954, p.382) sê in hierdie verband *...under usual testing conditions K-R formula 21 gives a lower bound estimate of  $r_{tt}$ .*

Die betroubaarkheidskoëffisiënte soos aangetoon in tabel 3.19 wissel van 0,76 tot 0,86 en is volgens Heinichen (1970, p.172) aanvaarbaar.

#### 3.4.5 Standaardmetingsfoute

'n Telling wat 'n leerling in 'n toets behaal, kan nie as eksak beskou word nie, want faktore soos geluk, raaiery, emosionele toestand en die manier waarop die toets toe-

gepas word, kan 'n rol speel (vergelyk Nunnally, 1964, p.71). Die telling is dus slegs 'n aanduiding van die ware telling.

Metingsfout word deur Ebel (1972, p.421) soos volg gedefinieer: *By the 'hypothetical true score' of an individual on a test is meant the average of a very large number of scores that might be obtained on similar tests, under similar conditions, for the same individual. The difference between the true score and an obtained score is called an 'error of measurement'.*

Die standaardafwyking van die verdeling van metingsfoute word die standaardmetingsfout genoem.

Die standaardmetingsfoute volgens tabel 3.19 wissel van 1,53 tot 1,71 vir toetse 1 en 2 van vorms A en B.

As die verkreë telling dus 12 uit 20 is, kan met 'n redelike mate van sekerheid aanvaar word dat die ware telling nie buite die omvang 10 tot 14 sal lê nie.

Vir die totaaltellings van vorms A en B is die standaardmetingsfoute onderskeidelik 2,99 en 2,92 wat impliseer dat as die verkreë telling 20 uit 40 is, met 'n redelike mate van sekerheid aanvaar kan word dat die ware telling in die orde van 17 tot 23 sal lê.

### 3.5 GEBRUIK VAN DIE TOETS

#### 3.5.1 Praktiese gebruik van die toets

Die toets kan veral gebruik word deur die onderwyser wat aan die begin van 'n jaar 'n groep standerd 5-leerlinge kry en nie weet wat die peil van hulle wiskundige bekwaamheid is nie. Nadat hy die toets op hulle toegepas het, kan hy hulle in groepe indeel en so meer aandag aan die swakker leerlinge bestee. Hy kan ook na aanleiding van toets 1 vasstel of die leerling se probleem by sy parate kennis lê en na aanleiding van toets 2 of dit 'n insigprobleem is.

n Verdere belangrike gebruik is dat dit as siftingstoets kan dien om swak presteerders te identifiseer en om leerlinge in homogene groepe in te deel. Dit sal meebring dat tyd en geld bespaar word by die toepassing van diagnostiese toetse, aangesien minder leerlinge dan daarby betrek hoef te word.

Deur gebruik te maak van die normtabelle kan die onderwyser vasstel of die leerling of groep leerlinge, ondergemiddeld, gemiddeld of bogemiddeld presteer.

### 3.5.2 Verdere navorsing

Soos wat die sillabusse hersien en verander word, sal die toetse ook hersien en aangepas moet word.

## HOOFSTUK 4

### 4 METODE BY DIE TWEDE GEDEELTE VAN HIERDIE ONDERSOEK

#### 4.1 INLEIDING

Volgens paragraaf 1.4.2 bestaan die ondersoek uit twee afdelings, naamlik

- (i) die saamstel van 'n Wiskundemeetmiddel waar daar nie verskillende hipoteses gestel is nie en
- (ii) die ondersoek van 'n spesifieke verband tussen wiskunde-prestasie, Verbale en Nie-verbale redeneringsvermoë by Swart leerlinge, waar verskeie hipoteses gestel is, wat in die tweede gedeelte van hierdie studie ondersoek is.

Die eerste gedeelte is in hoofstuk 3 beskryf. Die metode van ondersoek van die tweede gedeelte word in hierdie hoofstuk bespreek en resultate en bespreking van resultate word in hoofstuk 5 gerapporteer.

Saam met die Wiskundeprestasietoets wat in hoofstuk 3 bespreek is, is twee toetse, naamlik Patroonvoltooiing en Verbale Redenering, van die Skolastiese Aanlegtoets vir Swartes in standerd 6 (SATB) in hierdie ondersoek gebruik.

#### 4.2 ONDERSOEKSGROEP

Die ondersoeksgroep is om ekonomiese redes beperk tot standerd 5-leerlinge van 38 skole, in die Mamelodi-woonbuurt digby Pretoria.

#### 4.3 STEEKPROEF

Die prosedure vir die steekproeftrekkings was dieselfde soos beskryf in paragraaf 3.3.2. Daar is 10 skole op sistematiese wyse uit 'n moontlike 38 skole geselekteer. Twee beamptes van die RGN het by elkeen van die 10 skole die leerlinge, wat in die steekproef verteenwoordig is, op 'n sis=

tematiese wyse uit die totale aantal standerd 5-leerlinge wat op die toetsdag aanwesig was, geselekteer (kyk tabel 4.1).

TABEL 4.1  
TEORETIESE STEEKPROEF VOLGENS SKOOL EN GESLAG

Skool	Seuns	Meisies	Totaal
Eastwood HPS	24	24	48
Emthunzini HPS	24	24	48
Mogalo HPS	24	24	48
Monare HPS	24	24	48
Refentse HPS	24	24	48
Sikhanyisele HPS	24	24	48
Sindawonye HPS	24	24	48
Somisanang HPS	24	24	48
Tshwane HPS	24	24	48
Umthombo HPS	24	24	48
Totaal	240	240	480

#### 4.4 MEETINSTRUMENTE

##### 4.4.1 Wiskundeprestasietoets

Aangesien hierdie toets reeds volledig in hoofstuk 3 bespreek is, word hier slegs 'n opsomming van die vernaamste aspekte gegee.

##### a. Doel

Die doel met die toets is om 'n aanduiding te kry van 'n standerd 5-leerling se wiskundige vermoë.

##### b. Rasionaal

Die rasionaliteit van die toets is gebaseer op die veronderstelling dat 'n standerd 5-leerling se vermoë om getalle te manipuleer en om wiskundige probleme op te los, 'n geldige krite-

rium is vir sy wiskundige vermoë.

c. Beskrywing

Die toets bestaan uit vorms A en B. Elke vorm bevat twee toetse, naamlik

- (i) Toets 1: Parate Kennis
- (ii) Toets 2: Begrip

d. Toetstye

Toets 1: Parate Kennis: 50 minute  
Toets 2: Begrip: 42 minute

4.4.2 Skolastiese Aanlegtoets vir Swart leerlinge (SATB)

4.4.2.1 Inleiding

Die toets is gestandaardiseer vir Swart standerd 6-leerlinge en norms bestaan vir sewe Swart taalgroepe.

Die toetsbattery bestaan uit ses toetse met drie en veertig items elk. Die eerste drie items in elke toets is oefen= voorbeelde.

Die toetse van die SATB wat in hierdie ondersoek gebruik is, is:

- (i) Toets 5: Patroonvoltooiingstoets wat 'n nie-verbale redeneringstoets is en
- (ii) Toets 6: Verbale Redeneringstoets.

4.4.2.2 Patroonvoltooiing

a. Doel

Die doel met hierdie toets is om 'n aanduiding te kry van 'n leerling se nie-verbale redeneringsvermoë.

b. Rasionaal

Die toets berus op die aanname dat die vermoë om deduktief te redeneer, dit wil sê om 'n algemene beginsel uit gegewens af te lei en met behulp daarvan 'n reeks te voltooi, 'n aanduiding is van nie-verbale redeneringsvermoë.

c. Beskrywing

Die items van die toets bestaan uit onvoltooide patrone wat met behulp van een van vyf gegewe antwoordmoontlikhede voltooi moet word.

d. Toetstyd

Patroonvoltooiingstoets: 25 minute

4.4.2.3 Verbale Redenering

a. Doel

Die doel met die toets is om 'n aanduiding te kry van 'n leerling se verbale redeneringsvermoë.

b. Rasionaal

Die toets is gebaseer op die aanname dat die vermoë om aan die hand van verbaal gestelde analogieë, klassifikasies, verhoudings en berekeningsprobleme, basiese bewerkings te kan doen en tot logiese gevolgtrekkings te kan raak, 'n geldige aanduiding van verbale redeneringsvermoë is.

c. Beskrywing

In hierdie toets word daar van verbale stof gebruik gemaak vir die meting van die vermoë om verbale begrippe te vorm en dit op logiese wyse te manipuleer. Die inhoud van die toets is heterogeen en sluit take soos die voltooiing van analogieë en getallerye, klassifikasie, logiese gevolgtrek-

kings en rekenkundige probleme in. Volgens Van Staden (1972, p. 144) word hierdie tipe toets ingesluit in die meeste intelligensietoetse.

d. Toetstyd

Verbale Redenering: 35 minute

4.4.4 Betroubaarheid

Die betroubaarheid van die toetse van die SATB is nie in die handleiding genoem nie. Van Staden (1972, p. 204) het die SATB egter toegepas in 'n ondersoek waarin 4 559 Swart standaard 5-leerlinge betrek is. Hierdie betroubaarheid sowel as die van die Wiskundemeetmiddel word in tabel 4.2 aangetoon.

TABEL 4.2  
BETROUBAARHEID VAN DIE TOETSE

Meetmiddel	Betroubaarheid (K-R 8)
(i) Wiskunde	
Vorm A: Toets 1	0,83
Toets 2	0,76
Toets 1 + 2	0,80*
Vorm B: Toets 1	0,86
Toets 2	0,76
Toets 1 + 2	0,81*
(ii) SATB	
Patroonvoltooiing	0,91
Verbale Redenering	0,71

\* K-R 21

4.5 DIE TOEPASSINGSPROSEDURE

Die toets is deur die skrywer en twee skoolinspekteurs van die Departement van Onderwys en Opleiding afgeneem.

Die administratiewe prosedure vir die toepassing van die Wiskundetoets is reeds in hoofstuk 3 beskryf.

Die administratiewe prosedure vir die Patroonvoltooiing- en Verbale Redeneringstoets het geskied soos uiteengesit in die handleiding van die SATB (1971, pp. 2 - 16).

Die toets is aan die begin van die jaar toegepas. Daar is elke dag by twee skole getoets.

Die Wiskundetoets is eers deur die skrywer afgeneem, waarna die Patroonvoltooiing- en Verbale Redeneringstoetse deur die twee inspekteurs afgeneem is. Die toetsaanwysings vir Wiskunde was in Engels en die vir die ander twee toetse in Noord-Sotho en Zoeloe.

#### 4.6 DIE ANTWOORDBLAD

Dieselfde antwoordblad (PBS 1634) as wat by die standaardisering van die Wiskundetoets gebruik is, is hier gebruik.

Nadat die leerlinge die toetse beantwoord het, is die antwoordblaaie deur klerke van die RGN gekontroleer. Die antwoordblaaie is daarna masjinaal nagesien en die data is op band gelaai vir verdere ontledings.

Die resultate word in hoofstuk 5 gerapporteer.

## HOOFSTUK 5

### 5 RESULTATE, BESPREKING EN GEVOLGTREKKING TEN OPSIGTE VAN DIE TWEDE GEDEELTE VAN HIERDIE ONDERSOEK

#### 5.1 INLEIDING

Om die hipoteses vir die tweede deel van die ondersoek, soos gestel in hoofstuk 1, paragraaf 1.4.2 te ondersoek, is die volgende statistieke met behulp van die rekenaar bereken:

- (a) Korrelasies tussen die tellings wat in die Wiskunde-prestasietoets, die Verbale Redeneringstoets en die Nie-verbale Redeneringstoets behaal is (kyk hipotese (a) op bladsy 6).
- (b) Die gemiddelde prestasie wat die seuns en meisies in die Wiskundeprestasietoets behaal het (kyk hipotese (b) op bladsy 7).
- (c) Korrelasies tussen die tellings wat in die Wiskunde-prestasietoets en die totaal van die Verbale en Nie-verbale Redeneringstoets behaal is (kyk hipotese (c) op bladsy 7).

Die verband wat daar mag bestaan tussen bogenoemde statistieke word in hierdie hoofstuk bestudeer.

#### 5.2 ONDERSOEK

##### 5.2.1 Korrelasies tussen Wiskunde, Verbale Redenering en Nie-verbale Redenering

Korrelasies is met behulp van die rekenaar en met Pearson se produk-momentkorrelasieformule bereken (vergelyk paragraaf 4.3.2). Die korrelasies word in tabel 5.1 aangetoon.

TABEL 5.1  
 KORRELASIES TUSSEN WISKUNDE, VERBALE REDENERING  
 EN NIE-VERBALE REDENERING (UNIVERSUM)

VORM A (N = 240)

	Verbale Redenering	Nie-verbale Redenering
Wiskunde	0,58	0,36
Verbale Redenering		0,26

VORM B (N = 240)

	Verbale Redenering	Nie-verbale Redenering
Wiskunde	0,48	0,29
Verbale Redenering		0,40

Die korrelasie tussen Verbale Redenering en wiskundeprestasie is 0,58 en die tussen Nie-verbale Redenering en wiskundeprestasie is 0,36 (vergelyk tabel 5.1).

Op die oog af blyk Verbale Redenering 'n beter voorspeller van wiskundeprestasie te wees. Om dit te bevestig, gaan die volgende hipoteses wat ook in hoofstuk 1, paragraaf 1.4.2 gestel is, ondersoek word.

Die nul-hipotese wat getoets word, is dat daar nie 'n verskil in die voorspellingswaarde van Verbale en Nie-verbale Redeneringsvermoë ten opsigte van wiskundeprestasie is nie. Die alternatiewe hipotese is dat daar wel 'n verskil in die voorspellingswaarde is, dit wil sê:

$$H_0: \rho_{xy} = \rho_{xz}$$

$$H_1: \rho_{xy} \neq \rho_{xz}$$

waar

$\rho_{xy}$  = korrelasie tussen Wiskunde (X) en Verbale Redenering (Y)

$\rho_{xz}$  = korrelasie tussen Wiskunde (X) en Nie-Verbale Redene-  
ring (Z)

Die toets om  $H_0$  teenoor  $H_1$  te toets is volgens Glass & Stan-  
ley (1972, p. 313) die volgende:

$$z = \frac{\sqrt{n} (r_{xy} - r_{xz})}{\sqrt{(1-r_{xy}^2)^2 + (1-r_{xz}^2)^2 - 2r_{yz}^2 - (2r_{yz} - r_{xy}r_{xz})(1-r_{xy}^2 - r_{xz}^2 - r_{yz}^2)}}$$

waar

n = grootte van steekproef

$r_{xy}$  = die steekproefkorrelasie van X en Y

$r_{xz}$  = die steekproefkorrelasie van X en Z

$r_{yz}$  = die steekproefkorrelasie van Y en Z

Die waarde van z word nou soos volg bereken:

$$\begin{aligned} z &= \frac{\sqrt{240} (0,58-0,36)}{\sqrt{(1-0,58^2)^2 + (1-0,36^2)^2 - 2(0,26)^2 - (2(0,26) - (0,58)(0,36))(1-0,58^2-0,36^2-0,26^2)}} \\ &= \frac{15,49(0,22)}{\sqrt{(1-0,3364)^2 + (1-0,1296)^2 - 2(0,0176) - (0,52-0,2088)(1-0,3364-0,1296-0,0676)}} \\ &= \frac{3,4078}{\sqrt{0,4404 + 0,7576 - 0,0352 - (0,3112)(0,4664)}} \\ &= \frac{3,4078}{\sqrt{1,1628 - 0,14514}} \\ &= \frac{3,4078}{\sqrt{1,01766}} \\ &= \frac{3,4078}{1,0088} \\ &= 3,378 \end{aligned}$$

Omdat  $z = 3,378 > 1,96$  word die nul-hipotese op die 5%-peil  
van betekenisvolheid verwerp. Die alternatiewe hipotese, naam-  
lik  $\rho_{xy} \neq \rho_{xz}$  kan dus aanvaar word, wat beteken dat daar 'n  
verskil is in die voorspellingswaarde van Verbale Redenering  
en Nie-verbale Redenering ten opsigte van wiskundeprestasie.

Word daar na die korrelasiekoëffisiënte van Verbale Redenering teenoor wiskundeprestasie (0,58) aan die een kant en Nie-verbale Redenering teenoor wiskundeprestasie (0,36) aan die ander kant gekyk, kan die afleiding gemaak word dat Verbale Redenering 'n beter voorspeller van wiskundeprestasie is as Nie-verbale Redenering. Dieselfde afleiding kan ook by vorm B gemaak word. Die berekende z-waarde by vorm B is 2,98 wat ook betekenisvol op die 5%-peil is.

In aansluiting hierby het Gericke (1963, p. 49) gevind dat studente (reeds 'n geselekteerde groep) wie se Verbale IK meer as 10 punte hoër is as die Nie-verbale IK, die beste presteerders in die B.Sc.-rigting is. Hy skryf verder:

*Die belangrikste implikasie is dat dit die verbale faktor by uitnemendheid is wat verantwoordelik is vir 'n bepaalde prestasie (p. 50).*

Gunderson en Feldt (1960, p. 170) het met die California Test of Mental Maturity (CTMM) bevind dat leerlinge met die hoogste verbale tellings in die kriterium, dit is skoolprestasie soos gemeet met gestandaardiseerde prestasietoetse, die beste presteer en die groep met die hoogste nie-verbale tellings, die swakste.

Dit blyk dus dat die verband wat daar bestaan tussen Verbale Redenering en akademiese prestasie, in die geval van wiskundeprestasie, groter is as die tussen Nie-verbale Redenering en akademiese prestasie.

Dit kan ook afgelei word uit die 33,6 persent gemeenskaplike variansie wat daar bestaan tussen Verbale Redenering en wiskundeprestasie teenoor die 13 persent gemeenskaplike variansie wat daar tussen Nie-verbale Redenering en wiskundeprestasie bestaan.

#### 5.2.2 Gemiddelde wiskundeprestasie ten opsigte van geslag

Die hipoteses wat hier ondersoek gaan word, is die hipoteses

wat in hoofstuk 1, paragraaf 1.4.2 gestel is, naamlik die nul-hipotese dat daar geen verskil tussen die gemiddelde wiskundeprestasie tussen seuns en meisies op standerd 5-vlak is nie teenoor die alternatiewe hipotese dat daar wel 'n verskil is. Dit wil sê:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

waar

$\mu_1$  = gemiddelde wiskundeprestasie van die seuns

$\mu_2$  = gemiddelde wiskundeprestasie van die meisies

Die gemiddeldes van die Wiskundeprestasietoets soos behaal deur die seuns en meisies is met behulp van die rekenaar verkry. Die gemiddeldes word in tabel 5.2 aangetoon.

TABEL 5.2  
GEMIDDELDE WISKUNDEPRESTASIE VAN SEUNS (N = 242)  
EN MEISIES (N = 238)

Wiskundetoets	Seuns	Meisies
Gemiddelde	21,62	21,23
Standaardafwyking	9,04	8,27

Om betekenis te kan heg aan die verskil tussen die gemiddelde punte van die seuns en meisies moet die standaardfout van die verskil tussen die gemiddeldes eers soos volg bereken word (vergeelyk paragraaf 4.2.1):

$$SF_{vg} = \sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}$$

waar

$SF_{vg}$  = die standaardfout van die verskil tussen gemiddeldes

$S_1^2$  en  $S_2^2$  = die variansie van die verskillende toetse

$N_1$  en  $N_2$  = die aantal leerlinge in elke toets

Dis  $SF_{vg}$  is dus

$$\begin{aligned} SF_{vg} &= \sqrt{\frac{9,04^2}{242} + \frac{8,27^2}{238}} \\ &= \sqrt{0,3377 + 0,2871} \\ &= \sqrt{0,6248} \\ &= 0,794 \end{aligned}$$

Om betekenis te kan heg aan  $SF_{vg} = 0,79$  moet dit omgesit word in 'n standaardvorm, naamlik 'n z-waarde. Dit word gesoen met behulp van die volgende formule:

$$z = \frac{\bar{X}_1 - \bar{X}_2}{SF_{vg}}$$

waar

$\bar{X}_1 - \bar{X}_2$  = die verskil tussen die gemiddeldes van die toetse

In die besondere geval is

$$\begin{aligned} z &= \frac{21,62 - 21,23}{0,7904} \\ &= 0,493 \end{aligned}$$

Die verkreeë z-waarde van 0,493 val ver onder die kritieke waarde van 1,96 op die 5%-peil van betekenisvolheid. Dit impliseer dat die nul-hipotese nie verwerp kan word nie en dat aanvaar kan word dat seuns en meisies se wiskundeprestasies nie betekenisvol van mekaar verskil op standerd 5-vlak nie.

Dit is in ooreenstemming met die bevinding in paragraaf 3.4.2.2 waar die ondersoek ook uitgevoer is (dit was in die standaardisasieproses van die wiskundemeetmiddel).

### 5.2.3 Intelligensie as voorspeller van wiskundeprestasie

Die nul-hipotese wat in die eerste hoofstuk gestel is, naamlik dat intelligensie nie 'n voorspeller van wiskundeprestasie is nie, teenoor die alternatiewe hipotese dat dit wel 'n voorspeller van wiskundeprestasie is, word hier ondersoek, dit wil sê:

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

In tabel 5.3 word die korrelasies tussen die Wiskundeprestasietoets en die totaalstelling van die Verbale en Nie-verbale Redeneringstoets (wat as intelligensie geneem is) aangetoon.

Omdat die meeste intelligensietoetse uit 'n Verbale en 'n Nie-verbale afdeling bestaan, is besluit om die totaalstelling wat die leerlinge in die Verbale en Nie-verbale Redeneringstoets behaal het te neem as 'n aanduiding van hulle intelligensievermoë.

TABEL 5.3  
KORRELASIE TUSSEN WISKUNDEPRESTASIE EN INTELLIGENSIE

	Wiskundeprestasie Vorm A (N = 240)	Wiskundeprestasie Vorm B (N = 240)
Intelligensie	0,548	0,427

Om te bepaal of hierdie korrelasies beduidend is, is die nul-hipotese gestel, naamlik  $r = 0$  teenoor die alternatiewe hipotese, naamlik  $r = 0,548$  vir vorm A en  $r = 0,427$  vir vorm B.

Fisher het, volgens Snedecor (1950, p. 150), aangetoon dat

daar 'n transformasie van  $r$  moontlik is, wat 'n waarde gee wat normaal en simmetries versprei is. Dit is bekend as Fisher se  $z_r$ , en word verkry uit

$$z_r = \frac{1}{2} \left[ \log_e (1+r) - \log_e (1-r) \right]$$

vir  $r = 0,548$  is

$$\begin{aligned} z_r &= \frac{1}{2} \left[ \log_e (1+0,548) - \log_e (1-0,548) \right] \\ &= \frac{1}{2} \left[ \log_e 1,548 - \log_e 0,452 \right] \\ &= \frac{1}{2} (1,2310) \\ &= 0,6155 \end{aligned}$$

Die standaardafwyking hiervan is

$$s_z = \frac{1}{\sqrt{N-3}}$$

Vir  $z_r = 0,6155$  is

$$\begin{aligned} s_{z_r} &= \frac{1}{\sqrt{240-3}} \\ &= \frac{1}{15,395} \\ &= 0,06496 \end{aligned}$$

Die betekenisvolheid van  $z_r$  is

$$\begin{aligned} t &= \frac{0,6155-0}{0,06496} \\ &= 9,476 \end{aligned}$$

Hierdie waarde is baie meer as die kritiese waarde van 1,96 op die 5%-peil van betekenisvolheid. Die nul-hipotese van geen verband word dus verwerp en die alternatiewe hipotese, dat die korrelasie beduidend is, word aanvaar.

Die gevolgtrekking kan dus gemaak word dat intelligensie 'n voorspeller van wiskundeprestasie is.

Vir vorm B is bereken dat  $t = 7,02$ , wat ook betekenisvol op die 5%-peil van betekenisvolheid is.

Dieselfde gevolgtrekking, naamlik dat intelligensie 'n voorspeller van wiskundeprestasie is, kan ook hier gemaak word.

Alhoewel hierdie hoofstuk gaan om die bespreking van resultate moet onthou word dat die eerste stel resultate reeds in die vorige hoofstuk bespreek is.

## HOOFSTUK 6

### 6 SAMEVATTING, BETEKENIS VAN DIE ONDERSOEK EN AANBEVELINGS

#### 6.1 SAMEVATTING

Die eerste doelstelling van hierdie ondersoek was om 'n wiskundemeetmiddel daar te stel om die wiskundige vermoëns van Swart standerd 5-leerlinge te bepaal. In hoofstuk 3 word beskryf hoe die wiskundige meetmiddel gestandaardiseer is.

By die standaardisering van die wiskundemeetmiddel is daar vasgestel dat die Wiskundeprestasietoets 'n geldige en betroubare meetmiddel is en ook 'n goeie voorspeller van wiskunde-prestasie aan die einde van die jaar is. Die volgende bevindinge het ook uit die ondersoek na vore gekom:

- (a) Daar is bevind dat daar nie 'n beteknisvolle verskil in die wiskundeprestasies van seuns en meisies is nie.
- (b) Daar is bevind dat daar nie 'n betekenisvolle verskil tussen die gemiddelde wiskundeprestasies van stedelike en plattelandse leerlinge is nie.
- (c) Daar is egter bevind dat sekere Swart taalgroepe betekenisvol beter (op die 5%-peil) in Wiskunde presteer. Volgens die gemiddeldes was die Xhosas die beste presteerders in Wiskunde.

Die tweede doelstelling in hierdie ondersoek was om te kyk of Verbale Redenering, Nie-verbale Redenering en intelligensie voorspellers van wiskunde-prestasie is en watter een van die twee, Verbale Redenering of Nie-verbale Redenering, die beste voorspeller is. Daar is ook gekyk in watter mate geslag 'n rol speel in wiskunde-prestasie. Laasgenoemde is reeds by (a) hierbo gerapporteer.

Die bevinding is dat Verbale Redenering 'n beter voorspeller van wiskundeprestasie is as Nie-verbale Redenering, en dat intelligensie, soos gemeet deur die Verbale Redeneringstoets en Nie-verbale Redeneringstoets saam, 'n goeie voorspeller van wiskundeprestasie is.

Dit blyk uit die ondersoek dat Verbale Redenering, Nie-verbale Redenering en intelligensie al drie voorspellers van wiskundeprestasie is met Verbale Redenering die beste voorspeller, gevolg deur intelligensie en Nie-verbale Redenering.

## 6.2 BETEKENIS VAN DIE ONDERSOEK

Om die resultaat van hierdie ondersoek en die betekenis daarvan te beklemtoon, word Van Staden (1975, p. 431) soos volg aangehaal: *Daar bestaan geen intelligensie-, aanleg- of ander persoonlikheidstoetse wat vir Swart leerlinge op daardie vlak gestandaardiseer is nie. Ten tye van die ondersoek was daar ook geen gestandaardiseerde prestasietoetse vir die verskillende vakke op daardie vlak beskikbaar nie, .....*

Hierdie ondersoek het 'n gestandaardiseerde Wiskundeprestasietoets opgelewer wat met vrug op standerd vyf-vlak gebruik kan word.

## 6.3 AANBEVELINGS

Ten tye van hierdie ondersoek moes van die totaaltellings van Verbale Redenering en Nie-verbale Redenering van die verouderde Skolastiese Aanlegtoetsbattery vir leerlinge in standerd ses gebruik gemaak word om intelligensie te kon meet, omdat daar, soos reeds deur Van Staden gemeld, geen intelligensietoets vir Swartes bestaan nie. Die standerdisering van sodanige toets wat op alle vlakke van die Junior en Sekondêre skool gebruik sal kan word, word dus aanbeveel. Daar word ook aanbeveel dat die Wiskundepres=

tasietoets ook hersien sal word soos wat sillabusse verander om by veranderde omstandighede aan te pas.

In paragraaf 4.2.4 kon nie uitsluitel gegee word hoekom sekere Swart taalgroepe beter as ander in Wiskunde presteer het nie. Daar word dus aanbeveel dat die rede waarom sekere taalgroepe beter as ander presteer het, verder ondersoek behoort te word.

'n Laaste aanbeveling wat gemaak word, is dat die gestandaardiseerde Wiskundeprestasietoets vroeg in die skooljaar gebruik sal word om die swak leerlinge in Wiskunde te identifiseer en dat die gestandaardiseerde Diagnostiese Wiskundetoets weer op die leerlinge wat swak presteer het, toegepas moet word om sodoende hulle spesifieke probleemarea in die Wiskunde te identifiseer sodat remediëring gedoen kan word.

As dit gedoen kan word, sal die standaard van Wiskunde in Swart skole moet verbeter.

B Y L A E A

TABEL 1  
 NAME VAN SKOLE IN EKSPERIMENTELE TOEPASSING

Skoolnommer	Skool
01	Refentše HPS
02	Mogalo HPS
03	Monare HPS
04	Eastwood HPS
05	Šomišanang HPS
06	Tshwane HPS
07	Sindawonye HPS
08	Umthombo HPS
09	Emthunzini HPS
10	Sikhanyisele HPS

TABEL 2  
NAME VAN SKOLE IN NORMBEPALINGTOEPASSING

Skoolnommer	Skool
01	Lady Grey B.C.
02	Esethu
03	Little Flower
04	Addo
05	Johnson Marwanqa
06	Ilungelo
07	Warden Community
08	Vrede Community School
09	Zastron Community School
10	Smithfield
11	Boitshoko
12	Sol Plaatje
13	Thabong
14	Tswelopele
15	Ballengeich Mine School
16	Kilbarchan Mine School
17	Mpolweni HPS
18	Langsyde
19	Machadodorp
20	Tshwenyane
21	Dikubu
22	Lehlasedi
23	Patogeng
24	Walton Jameson
25	Madingoane
26	Tshepisa HPS
27	Seatile CPS
28	Maputle HPS
29	Matlape
30	Ditsobotla
31	Lesego CP School
32	Phaladi CP School
33	Setsing HPS
34	De Bruyn HPS
35	Driekoppies
36	Boschfontein
37	Lebowa HPS
38	Montshepetšabošego HPS
39	Pithuthalushaka HPS
40	Mambo HPS
41	Ditau
42	Itshepeng HPS
43	Usindiso HPS
44	Khulangolwazi HPS
45	Tiyani HPS
46	Khindlimukani HPS
47	Emthonjeni
48	Zola HPS
49	Shomang
50	Mokgolokwane

B Y L A E B

TABEL 1

VORM a

STATISTIESE GEGEWENS VAN DIE EKSPERIMENTELE TOEPASSING

	Toets 1			Toets 2			Toets 3					
	Nr.*	M-waarde	$r_{it}$	%NB	Nr.	M-waarde	$r_{it}$	%NB	Nr.	M-waarde	$r_{it}$	%NB
	4	83,5	0,497	0,8	3	28,0	0,410	0,0	3	36,7	0,418	0,0
	5	45,4	0,518	0,8	4	53,8	0,446	0,0	4	50,0	0,523	0,0
	6	45,5	0,512	0,8	5	<u>17,4</u>	<u>0,393</u>	0,0	5	32,8	0,486	0,0
	7	55,1	0,561	0,9	6	47,2	0,433	0,0	6	42,3	0,488	0,0
	8	53,8	0,491	0,8	7	58,2	0,314	0,0	7	<u>34,4</u>	<u>0,179</u>	0,0
	9	24,4	0,423	0,8	8	<u>22,7</u>	<u>0,174</u>	0,0	8	65,9	0,528	0,0
	10	50,4	0,289	0,8	9	28,9	0,483	0,0	9	<u>14,2</u>	<u>0,442</u>	0,0
	11	42,0	0,364	0,8	10	52,5	0,304	0,0	10	37,4	0,616	0,0
	12	32,0	0,260	0,8	11	<u>14,9</u>	<u>0,349</u>	0,0	11	43,5	0,438	0,0
	13	<u>12,5</u>	<u>0,174</u>	0,8	12	28,7	0,333	0,0	12	50,8	0,516	0,8
	14	80,8	0,377	1,7	13	<u>9,0</u>	<u>0,230</u>	0,0	13	67,5	0,433	0,8
	15	58,0	0,631	1,7	14	50,0	0,447	0,8	14	27,9	0,352	0,8
	16	30,8	0,394	1,7	15	45,5	0,441	0,8	15	26,0	0,440	0,8
	17	60,8	0,680	2,5	16	<u>13,2</u>	<u>0,288</u>	0,8	16	<u>9,9</u>	<u>0,189</u>	0,8
	18	46,2	0,629	2,6	17	47,5	0,397	1,7	17	75,6	0,421	0,8
	19	60,8	0,385	2,5	18	20,2	0,271	2,5	18	53,3	0,654	1,7
	20	45,4	0,493	2,5	19	49,6	0,519	2,6	19	<u>11,8</u>	<u>0,140</u>	1,7
	21	36,7	0,215	2,5	20	33,6	0,366	2,5	20	47,5	0,345	2,5
	22	<u>30,8</u>	<u>0,153</u>	2,5	21	<u>6,7</u>	<u>0,013</u>	4,2	21	26,9	0,486	3,4
	23	54,6	0,270	4,2	22	67,2	0,503	6,9	22	45,0	0,536	3,3
Gemid= delde	47,48				34,74				40,00			
Steek= proef= grootte	124				124				124			
K-R 8	0,837				0,761				0,853			
Stan= daard= metings= fout	1,893				1,828				1,792			

\*Nr. = nommer van item  
M-waarde = moeilikheidswaarde  
 $r_{it}$  = diskriminasiewaarde  
%NB = persentasie nie bereik nie

TABEL 2

VORM b

STATISTIESE GEGEWENS VAN DIE EKSPERIMENTELE TOEPASSING

	Toets 1			Toets 2			Toets 3				
	Nr. M-waarde	$r_{it}$	%NB	Nr. M-waarde	$r_{it}$	%NB	Nr. M-waarde	$r_{it}$	%NB		
4	68,9	0,359	0,0	3	33,0	0,398	0,0	3	50,9	0,202	0,0
5	47,1	0,550	0,0	4	59,7	0,372	0,0	4	61,3	0,278	0,0
6	37,1	0,509	0,0	5	45,8	0,468	0,0	5	43,0	0,606	0,0
7	40,2	0,450	0,0	6	<u>15,5</u>	<u>0,057</u>	0,0	6	60,2	0,456	0,0
8	53,4	0,305	0,0	7	23,3	0,560	0,0	7	40,2	0,265	0,0
9	27,0	0,304	0,0	8	<u>14,0</u>	<u>0,085</u>	0,0	8	<u>35,3</u>	<u>0,107</u>	0,0
10	28,2	0,309	0,0	9	44,4	0,289	0,0	9	39,3	0,493	0,0
11	23,7	0,295	0,0	10	<u>6,9</u>	<u>0,254</u>	0,0	10	<u>7,6</u>	<u>0,329</u>	0,0
12	44,9	0,345	0,8	11	44,8	0,524	0,0	11	47,0	0,508	0,0
13	24,1	0,461	0,9	12	<u>10,2</u>	<u>0,133</u>	0,0	12	50,4	0,415	1,7
14	65,3	0,442	0,9	13	42,2	0,421	0,9	13	22,1	0,415	1,8
15	47,5	0,555	0,8	14	<u>15,4</u>	<u>0,442</u>	1,7	14	47,0	0,468	1,7
16	<u>6,9</u>	<u>0,050</u>	1,7	15	22,6	0,255	1,7	15	<u>17,2</u>	<u>0,328</u>	1,7
17	59,0	0,523	1,7	16	48,7	0,259	1,7	16	52,1	0,308	1,7
18	28,4	0,233	1,7	17	<u>6,1</u>	<u>0,126</u>	2,6	17	31,6	0,510	1,8
19	36,2	0,362	1,7	18	27,0	0,216	2,6	18	<u>19,0</u>	<u>0,471</u>	1,7
20	41,9	9,423	1,7	19	47,4	0,437	2,6	19	<u>19,1</u>	<u>0,496</u>	1,7
21	36,0	0,428	2,6	20	64,3	0,334	2,6	20	31,6	0,467	3,5
22	<u>25,2</u>	<u>0,090</u>	3,5	21	<u>18,9</u>	<u>0,345</u>	4,5	21	<u>9,2</u>	<u>0,208</u>	9,2
23	27,0	0,404	6,3	22	15,0	0,320	5,3	22	20,4	0,293	10,2
Gemid= delde	38,3			30,26				35,22			
Steek= proef= grootte	119			119				119			
K-R 8	0,779			0,692				0,785			
Stan= daard= metings= fout	1,908			1,793				1,840			

TABEL 3

VORM c

STATISTIESE GEGEWENS VAN DIE EKSPERIMENTELE TOEPASSING

	Toets 1				Toets 2				Toets 3			
	Nr.	M-waarde	$r_{it}$	%NB	Nr.	M-waarde	$r_{it}$	%NB	Nr.	M-waarde	$r_{it}$	%NB
4	72,0	0,395	0,0	3	42,1	0,286	0,0	3	22,6	0,484	0,0	
5	51,7	0,609	0,0	4	53,6	0,409	0,0	4	21,4	0,475	0,0	
6	47,0	0,475	0,0	5	35,3	0,349	0,0	5	47,9	0,372	0,0	
7	<u>10,4</u>	<u>-0,073</u>	0,0	6	28,7	0,373	0,0	6	<u>12,8</u>	<u>0,122</u>	0,0	
8	69,2	0,422	0,0	7	<u>9,3</u>	<u>0,079</u>	0,0	7	<u>5,9</u>	<u>0,138</u>	0,0	
9	37,1	0,442	0,0	8	24,3	0,337	0,0	8	<u>15,5</u>	<u>0,389</u>	0,0	
10	35,3	0,456	0,0	9	26,7	0,306	0,0	9	62,9	0,315	0,0	
11	<u>18,1</u>	<u>0,261</u>	0,0	10	<u>17,2</u>	<u>0,265</u>	0,0	10	<u>17,9</u>	<u>0,489</u>	0,0	
12	34,8	0,323	0,9	11	<u>20,5</u>	<u>0,063</u>	0,0	11	69,8	0,427	0,0	
13	36,5	0,305	0,9	12	<u>10,4</u>	<u>0,008</u>	0,9	12	<u>5,1</u>	<u>-0,008</u>	0,0	
14	77,6	0,430	0,9	13	<u>18,4</u>	<u>0,388</u>	0,9	13	<u>19,0</u>	<u>0,489</u>	0,0	
15	51,7	0,614	0,9	14	<u>19,7</u>	<u>-0,008</u>	0,9	14	<u>17,1</u>	<u>0,111</u>	0,0	
16	29,3	0,514	0,9	15	69,0	0,418	0,9	15	<u>15,4</u>	<u>0,175</u>	0,0	
17	46,1	0,574	0,9	16	<u>17,9</u>	<u>0,110</u>	0,9	16	<u>9,4</u>	<u>0,327</u>	0,0	
18	<u>18,1</u>	<u>0,140</u>	0,9	17	<u>8,5</u>	<u>0,157</u>	0,9	17	<u>16,2</u>	<u>0,264</u>	0,0	
19	<u>10,3</u>	<u>-0,089</u>	0,9	18	84,3	0,386	0,9	18	34,5	0,580	0,9	
20	43,1	0,289	0,9	19	33,3	0,330	0,9	19	41,0	0,507	0,9	
21	48,7	0,382	0,9	20	38,5	0,223	0,9	20	<u>16,4</u>	<u>0,289</u>	0,9	
22	28,4	0,409	0,9	21	26,1	0,235	2,6	21	34,2	0,339	0,9	
23	38,8	0,378	1,7	22	<u>32,5</u>	<u>0,060</u>	3,5	22	13,8	0,408	1,7	
Gemid= delde	40,21			30,8				24,9				
Steek= proef= grootte	118			118				118				
K-R 8	0,794			0,448				0,742				
Stan= daard= metings= fout	1,866			1,858				1,650				

TABEL 4

VORM d

STATISTIESE GEGEWENS VAN DIE EKSPERIMENTELE TOEPASSING

	Toets 1				Toets 2				Toets 3			
	Nr.	M-waarde	$r_{it}$	%NB	Nr.	M-waarde	$r_{it}$	%NB	Nr.	M-waarde	$r_{it}$	%NB
	4	80,7	0,306	0,0	3	69,8	0,498	0,0	3	50,0	0,487	0,0
	5	47,9	9,634	0,0	4	73,7	0,463	0,0	4	37,9	0,475	0,0
	6	37,1	0,508	0,0	5	72,6	0,276	0,0	5	20,0	0,435	0,0
	7	37,9	0,605	0,0	6	29,3	0,391	0,0	6	45,3	0,548	0,0
	8	62,7	0,402	0,0	7	71,4	0,412	0,0	7	58,6	0,443	0,0
	9	40,3	0,531	0,0	8	76,9	0,343	0,0	8	23,3	0,368	0,0
	10	33,9	0,577	0,0	9	43,1	0,409	0,0	9	<u>19,8</u>	<u>0,307</u>	0,0
	11	26,3	0,337	0,0	10	34,7	0,502	0,0	10	20,9	0,525	0,0
	12	27,6	0,561	0,9	11	71,2	0,421	0,0	11	<u>8,4</u>	<u>0,502</u>	0,0
	13	<u>29,1</u>	<u>0,083</u>	0,9	12	<u>17,8</u>	<u>0,004</u>	0,0	12	48,7	0,482	0,0
	14	70,3	0,393	0,8	13	36,1	0,313	0,0	13	21,2	0,290	0,0
	15	44,1	0,631	0,8	14	33,1	0,435	0,0	14	<u>13,4</u>	<u>0,079</u>	0,0
	16	22,9	0,306	0,8	15	83,9	0,254	0,0	15	66,7	0,488	0,0
	17	35,3	0,488	0,9	16	<u>16,9</u>	<u>0,197</u>	0,0	16	37,6	0,499	0,0
	18	<u>19,5</u>	<u>0,310</u>	0,9	17	57,1	0,538	0,0	17	<u>19,3</u>	<u>0,359</u>	0,0
	19	22,4	0,259	0,9	18	40,9	0,443	0,0	18	33,3	0,609	0,9
	20	35,7	0,411	1,8	19	<u>10,1</u>	<u>0,270</u>	0,0	19	<u>8,0</u>	<u>0,319</u>	3,5
	21	33,9	0,454	1,7	20	31,4	0,300	0,8	20	42,6	0,458	3,5
	22	35,1	0,352	4,4	21	37,6	0,283	0,9	21	33,6	0,465	7,3
	23	<u>3,5</u>	<u>0,108</u>	5,3	22	29,7	0,482	0,8	22	22,0	0,209	9,2
Gemid= delde	37,06				46,87				31,53			
Steek= proef= grootte	119				119				119			
K-R 8	0,847				0,764				0,842			
Stan= daard= metings= fout	1,793				1,843				1,723			

TABEL 5  
VORM A  
STATISTIESE GEGEWENS VAN DIE GESELEKTEERDE ITEMS

	Toets 1						Toets 2					
	Nr.	M- waarde	r <sub>it</sub>	Waar gevind	Item= nommer in nuwe toets	%NB	Nr.	M- waarde	r <sub>it</sub>	Waar gevind	Item= nommer in nuwe toets	%NB
1	83,5	0,497	T1 a4*	1	0,8	1	84,3	0,386	T2 c18	10	0,9	
2	77,6	0,430	T1 c14	2	0,9	2	83,9	0,254	T2 d15	8	0,0	
3	69,2	0,422	T1 c8	11	0,0	3	71,2	0,421	T2 c11	13	0,0	
4	59,0	0,523	T1 b17	3	1,7	4	69,8	0,498	T2 d3	1	0,0	
5	58,0	0,631	T1 a15	7	1,7	5	67,5	0,433	T3 a13	18	0,8	
6	48,7	0,382	T1 c21	8	0,9	6	66,7	0,488	T3 d15	20	0,0	
7	47,9	0,634	T1 d5	6	0,0	7	65,9	0,528	T3 a8	17	0,0	
8	45,5	0,512	T1 a6	12	0,8	8	58,2	0,314	T2 a7	7	0,0	
9	44,1	0,631	T1 d15	9	0,8	9	53,8	0,446	T2 a4	2	0,0	
10	42,0	0,364	T1 a11	5	0,8	10	53,6	0,409	T2 c4	4	0,0	
11	46,2	0,629	T1 a18	13	2,6	11	52,5	0,304	T2 a10	11	0,0	
12	38,8	0,378	T1 c23	20	1,7	12	50,0	0,447	T2 a14	12	0,8	
13	37,9	0,605	T1 d7	15	0,0	13	45,8	0,468	T2 b5	3	0,0	
14	37,1	0,442	T1 c9	14	0,0	14	43,1	0,409	T2 d9	15	0,0	
15	36,2	0,362	T1 b19	4	1,7	15	41,0	0,507	T2 c19	19	0,9	
16	35,3	0,488	T1 d17	17	0,9	16	37,6	0,283	T2 d21	16	0,9	
17	35,3	0,456	T1 c10	10	0,0	17	29,3	0,391	T2 d6	6	0,0	
18	35,1	0,352	T1 d22	18	4,4	18	28,9	0,483	T2 a9	5	0,0	
19	34,8	0,323	T1 c12	16	0,9	19	28,7	0,333	T2 a12	14	0,0	
20	28,4	0,409	T1 c22	19	0,9	20	24,3	0,337	T2 c8	9	0,0	
Gemid= delde	47,02	0,449				52,8	0,427					

\* T1 = Toets 1

a4 = Vorm a item 4

TABEL 6  
VORM B  
STATISTIESE GEGEWENS VAN DIE GESELEKTEERDE ITEMS

	Toets 1					Toets 2					
	Nr.	M- waarde	$r_{it}$	Waar gevind	Item= nommer in nuwe toets	%NB	Nr.	M- waarde	$r_{it}$	Waar gevind	Item= nommer in nuwe toets
1	80,8	0,377	T1 a14	2	1,7	1	76,9	0,343	T2 d8	9	0,0
2	68,9	0,359	T1 b4	1	0,0	2	75,6	0,421	T3 a17	20	0,8
3	65,3	0,442	T1 b14	3	0,8	3	73,7	0,463	T2 d4	2	0,0
4	62,7	0,402	T1 d8	11	0,0	4	72,6	0,276	T2 d5	4	0,0
5	60,8	0,680	T1 a17	9	2,5	5	69,8	0,498	T2 d3	1	0,0
6	55,1	0,561	T1 a7	15	0,8	6	69,8	0,427	T3 c11	18	0,0
7	51,7	0,614	T1 c15	7	0,9	7	64,3	0,334	T2 b20	10	2,6
8	51,7	0,609	T1 c5	6	0,0	8	62,9	0,315	T3 c9	17	0,0
9	47,0	0,475	T1 c6	12	0,0	9	58,2	0,314	T2 a7	7	0,0
10	45,4	0,493	T1 a20	17	2,5	10	57,1	0,538	T2 d17	16	0,0
11	44,9	0,345	T1 b12	16	0,8	11	50,4	0,415	T3 b12	19	1,7
12	41,9	0,423	T1 b20	8	1,7	12	49,6	0,519	T2 a19	15	2,6
13	40,3	0,531	T1 d9	10	0,0	13	47,2	0,433	T2 a6	6	0,0
14	36,0	0,428	T1 b21	13	2,6	14	45,5	0,441	T2 a15	8	0,8
15	30,7	0,411	T1 d20	4	1,8	15	35,3	0,349	T2 c5	3	0,0
16	35,1	0,352	T1 d22	18	4,4	16	34,7	0,502	T2 d10	11	0,0
17	33,9	0,577	T1 d10	14	0,0	17	33,3	0,330	T2 c19	13	0,9
18	28,4	0,409	T1 c22	19	0,9	18	29,7	0,482	T2 d22	12	0,8
19	27,6	0,561	T1 d12	5	0,9	19	26,1	0,235	T2 c21	14	2,6
20	25,0	0,404	T1 b23	20	6,3	20	23,3	0,560	T2 b7	5	0,0
Gemid= delde	46,7	0,415				52,8	0,424				

TABEL 7

VORM A

STATISTIESE GEGEWENS VAN DIE GESELEKTEERDE ITEMS VAN DIE  
FINALE TOETS

	Toets 1				Toets 2			
	Nr.	M-waarde	$r_{it}$	%NB	Nr.	M-waarde	$r_{it}$	%NB
1	70,1	0,381	0	1	61,6	0,266	0	
2	74,6	0,379	0	2	53,2	0,389	0	
3	57,4	0,468	0	3	36,0	0,429	0	
4	52,9	0,418	0	4	29,9	0,334	0	
5	30,1	0,241	0	5	28,3	0,397	0	
6	40,4	0,576	0	6	29,7	0,247	0	
7	53,9	0,556	0	7	58,2	0,253	0	
8	56,5	0,451	0	8	82,4	0,399	1	
9	46,4	0,583	0	9	30,2	0,225	1	
10	39,3	0,462	1	10	70,3	0,429	1	
11	67,1	0,396	2	11	51,3	0,274	2	
12	38,9	0,453	6	12	41,8	0,429	2	
13	43,9	0,505	8	13	72,2	0,415	3	
14	25,9	0,375	13	14	36,5	0,208	5	
15	34,6	0,438	17	15	47,0	0,369	5	
16	19,9	0,212	19	16	33,5	0,291	6	
17	33,2	0,415	20	17	64,7	0,529	7	
18	38,8	0,331	21	18	60,3	0,551	11	
19	26,1	0,157	23	19	46,7	0,473	17	
20	34,9	0,411	26	20	61,9	0,460	23	
Gemid= delde	44,2			49,8				

Gemiddelde M-waarde van toets 1 + toets 2 = 47,0

TABEL 8

VORM B

STATISTIESE GEGEWENS VAN DIE GESELEKTEERDE ITEMS VAN DIE  
FINALE TOETS

	Toets 1			Toets 2			
	Nr. M-waarde	$r_{it}$	%NB	Nr. M-waarde	$r_{it}$	%NB	
1	64,6	0,423	0	1	50,2	0,328	0
2	64,8	0,407	0	2	64,0	0,410	0
3	73,1	0,412	0	3	22,6	0,362	0
4	20,0	0,331	0	4	58,8	0,440	0
5	28,8	0,483	0	5	18,4	0,294	0
6	44,8	0,606	0	6	27,6	0,339	1
7	54,1	0,577	1	7	64,0	0,227	1
8	43,6	0,496	1	8	44,7	0,311	2
9	49,2	0,614	1	9	59,4	0,371	2
10	42,7	0,412	1	10	53,3	0,450	2
11	51,1	0,439	4	11	21,5	0,325	6
12	42,4	0,397	7	12	20,0	0,223	6
13	34,0	0,511	8	13	37,7	0,381	7
14	29,2	0,488	11	14	18,9	0,271	9
15	46,8	0,466	13	15	44,2	0,399	9
16	31,7	0,365	15	16	55,1	0,418	9
17	44,3	0,542	20	17	52,8	0,482	11
18	22,8	0,272	21	18	62,4	0,487	12
19	19,9	0,212	22	19	48,1	0,488	14
20	29,9	0,347	23	20	60,9	0,452	19
Gemid= delde	41,9			44,2			

Gemiddelde M-waarde van toets 1 + toets 2 = 43,1

TABEL 9

VERDELING VAN PUNTE DEUR DIE LEERLINGE BEHAAL VIR VORM A  
(TOETSE 1 EN 2) (N = 495)

Waarde	Frekwensie	Kummulatiewe frekwensie	Persentasie frekwensie	Persentasie kummulatiewe frekwensie
2	1	1	0,202	0,202
3	1	2	0,202	0,404
4	2	4	0,404	0,808
5	5	9	1,010	1,818
6	3	12	0,606	2,424
7	7	19	1,414	3,838
8	3	22	0,606	4,444
9	13	35	2,626	7,071
10	16	51	3,232	10,303
11	12	63	2,424	12,727
12	33	96	6,667	19,394
13	26	122	5,253	24,646
14	37	159	7,475	32,121
15	31	190	6,263	38,384
16	32	222	6,465	44,848
17	27	249	5,455	50,303
18	26	275	5,253	55,556
19	31	306	6,263	61,818
20	26	332	5,253	67,071
21	22	354	4,444	71,515
22	18	372	3,636	75,151
23	16	388	3,232	78,384
24	19	407	3,838	82,222
25	11	418	2,222	84,444
26	12	430	2,424	86,869
27	14	444	2,828	89,697
28	10	454	2,020	91,717
29	10	464	2,020	93,737
30	11	475	2,222	95,959
31	4	479	0,808	96,768
32	8	487	1,616	98,384
33	1	488	0,202	98,586
34	3	491	0,606	99,192
35	1	492	0,202	99,394
36	1	493	0,202	99,596
37	1	494	0,202	99,798
39	1	495	0,202	100,000

Aantal gevalle = 495

Gemiddelde = 18,21

Standaardafwyking = 6,575

Skeefheid = 0,380

Kurtose = 2,808

## BRONNELYS

- ADAMS, G.S. *Measurement and evaluation in education and guidance*. New York: Holt Reinhart and Winston, 1964.
- ADLER, I. Some thoughts about curriculum revision. *The Mathematics Teacher*, Vol 56(7), 1963.
- AHMAND, J., STANLEY and GLOCK, MARVIN, D. *Evaluating pupil growth. Principles of test measurement*. Boston: Allyn and Bacon, Inc., 1967.
- AHMANN, J.S. *Testing student achievement and aptitudes*. Washington: The Center for Applied Research in Education Inc., 1962.
- AIKEN, L.R. *Psychological and educational testing*. Boston: Allyn and Bacon Inc., 1971.
- ALBERTS, N.F. *Die ontwikkeling van 'n beroepsbelangstellingsvraelys, gebaseer op aktiwiteitsbelangstelling*. D.Phil.-proefskrif, Universiteit van Pretoria, 1967.
- AMERICAN PSYCHOLOGICAL ASSOCIATION. *Standards for educational and psychological tests and manuals*. Washington: APA, 1966.
- ANASTASI A. *Psychological testing*. New York: Macmillan Co. Ltd, 1964.
- ANASTASI A. *Psychological testing*. New York: Macmillan Co. Ltd, 1971.
- ANSTEY E. The d-method of item-analysis. *The British Journal of Psychology* (Statistical Section), 1948: 167-177
- ANSTEY, E. *Psychological tests*. London: Nelson, 1966.
- BROWN, F.G. *Principles of Educational and Psychological Testing*. Illinois: The Dryden Press, Inc., 1970.

- CHAMBERLAIN, J.C. en GERICKE, F.W. *'n Handleiding vir die Bepanning van Objektiewe Skolastiese Prestasietoetse en vir die skryf van Veelkeusige Items.* Raad vir Geesteswetenskaplike Navorsing, 1975.
- CHAMBERLAIN, J.C. *The Role of Certain Biographical Variables in the Academic Proficiency of Bantu Students and the standardization of a Proficiency test battery.* D.Ed.-proefskrif, Pretoria: Universiteit van Suid-Afrika, 1977.
- CONRAD, H.S., "Characteristics and uses of item-analysis data" *Psychological Monographs* 62 (295), 1945, 1-48.
- CRONBACH, L.J. Test validation. In: Thorndike R.L. ed. *Educational Measurement.* Washington: American Council on Education, 1971.
- DAVIS, B. Item selection techniques. In: Lindquist, E.F. ed. *Educational measurement.* Washinton: American Council on Education, 1963. 266-328.
- DEPARTEMENT VAN INLIGTING. *The progress of the Bantu peoples towards nationhood.* Pretoria: Staatsdrukker, (-).
- DEPARTEMENT VAN ONDERWYS EN OPLEIDING. *Handleiding en Sillabusse vir die Hoër Primêre Eksamen.* Pretoria, 1975.
- DEPARTEMENT VAN ONDERWYS EN OPLEIDING. *Jaarverslag 1976.* Uitgegee op gesag 1977.
- DEPARTEMENT VAN ONDERWYS EN OPLEIDING. *Jaarverslag.* Pretoria, 1978.
- DREVER, J. *A Dictionary of Psychology,* Great Britain 1974.
- DU TOIT, J.M. *Statistiese Oefeninge en Tabelle vir die gebruik saam met Statistiese Metodes.* Kosmo Uitgewery Edms. Bpk., Stellenbosch, 1971.
- EBEL, R.L. *Measuring educational achievement.* Englewood Cliffs, New Jersey, Prentice-Hall, Inc., 1965.

- EBEL, R.L. *Essentials of educational measurement*. Englewood Cliffs, New Jersey: Prentice-Hall., 1972.
- FOURIE, A.B. *Ondersoek na die implikasie van gestandaardiseerde toetse vir St. vi-Bantoeleerlinge*. Pretoria: Ongepubliseerde D.Phil-proefskrif, Universiteit van Pretoria,
- GARRET, H.E. and WOODWORTH, R.S. *Statistics in Psychology and Education*. Toronto: Longmans, Green and Co., 1964.
- GERICKE, F.W. *Die Toepasbaarheid van die ATTR (Algemene Toetse in Taal en Rekenkunde) op Studente aan Bantoe-Onderwyskolleges*. M.Ed-verhandeling, Potchefstroom: Potchefstroomse Universiteit vir Christelike Hoër Onderwys, 1972.
- GERICKE, J.S. *Verbale en Nie-verbale IK ten opsigte van prestasiekriteria*. M.A.-verhandeling, Stellenbosch: Universiteit van Stellenbosch, 1963.
- GHISELLI, E.E. *Theory of psychological measurement*. New York: McGraw-Hill, 1964.
- GLASS, G.V. and STANLEY, J.C. *Statistical methods in Education and Psychology*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970.
- GODDARD, H.H. *Feeble-Mindedness its causes and consequences*. New York, The Macmillan Company, 1920.
- GOWNS, L.A. e.a. *Psigologie-Woordeboek*. McGraw-Hill Boekmaatskappy, Johannesburg, 1979.
- GRONLUND, Norman E. *Measurement and evaluation in teaching*. New York: Macmillan, 1965.
- GRUNDLING, D.L. *Ondersoek na die betroubaarheid en geldigheid van die nuwe Suid-Afrikaanse individuele skaal by 'n groep verstandelik vertraagde en swaksinnige seuns*. Potchefstroomse Universiteit vir Christelike Hoër Onderwys, 1967 (MSc-verhandeling.)

- GUILFORD, J.P. *Psychometric Methods*. New York: McGraw-Hill Book Company, 1954.
- GUILFORD, J.P. *Fundamental statistics in Psychology and Education, 3rd edition*. New York, McGraw-Hill, 1956.
- GUILFORD, J.P. *Fundamental statistics in Psychology and Education*. New York: McGraw-Hill Book Company, 1965.
- GUILFORD, J.P. *Psychometric Methods*. New York: McGraw-Hill, 1966.
- GUILFORD, J.P. *The Nature of human intelligence*. New York, McGraw-Hill Book Company, 1967.
- GUILFORD, J.P. "The structure of intelligence" in *Handbook of measurement and assessment in behavioral sciences*. edited by Dean, K. Whitla, Reading. Massachusetts: Addison Wesley Publishing Co., 1968.
- GUNDERSON, R.O. and FELDT, L.S. Relationship between Verbal and Non Verbal Intelligence scores to achievement. *Journal of Educational Psychology*, 51. 1960, 115-121.
- HANDLEIDING vir die Skolastiese aanlegtoets vir Bantoes (SATB). Raad vir Geesteswetenskaplike Navorsing, 1971.
- HAYS, W.L. *Statistics for psychologist*. New York: Holt, Reinhart and Winston, 1965.
- HEINICHEN, F.W.O. *Die standaardisering van 'n Groepsintelligensie-toets vir Indiërleerlinge in die Senior klasse van die Hoërskool*. D.Ed.-proefskrif, Pretoria: Universiteit van Suid-Afrika, 1970.
- HELMSTADTER, G.C. *Principles of psychological measurement*. New York: Appleton-Century-Crofts, 1964.
- HELMSTADTER, G.C. *Research concepts in human behavior*. New York: Appleton-Century-Crofts, 1970.
- HUYSAMEN, G.K. *Beginsels van Sielkundige meting*. Pretoria: Academica, 1978.
- KERLINGER, FRED N., *Foundation of behavioral research*. New York: Holt Reinhart and Winston, 1964.

- KUDER, G.F. and RICHARDSON, M.W. The theory of the estimation of test reliability. *Psychometrika*, Vol. 2, 1937.
- LÖTZ, J. VAN Z. *Die standaardisering van individuele leestoetse vir Afrikaanssprekende leerlinge in die junior klasse*. D.Ed.-proefskrif, Pretoria: Universiteit van Suid-Afrika, November 1967.
- LYMAN, HOWARD B. *Test scores and what they mean*. Englewood-Cliffs, New Jersey: Prentice-Hall, 1963.
- MADGE, E.M. *Geldigheid*. Interne Opleidingslesing, Raad vir Geesteswetenskaplike Navorsing, 1976.
- MICHAU, J.M.Z. Problem Areas in the Acquisition of Mathematical concepts by Black children in South Africa. *Journal of Education*, 10, University of Natal, 1978.
- MORGAN, C.T. *Introduction to psychology*. The University of Wisconsin. McGraw-Hill Book Company, Inc., New York, Toronto, London: 1961.
- NUNNALLY, J.C. *Educational measurement and evaluation*. New York: McGraw-Hill Book Company, 1964.
- NUNNALLY, J.C. *Psychometric theory*. New York: McGraw-Hill Book Company, 1967.
- NUNNALLY, J.C. *Introduction to psychological measurement*. New York: McGraw-Hill Book Company, 1970.
- OOSTHUIZEN, S. *Geldigheid*. Interne Opleidingslesing, Pretoria: Raad vir Geesteswetenskaplike Navorsing, 1973.
- OWEN, K. *Opstelling en Standaardisering van die Akademiese - Tegniese Aanlegtoetse vir Kleurlinge (tegniese verslag)*. Pretoria: Raad vir Geesteswetenskaplike Navorsing, 1971.
- RAAD VIR GEESTESWETENSKAPLIKE NAVORSING. *Katalogus van Toetse*. Instituut vir Psigometriese Navorsing 1977.
- REMMERS, H.H. and GAGNE, N.L. *Educational measurement and evaluation*. Revised edition. New York: Harper and Brothers, 1955.
- SNEDECOR, G.W. *Statistical methods*. The Iowa State College Press, 1950.
- SONNEKUS, M.C.H. *Akademiese prestasietoetse en hulle betekenis by die vakkeuse van universiteitbeginners*. Kaapstad: H.A.U.M. 1958-

STEYN, G.H.A. "Die Plek van Wiskunde en Algemene Wiskunde in die huidige Onderwysprogram". *Onderwysbulletin*. Vol. 5(4), 1960.

STILSON, D.W. *Probability and statistics in psychological research and theory*. San Francisco, California: Holden-Day Inc., 1966.

SWANEPOEL, H.F. *’n Psigometriese ondersoek na die geldigheid en gebruik van die voorligtingstoets vir Junior Sekondêre Bantoeleerlinge in Vorm III*. D.Phil-proefskrif, Potchefstroom: Potchefstroomse Universiteit vir Christelike Hoër Onderwys, 1975.

SWART, D.J. *Itemontleding en -seleksie*. Raad vir Geesteswetenskaplike Navorsing, 1978.

SWART, D.J. *Betroubaarheid*. Raad vir Geesteswetenskaplike Navorsing, 1980.

THORNDIKE, R.L. Some methodological issues in the study of creativity. In: *Proceedings of the 1962 invitational conference on testing problems*. Princeton, N.J.: Educational Testing Service, 1963.

VAN DER WALT, J.S. *Opvoedkundige en Psigologiese Meting*. *’n Psigometries Statistiese Analise*. Stellenbosch, 1970.

VAN ROOY, A.J. *Die onderwys van Wiskunde, Algemene Wiskunde en Rekeningkunde aan die openbare middelbare en hoërskole vir Blankes in Suid-Afrika*. Pretoria: Nasionale Buro vir Opvoedkundige en Maatskaplike Navorsing (Navorsingsreeks nr. 24), 1965.

VAN STADEN, J.D. *Die Keuring van ’n groep begaafde standerd vyf bantoeleerlinge vir versnelde bevordering en die implikasies daarvan*. Ongepubliseerde M.Ed.-verhandeling, Pretoria: Universiteit van Suid-Afrika, 1972.

VAN STADEN, J.D. *Die Begaafde Bantoeleerling - ’n Psigodiagnostiese studie*. D.Ed.-proefskrif, Pretoria: Universiteit van Suid-Afrika, 1975.

VERWEY, F.A. *Die Betroubaarheid van Toetse*. Interne lesing, Raad vir Geesteswetenskaplike Navorsing, 1975.

VON MOLLENDORF, J.W. *Sielkundige Meting en Aspekte van Skoolastiese Prestasie in die ontwikkeling van 'n voorligtingsprosedure vir Swart skoliere in vorm V*. D.Phil-proefskrif, Potchefstroom: Potchefstroomse Universiteit vir Christelike Hoër Onderwys, 1978.

WECHSLER, D. *The measurement and appraisal of adult intelligence*. Fourth edition. London: Baillière, 1958.

WHITLA, D.K. *Handbook of Measurement and Assessment in Behavioral Sciences*. Addison-Wesley Publishing Co., California, 1968.

WILLMOTT, S. and NUTTEL, D.L. *Reliability of examinations at 16*. School Council Research Studies. Macmillan 1975.

\*\*\*

## SUMMARY

### ACHIEVEMENT IN MATHEMATICS, VERBAL AND NON-VERBAL REASONING OF BLACK STANDARD 5 PUPILS: A PSYCHOMETRIC INVESTIGATION

#### 1. INTRODUCTION

According to Van Staden (1975, p.viii) there is an urgent need for psychological and scholastic achievement tests for Black ethnic groups. The Department of Education and Training was aware of this need and approached the HSRC to standardize achievement tests in Mathematics to be used for initial evaluation in the various standards.

This study makes a contribution towards meeting this need by supplying such a test in Mathematics.

#### 2. AIMS

##### 2.1 ACHIEVEMENT TEST IN MATHEMATICS

This investigation consists of two sections. The first section, which constitutes the greater part of this study, consists of the development of an achievement test in Mathematics which is used as a measuring instrument in the second part of the study. This implies that various hypotheses are not stated in the first section.

2.2 In the second section of the study the relationship between Verbal and Non-Verbal Reasoning on the one hand and achievement in Mathematics on the other is investigated.

In this section the following hypotheses will be investigated:

- a)  $H_0 : \rho_1 = \rho_2$  The null hypothesis is stated that there is no difference in the predictive power of Verbal and Non-verbal Reasoning for achievement in Mathematics by Black pupils.

- $H_1 : \rho_1 \neq \rho_2$  The alternative hypothesis is stated that there is a difference in the predictive power of Verbal and Non-Verbal Reasoning for achievement in Mathematics by Black pupils.
- b)  $H_0 : \mu_1 = \mu_2$  A second null hypothesis is stated that there is no difference in the mean achievement in Mathematics by Black boys and girls in standard 5.
- $H_1 : \mu_1 \neq \mu_2$  The alternative hypothesis is stated that there is a difference in the mean achievement in Mathematics by Black boys and girls in standard 5.
- c)  $H_0 : \rho = 0$  A third null hypothesis is stated that intelligence is not a predictor of achievement in Mathematics by Black pupils.
- $H_1 : \rho \neq 0$  The alternative hypothesis is stated that intelligence is a predictor of achievement in Mathematics by Black pupils.

### 3. RATIONALE OF THE MEASURING INSTRUMENTS

#### 3.1 RATIONALE OF THE ACHIEVEMENT TEST IN MATHEMATICS

The rationale of the test is based on the assumption that the ability of a Standard 5 pupil to manipulate numbers and to solve mathematical problems is a valid criterion of his achievement in Mathematics.

#### 3.2 RATIONALE OF THE VERBAL REASONING TEST

The test is based on the assumption that the ability to do basic operations and/or reach logical conclusions in connection with verbally stated analogies, classifications, ratios and computational problems is a valid indication of verbal reasoning ability.

### 3.3 RATIONALE OF THE NON-VERBAL REASONING TEST

The test is based on the assumption that the ability to reason deductively, i.e. to arrive at a general principle from given data and to complete a series by applying the principle, is an indication of non-verbal reasoning ability.

## 4. DESCRIPTION OF THE MEASURING INSTRUMENTS

### 4.1 MATHEMATICS ACHIEVEMENT TEST

The test consists of Forms A and B. Each form consists of two tests:

- (i) Test 1: Ready Knowledge
- (ii) Test 2: Comprehension

### 4.2 VERBAL REASONING TEST

In this test verbal subject-matter is used to measure the ability to form and logically manipulate verbal concepts. The content is heterogeneous and it includes tasks such as the completion of analogies and number rows, classifications, logical conclusions and arithmetical problems.

### 4.3 NON-VERBAL REASONING TEST (Pattern Completion)

The items consist of incomplete patterns to be completed by choosing one of five given options.

## 5. METHOD OF INVESTIGATION

### 5.1 METHOD USED IN THE FIRST PART OF THIS INVESTIGATION

After a comprehensive literature study on aspects concerning this study had been done, the standardization of the Mathematics Achievement Test (chapter 3) was commenced.

The Mathematics Achievement Test was compiled by the researcher

and submitted to the internal committee for Mathematics of the HSRC for comments and approval. After the test had been edited according to recommendations, it was applied to a representative sample of Standard 5 pupils in the Pretoria East area. The results were processed, item analysis and item selection were done and approved by the internal committee for Mathematics.

The selected items for the establishment of norms were applied to a representative, stratified sample of Black Standard 5 pupils in March 1979. The results were processed and norms established.

## 5.2 METHOD USED IN THE SECOND PART OF THIS INVESTIGATION

The investigation group was for purposes of economy limited to Standard 5 pupils in 38 schools in the Mamelodi residential area near Pretoria.

After the application of the Mathematics test by the researcher, the Pattern Completion and Verbal Reasoning Tests were applied by two Black inspectors. The test instructions for Mathematics were in English and those of the other two tests in Northern Sotho and Zulu.

The same answer sheet, which is scored mechanically, was used for both parts of this investigation.

## 6. DISCUSSION OF RESULTS

The standardization of the Mathematics measuring instrument established that the Mathematics Achievement Test was valid and reliable measuring instrument as well as a good predictor of mathematics achievement at the end of the year. The following findings have also been made:

- a) There is no significant difference between the mathematical achievement of boys and girls.
- b) There is no significant difference between the mean mathe=

mathematical achievement of urban and rural pupils.

- c) However, certain Black language groups achieve significantly higher (on the 5%-level) in Mathematics. According to the means, the Xhosas were the best achievers in Mathematics.

The second objective of this investigation was to establish whether Verbal Reasoning, Non-Verbal Reasoning and Intelligence are predictors of mathematics achievement and to establish which of the two, Verbal Reasoning and Non-Verbal Reasoning, is the better predictor. The role of sex in mathematics achievement was also investigated. The latter has already been reported in (a).

It has been found that Verbal Reasoning is a better predictor of mathematics achievement than Non-Verbal Reasoning and that Intelligence, as jointly measured by the Verbal Reasoning Test and Non-Verbal Reasoning Test, is a good predictor of mathematics achievement.

It seems from this investigation that Verbal Reasoning, Non-Verbal Reasoning and Intelligence are all predictors of mathematics achievement, with Verbal Reasoning as the best predictor, followed by Intelligence and then Non-Verbal Reasoning.