

HOOFSTUK 5

METODE VAN ONDERSOEK

5.1 INLEIDING

Hierdie hoofstuk gee 'n uiteensetting van die algemene ontwikkeling van die navorsingsprogram, die ondersoekgroepe, instrumente, statistiese tegnieke en algemene metodologiese oorwegings. Voorlopige resultate word weergegee en ontleed met die oog op die vasstelling van gewenste procedures en tegnieke vir die finale uiteensetting van resultate, stel van gevolgtrekkings en formulering van riglyne wat musikale aanlegtoetsing in die toekoms betref.

Die drie stadia waarin die navorsingsprogram verloop het, was

- (1) voorbereiding: verkenningsstudie of propedeuse,
 - (2) vooronderzoek en
 - (3) hoofonderzoek wat ooreenstem met die drie verskillende toetstoepassings wat Henrysson (in Thorndike, 1971, p.131) die "pre tryout, tryout and trial administration" noem.
- Die propedeuse of voorbereiding behels die voorafgaande informele toepassing van toetsitems op 'n relatief klein aantal toetslinge uit die populasie waarop die volle toetsprogram uiteindelik uitgevoer sal word. Die algemene doel was om die Bentleytoets se algemene bruikbaarheid en betrouwbaarheid in terme van die Suid-Afrikaanse skoolsituasie te probeer bepaal.
 - Die vooronderzoek: 'n formele ondersoek word onderneem om meer akkurate inligting oor die Bentleytoets as 'n geheel, die vier subtoetse en sestig toetsitems te verkry. 'n Ondersoekgroep word gebruik wat groot en verteenwoordigend genoeg is vir verskillende vlakke van ouderdom en standerd, sowel as vergelyking ten opsigte van taal en geslag. Sodoende kan 'n gedaallieerde statistiese analise uitgevoer en moontlike gebreke in die meetingsprosedure (bv. dubbelsinnige instruksies) na vore gebring word.
 - Die hoofonderzoek. Met die voorafgaande versamelde inligting, toetsresultate en ervaring ten opsigte van toetsadministrasie en diverse praktiese omstandighede as basis, word 'n gestratifiseerde steekproef gebruik en die

finale deel van die navorsingsprogram onderneem. Verskillende toetse word vir vergelykingsdoeleindes gebruik en volledige statistiese bewerkings gedoen. Hieruit volg die ontleding van resultate, gevolgtrekkings en aanbevelings.

5.2 PROPEDEUSE - VOORBEREIDENDE NAVORSING

Verskeie eksperimente is met groepe skoolleerlinge en studente uitgevoer met die Bentleytoets en -handleiding as belangrikste navorsingsmateriaal en algemene leidraad. Drie sodanige proefnemings word hieronder gerapporteer.

5.2.1 Eksperimentele groep I

5.2.1.1 Doelstellings

- Om die algemene verspreiding van musikale aanleg binne 'n bepaalde klasgroep, soos in die toetsprestasies weerspieël, te ondersoek.
- 'n Ondersoek na geslag as faktor by toetsprestasie en algemene skolastiese prestasie.
- 'n Ondersoek na die moontlike verband tussen IK, vakprestasie en musikale aanleg soos blyk uit toetsresultate (e.g. volgens gegewens deur die betrokke skool verskaf).
- Die bepaling van die mate van korrelasie tussen subtoetstellings onderling, asook tussen subtoetstellings en die toetstotaal ten einde die mate van onafhanklikheid van die subtoetse vas te stel.
- 'n Ondersoek na intervalgrootte as faktor in toonhoogteonderskeiding.

5.2.1.2 Die ondersoeksgroep

Die ondersoeksgroep het bestaan uit 'n standerd 4- klasgroep met vyftien seuns en vyftien dogters van wie die gemiddelde ouderdom elf jaar en drie maande was. Wat algemene verstandsvermoë en akademiese prestasie betref, blyk die betrokke klasgroep betreklik heterogene te wees, soos uit tabel 5.1 blyk. Volgens Guilford (1956, p.73) verskaf die verhouding variasiebreedte/

standaardafwyking naamlik 'n aanduiding van heterogeniteit:

$$\frac{\text{var.br.}}{s} > 3,5 = \text{heterogene verspreiding}$$

TABEL 5.1

VARIASIEBREEDTE EN STANDAARDAFWYKING IN IK EN VAKGEMIDDELDES AS MAATSTAF VAN HETEROGENITEIT (N = 30)

	\bar{x}^*	Min.	Maks.	Var.br.	s^*	$\frac{\text{var.br.}}{s}$
IK	108,433	90	131	41	7,788	5,265
Afrikaans %	67,300	51	88	37	9,668	3,827
Engels %	64,500	47	88	41	10,187	4,025
Rekenkunde %	68,267	42	92	50	13,299	3,760
Ander vakke %	69,267	44	96	52	12,974	4,008
Gemiddelde totaal %	67,467	46	90	44	10,626	4,141

5.2.1.3 Die instrument (vide p.91, 138)

Die Bentleytoets bestaan uit vier subtoetse:

- I. Toonhoogteonderskeiding: 20 items; leerling moet na 2 opeenvolgende tone luister, dit vergelyk en aandui of die tweede toon hoër, laer of dieselfde as die eerste is (maksimumtelling: 20).
- II. Melodiese geheue: 10 vyftoonmelodieë word voorgespeel en telkens direk herhaal; leerlinge moet aandui of die herhaling van elke melodie diezelfde is, indien nie, watter toon verander het (maksimumtelling: 10).
- III. Akkoordanalise: 20 akkoorde word voorgespeel en leerling moet aandui of elke akkoord uit twee, drie of vier tone saamgestel is (maksimumteling: 20).
- IV. Ritmiese geheue: 10 ritmiese patronen, wat elk uit 4 maatslae bestaan, word voorgespeel en dan direk herhaal; leerling moet aandui of die her-

*) \bar{x} = gemiddelde telling
 s = standaardafwyking (vide p.141)

haling dieselfde is, indien nie, watter maatslag verander het (maksimumtelling 10).

Die toets is in sy geheel met behulp van 'n klavier gegee, wat afwyk van die oorspronklike Bentleytoets in die volgende opsigte:

By Subtoets I maak Bentley van spesiaal gekalibreerde sinusgolf-ossillators gebruik en word intervalle so klein as $\frac{3}{26}$ ste van 'n halftoon ingesluit. By die klavier kan geen interval kleiner as 'n halftoon weergegee word nie, terwyl die halftoon die grootste interval in die Bentleytoets is. Vir die huidige doel is dus van voorbeeld van toetsitems uit Bentley se handleiding (1966a, p.53), wat hy gebruik het by die ontwikkeling van sy toets, gebruik gemaak. By Subtoetse II, III en IV gebruik Bentley 'n pyporrel in plaas van 'n klavier. Verder is die toetsinstruksies mondelings gegee en nie met behulp van die Bentley-toetskankbaan nie.

5.2.1.4 Resultate en gevolgtrekkings

- Die gemiddelde toetstellings, variasiebreedte en standaardafwyking word in tabel 5.2 weergegee. Hieruit blyk die volgende:

Met die uitsondering van Subtoets I, toon die toetsresultate 'n betreklik heterogene verspreiding en wye gradering van musikale aanleg in soverre as wat die toets as maatstaf geld.

Subtoets I blyk die maklikste te wees, gevvolg deur Subtoets IV, met Subtoets II en III naby die teoretiese gemiddelde.

Die klavier as medium is blykbaar, in die lig van die hoë gemiddelde prestasie (85,5%), nie diskriminerend genoeg vir toonhoogteonderskeiding (Subtoets I) nie.

TABEL 5.2

TOETSRESULTATE: GEMIDDELDES, VARIASIEBREEDTE EN STANDAARDAFWYKING (N = 30)

Subtoets	\bar{X}	\bar{X} as %	Min.	Maks.	Var.br.	s	$\frac{\text{var.br.}}{s}$
I	17,111	85,555	12	20	8	2,363	3,385
II	4,844	48,440	3	7	4	1,102	3,630
III	9,467	47,335	4	15	11	3,028	3,633
IV	6,556	65,560	3	10	7	1,747	4,007

- Geslag as faktor by musikale aanleg, skolastiese prestasie en IK

Uit tabel 5.3 blyk die volgende: Hoewel die dogters, met die uitsondering van Subtoets IV, 'n hoër gemiddelde prestasie as die seuns gelewer het, is hierdie verskil nie betekenisvol nie.

TABEL 5.3

VERGELYKING TUSSEN TOETSPRESTASIES VAN SEUNS EN DOGTERS

Subtoets	Dogters (N = 15)		Seuns (N = 15)		$\bar{X}_1 - \bar{X}_2$	z *	Betekenisvolheid van verskil tussen gemiddeldes
	\bar{X}_1	s ₁	\bar{X}_2	s ₂			
I	17,423	2,256	16,800	2,426	0,623	0,728	$z \geq 1,960$ (0,05-peil)
II	4,978	1,286	4,711	0,860	0,267	0,668	
III	10,045	2,834	8,889	3,105	1,156	1,065	2,576 (0,01-peil)
IV	6,267	2,004	6,845	1,387	-0,578	0,629	
Totaal	38,713	4,766	37,245	4,987	1,468	0,824	

'n Soortgelyke situasie word aangetref as seuns en dogters ten opsigte van skoolprestasie en IK vergelyk word. Tabel 5.4 toon naamlik dat slegs in die geval van die vakke Afrikaans en Engels, daar 'n betekenisvolle verskil tussen dogters en seuns bestaan.

*) Verskil uitgedruk in standaardeenhede (vide p.143)

TABEL 5.4

VERGELYKING TUSSEN IK EN VAKPRESTASIES VAN SEUNS EN DOGTTERS

	Dogsers (N = 15)		Seuns (N = 15)		$\bar{X}_1 - \bar{X}_2$	z
	\bar{X}_1	s_1	\bar{X}_2	s_2		
IK	107,200	7,521	109,667	7,854	-2,467	-0,712*
Afrikaans %	71,733	9,015	62,867	8,148	8,866	2,826
Engels %	68,267	9,936	60,733	8,970	7,534	2,180**
Rekenkunde %	70,467	12,230	66,067	13,945	4,400	0,919
Ander vakke %	72,467	11,667	66,067	13,419	6,400	1,394
Gemiddelde vaktotaal %	70,867	10,118	64,067	10,016	6,800	1,850

* = beduidend op 0,01-peil

** = beduidend op 0,05-peil

- Verband tussen toetsresultate, IK en vakprestasies

Uit die onderstaande korrelasiematrikse (tabelle 5.5 en 5.6) blyk die volgende: Toetsprestasie en IK

Slegs in die volgende gevalle is 'n beduidende verband aan te toon:

NV-IK en Subtoets II by dogsers ($r = 0,544$) en volle klasgroep ($r = 0,411$); NV-IK en toetstotaal, volle groep ($0,363$).

In verskeie gevalle is die korrelasies by die dogsers negatief, bv. Subtoets I.

Skoolvakprestasie en IK

Skoolvakke en skoolvakkombinasies korreleer in alle gevalle beduidend met totale IK by dogsers en in die meeste gevalle ook by die volle klasgroep (op 0,05-peil).

Toetsprestasie teenoor vakprestasie

Slegs Subtoets II toon 'n beduidende verband met 'ander vakke' en gemiddelde vaktotaal by die volle klasgroep. Subtoets III en IV korreleer deurgaans negatief met skoolvakprestasies by die dogsers. Daarteenoor interkorreleer

TABEL 5.5

KORRELASIES^④: IK, VAK- EN TOETSPRESTASIES VOLGENS GESLAG

	IK NV	IK V	IK Totaal	Afrikaans	Engels	Rekenkunde	Ander Vakke	Vak= totaal	Bentley				Totaal
									I	II	III	IV	
D O G T E R S (N = 15)													
IK-NV	-	+508	838*	457	429	483	337	458	-286	544**	-138	213	174
IK-V	492	-	892*	453	443	455	545**	499	-074	198	164	-144	033
IK-Totaal	870*	853*	-	552**	515**	559**	538**	574**	-184	394	018	014	101
Afrikaans	-011	503	307	-	932*	857*	889*	970*	272	280	-320	-117	-058
Engels	-102	510	249	958*	-	749*	908*	943*	225	383	-364	-064	030
Rekenkunde	-376	442	475	606**	595**	-	767*	909*	250	423	-063	-156	196
Ander vakke	049	221	182	860*	791*	636**	-	943*	188	340	-204	-034	084
Gem. vaktotaal	122	449	346	926*	899*	833*	917*	-	242	396	-242	-049	098
Subtoets I	336	337	387	041	-023	089	-248	-072	-	-132	157	113	403
Subtoets II	340	204	338	337	267	133	420	315	-159	-	-044	692*	705*
Subtoets III	334	055	235	212	171	317	248	272	110	407	-	-138	361
Subtoets IV	172	414	328	439	357	390	444	457	222	292	156	-	809*
Toetstotaal	464	397	504	408	307	369	339	396	496	543**	719*	686*	-
S E U N S (N = 15)													

+ = desimale tekens weggelaat

* = betekenisvol op 0,01-peil

** = betekenisvol op 0,05-peil

^④ Vir die berekening van die produk-moment-korrelasiekoeffisiënte, sien p.140.

TABEL 5.6

KORRELASIES: IK, VAK- EN TOETSSTREKASIES - VOLLE GROEP (N = 30)

	IK NV	IK V	IK Totaal	Afrikaans	Engels	Rekenkunde	Ander vakke	Vak= totaal	I	II	Bentley III	IV	Totaal
IK-NV	-												
IK-V	+673*												
IK-Totaal	908*	917*											
Afrikaans	216	298	361**										
Engels	180	422**	341	949*									
Rekenkunde	399**	423**	461**	711*	672*								
Ander vakke	194	355	319	861*	850*	703*							
Gem. vaktotaal	279	436**	407**	945*	929*	863*	929*						
Subtoets I	142	211	198	199	143	136	-018	117					
Subtoets II	411**	217	340	318	355	302	378**	375**	-121				
Subtoets III	122	118	132	040	-018	175	093	085	153	162			
Subtoets IV	232	138	194	085	028	053	124	091	130	534*	-042		
Toetstotaal	363**	272	344	199	175	291	226	252	452**	641*	540*	728*	-

+ = desimale tekens weggelaat

* = betekenisvol op 0,01-peil

** = betekenisvol op 0,05-peil

die skoolvakke deurgaans beduidend tot hoogs beduidend (0,595 tot 0,958).

- Korrelasie tussen subtoetse onderling en tussen subtoetse en toetstotaal

Slegs Subtoets II en IV korreleer betekenisvol in die geval van die dogters-groep (0,692) en volle klasgroep (0,534). Die subtoetse korreleer almal beduidend met die toetstotaal in die geval van die volle toetsgroep.

Klaarblyklik is die betrokke klasgroep te klein om werklik betekenisvolle afleidings te kan maak. Tog kan die volgende voorlopige gevolgtrekkings moontlik gemaak word:

Die Bentleytoets meet op 'n heel ander terrein as die normale skoolvakesamens en toon weinig verband met IK-tellings en vakprestasies, terwyl skoolvakke beduidend met IK korreleer asook onderling. Een moontlike afleiding wat gemaak kan word, is dat die Bentleytoets 'n meer objektiewe meetmiddel is as die normale skoolvakesamens, waar die subjektiewe element evaluering in 'n redelike mate mag beïnvloed.

Die verskillende subtoetse (met die uitsondering van Subtoets II en IV) vertoon redelik onafhanklik in funksie wat gemeet word, maar korreleer beduidend met die toetstotaal sodat daar duidelik van 'n gemeenskaplike faktor sprake is.

- Intervalgrootte as faktor in toonhoogteonderskeiding

'n Ontleding van die moeilikheidswaarde van items in Subtoets I (aanduiding of die toonhoogte styg, daal of dieselfde bly) word in tabel 5.7 aangegee. Dit kan verwag word dat groter intervalle in die algemeen makliker as kleiner intervalle te onderskei is (Bentley, 1966, p.52). Die resultate toon dan ook dat die halftoon (mineur 2de, stygend of dalend) die moeilikste vir die leerlinge was om te onderskei. Naas die unison was 'n stygende mineur 6de daarteenoor die maklikste terwyl 'n dalende oktaaf egter relatief moeilik onderskeibaar was. 'n Ontleding van die trillingsgetal as maatstaf van die eksakte toonafstand werp ook nie meer lig op die saak nie: 'n trillingsverskil

van 48 (heeltoon) is volgens die tabel maklik herkenbaar terwyl 'n verskil van 62,3 (mineur 2de) veel moeiliker geblyk het. Soos reeds genoem, is die subtoets klaarblyklik te maklik en gevoldglik nie diskriminerend genoeg nie. In die latere navorsingsprogram word egter verder hierop ingegaan (vide p. 186).

TABEL 5.7

INTERVALGROOTTE EN MOEILIKHEIDSWAARDE (SUBTOETS I)

Interval	Betrokke note	Trillings per sekonde	Verskil in trillings per sekonde	Moeilikhedswaarde (% korrek) in rangorde
Unison	$f^1 - f^1$	349,2	0	100
	{ $a^2 - a^2$	880,0	0	100
	$a - a$	220,0	0	100
Stygende mineur 6de	$g^{*1} - e^2$	415,4 - 659,3	243,9	100
Unison	{ $c^2 - c^2$	523,3	0	96,67
	$b - b$	246,9	0	96,67
Stygende heeltoon	$g^1 - a^1$	392,0 - 440,0	48	96,67
Dalende majeur 3de	$e^2 - c^2$	659,3 - 523,3	136	96,67
Dalende rein 5de	$c^2 - f^1$	523,3 - 349,2	174,1	86,67
Stygende mineur 3de	$c^{*1} - e$	277,2 - 329,7	52,5	83,33
Stygende mineur 7de	$d^1 - c^2$	293,7 - 523,3	229,6	83,33
Dalende majeur 7de	$f^{*2} - g^1$	740,0 - 392,0	348	80
Stygende mineur 2de	$c^2 - d^{*2}$	523,3 - 554,4	31,1	76,67
Stygende majeur 2de	$a^{*2} - b^{*2}$	830,7 - 932,4	97,7	73,33
Dalende majeur 2de	$g^2 - f^2$	784,0 - 698,4	85,6	70
Dalende mineur 2de	$c^2 - b^1$	523,3 - 493,9	29,4	66,67
Dalende rein oktaaf	$b^2 - b^1$	987,8 - 493,9	493,9	63,33
Stygende mineur 2de	$a - b^1$	220,0 - 233,1	13,1	60
Dalende mineur 2de	$d^{*3} - c^3$	1108,8 - 1046,5	62,3	53,33
Dalende mineur 2de	$g^1 - f^{*1}$	392,0 - 370,0	22	46,67

5.2.2 Eksperimentele groep II

5.2.2.1 Doelstellings

- Die algemene doel was om die Bentleytoets op 'n groter groep leerlinge as in die vorige eksperiment toe te pas en leerlinge van verskillende skole en ouderdomsgroepe te betrek. Verdere doelstellings was die volgende.
- Om die betrouwbaarheid van die toets te ondersoek wat betref tydstabiliteit (toets-hertoets-betrouwbaarheid) en interne konsekwentheid (bereken volgens Kuder-Richardsonformule-21).
- Die ondersoek na die verskil in toetsprestasie volgens ouderdom, standerd en geslag.
- Om die subtoetse onderling te vergelyk wat betref algemene moeilikheidswaarde asook die moontlike verband tussen die subtoetse vas te stel.
- Om deur middel van 'n itemontleding objektiewe, kwantitatiewe data te verkry met betrekking tot die toetssamestelling en die afsonderlike toets-items te evalueer.
- Om die metingskaal wat Bentley verskaf te vergelyk met die huidige resultate (teen die agtergrond van verskillende normskale).

5.2.2.2 Die ondersoekgroep

Die groep het bestaan uit 115 leerlinge van ses verskillende skole wat vrywillig aan die toetsprogram deelgeneem het. Die volgende tabel toon die samestelling van die groep volgens standerd, ouderdom en geslag.

TABEL 5.8

SAMESTELLING VAN ONDERSOEKGROEP VOLGENS STANDERD, OUDERDOM EN GESLAG

	Standerd					Ouderdom (jare)					Totaal
	2	3	4	5	6	9	10	11	12	13	
Seuns	7	9	9	3	0	6	10	9	3	0	28
Dogters	21	21	20	18	7	16	23	22	13	13	87
Totaal	28	30	29	21	7	22	33	31	16	13	115

Met die variasiebreedte en standaardafwyking as maatstaf van heterogeniteit toon die ondersoekgroep 'n heterogene verspreiding wat alle toetstellings betref (vide p.83).

TABEL 5.9

VARIASIEBREEDTE EN STANDAARDAFWYKING IN TOETSTELLINGS AS MAATSTAF VAN HETEROGENITEIT (EERSTE TOETSING, N = 115)

	\bar{X}	Min.	Maks.	Var.br.	s	$\frac{\text{var.br.}}{s}$
Subtoets I	12,113	4	19	15	3,938	3,809
Subtoets II	6,000	0	10	10	2,544	3,931
Subtoets III	8,409	0	16	16	3,683	4,344
Subtoets IV	5,757	1	10	9	2,125	4,235
Toetstotaal	32,278	9	52	43	9,214	4,667

Om vas te stel of die ondersoekgroep 'n normale verspreiding verteenwoordig, is die Chi-kwadraattoets*) vir normaliteit op die totale toetstellings (eerste toetsing) toegepas. Hiervolgens is die verspreiding wel normaal:

$$\chi^2 = 5,935 \text{ (Kritiese waardes van } \chi^2 : \chi^2 \geq 11,071 \text{ (0,05-peil)} \\ \chi^2 \geq 15,086 \text{ (0,01-peil))})$$

5.2.2.3 Die instrument (vide p.83 hierbo en p.138 hieronder)

By die vorige eksperiment is van 'n klavier gebruik gemaak en toetsinstruksies mondeling gegee. By hierdie toepassing is van die oorspronklike klankbaan gebruik gemaak met dié verskil dat die oorspronklike Engelse aanwysings vervang is met Afrikaanse instruksies (vergelyk hieronder, p.146). Daar is naamlik gepostuleer dat die taalfaktor 'n invloed op die leerlinge se toetsprestasie kon uitoefen.

*) vide p.143.

5.2.2.4 Metodologiese oorwegings

Vir die bepaling van toets-hertoets-betrouwbaarheid is 21 van die 115 leerlinge twee dae later weer getoets, veral om die moontlike invloed van vertroudeheid met die toetsinhoud te probeer bepaal. Verder is 46 leerlinge twee maande later, en 48 leerlinge vier maande later hertoets. Verskillende tydsintervalle is naamlik gebruik vir vergelykingsdoeleindes.

Op grond van die ondervinding opgedoen met die eerste toetsing, is enkele wysisings in die klankbaan aangebring en dié gewysigde klankbaan vir die herhalings van die toets gebruik. So is 'Op' en 'Af' byvoorbeeld vervang met 'Hoër' en 'Laer', wat meer in ooreenstemming is met die bestaande skoolmusiekterminologie in Suid-Afrika.

Verder is die antwoordblad ietwat gewysig, sodat die leerling elke respons met 'n kruisie moes aandui, in plaas van telkens 'n lettersimbool of syfer neer te skryf (vide p.146).

5.2.2.5 Resultate van die ondersoek

• Vergelyking tussen resultate van verskillende toetstoepassings

Uit tabel 5.10 blyk die volgende:

Hertoetsing na twee dae.

Subtoets III (akkoordanalise) toon die grootste verbetering in gemiddelde toetstelling by herhaling (5,24%) terwyl Subtoets I (toonhoogteonderskeiding) 'n verswakking van 5,24% gelewer het. In totaal is die verbetering gering (0,6%). Slegs by Subtoets I is daar 'n betekenisvolle verskil tussen die gemiddelde tellings.

Toetsbetrouwbaarheid, volgens die KR-formule-^{21*)} bereken, is onbevredigend (0,5 as minimum geneem ^{**) terwyl toets-hertoets-betrouwbaarheid slegs in die geval van Subtoets I en die toetstotaal betekenisvol is.}

*) vide p.142.

**) vide Helmstadter, 1964, p.84.

TABEL 5.10

VERGELYKING VAN GEMIDDELDE TOETSTELLINGS BY HERTOETSING

5.10.1 HERTOETSING NA TWEE DAE (N = 21)

Subtoets	n	Eerste toetsing			Tweede toetsing			r_{pm}	$\bar{X}_1 - \bar{X}_2$	z	Gemiddelde % verbetering
		\bar{X}_1	s_1	KR-21	\bar{X}_2	s_2	KR-21				
I	20	15,286	1,955	+060	14,238	2,844	+519	+579*	1,048	1,392	-5,240
II	10	7,762	1,231	-163	8,095	1,269	047	380	-0,333	-0,863	3,330
III	20	10,952	2,609	286	12,000	2,488	236	345	-1,048	-1,332	5,240
IV	10	6,905	1,601	185	6,952	1,647	243	378	-0,047	-0,049	0,470
Totaal	60	40,905	4,116	236	41,286	5,138	521	672*	-0,381	-0,265	0,635

5.10.2 HERTOETSING NA TWEE MAANDE (N = 46)

I	20	10,283	4,116	742	11,870	4,387	789	709*	-1,857	-2,094**	7,935
II	10	5,239	2,495	666	5,870	2,173	541	459*	-0,631	-1,293	6,310
III	20	7,130	3,949	743	9,500	3,519	629	582*	-2,370	-3,039*	11,850
IV	10	5,152	2,064	460	6,043	2,186	555	673*	-0,891	-2,010**	8,910
Totaal	60	27,804	9,062	832	33,283	9,252	841	774*	-5,479	-2,869*	9,132

5.10.3 HERTOETSING NA VIER MAANDE (N = 48)

I	20	12,479	3,379	620	13,125	4,003	756	655*	-0,646	-0,854	3,230
II	10	5,958	2,638	729	6,292	2,458	682	748*	-0,334	-0,642	3,340
III	20	8,521	3,188	546	10,063	3,132	516	374*	-1,542	-2,390**	7,710
IV	10	5,833	2,163	534	7,083	1,880	462	557*	-1,250	-3,022*	12,500
Totaal	60	32,792	8,080	785	36,563	8,488	815	757*	-3,771	-2,229**	6,285

5.10.4 HERTOETSING - VOLLE GROEP (N = 115)

I	20	12,113	3,938	728	13,122	4,171	780	701*	-1,009	-1,886	5,045
II	10	6,000	2,544	699	6,461	2,323	640	652*	-0,461	-1,435	4,610
III	20	8,409	3,683	674	10,070	3,248	554	434*	-1,661	-3,627*	8,305
IV	10	5,757	2,125	510	6,800	2,164	595	597*	-1,043	-3,688	10,430
Totaal	60	32,278	9,214	838	36,452	9,012	838	788*	-4,174	-3,473*	6,957

+ = desimale tekens weggelaat * = betekenisvol op 0,01-peil r_{pm} = toets-hertoets-betrouwbaarheid

Hertoetsing na twee maande

Subtoets III toon weer eens die grootste verbetering in gemiddelde toetstelling by herhaling (11,85%). Met die uitsondering van Subtoets II (melodiese geheue) is die verskille tussen die gemiddeldes almal beduidend.

Behalwe by Subtoets IV (ritmiese geheue), eerste toetsing, is al die betrouwbaarheidsyfers bevredigend terwyl die toets-hertoets-betrouwbaarheid deurgaans beduidend is.

Hertoetsing na vier maande

Subtoets IV toon die grootste verbetering in gemiddelde toetstelling. Die verskil tussen gemiddeldes is in die geval van Subtoets III en IV en die toets-totaal beduidend.

Behalwe by Subtoets IV, tweede toetsing, is al die betrouwbaarheidsyfers (KR-21) bevredigend en toets-hertoets-betrouwbaarheid in alle gevalle beduidend.

Hertoetsing - volle ondersoekgroep

Subtoets IV toon die grootste verbetering, gevvolg deur Subtoets III. Die verskil tussen gemiddeldes is in die geval van Subtoets III en IV en die toets-totaal beduidend.

Betrouwbaarheid volgens KR-21 is deurgaans bevredigend en die toets-hertoets-betrouwbaarheidsindeks in alle gevalle beduidend.

Gevolgtrekkings

Die verbeterde toetsprestasie met die herhaling, kan aan verskeie moontlike faktore toegeskryf word:

Veral by die jonger kind is die vreemdheid van die toetssituasie en toetsinhoud 'n faktor wat tot 'n minder getroue weergawe van die leerling se prestasievermoë die eerste keer kon lei.

Die gebruik van 'n verbeterde Afrikaanse klankbaan die tweede keer, mag 'n bydrae gelewer het, veral wat die algemene begrip van die toetsvereistes betref.

TABEL 5.11

TOETSPRESTASIES VAN SEUNS TEENOOR DOGTERS (N = 115)

Sub-toets	n	\bar{X}_1	s_1	KR-21	\bar{X}_2	s_2	KR-21	$\bar{X}_1 - \bar{X}_2$	z	% verskil
SEUNS (N = 28)										
DOGTERS (N = 87)										
I	20	11,889	4,200	+765	12,140	3,926	+727	-0,251	-0,276	1,255
II	10	5,630	2,356	619	6,081	2,638	731	-0,451	-0,845	4,510
III	20	7,852	3,830	710	8,581	3,705	677	-0,729	-0,871	3,645
IV	10	5,852	1,791	270	5,698	2,234	565	0,154	0,367	1,540
Totaal	60	31,222	9,061	831	32,500	9,413	846	-1,278	-0,634	2,130
TWEEDE TOETSING										
I	20	12,000	4,587	812	13,483	4,014	766	-1,483	-1,510	7,415
II	10	5,929	2,193	553	6,632	2,363	667	-0,703	-1,429	7,030
III	20	8,714	3,065	502	10,506	3,216	545	-1,792	-2,624*	8,960
IV	10	6,786	2,500	723	6,805	2,073	549	-0,019	-0,036	0,190
Totaal	60	33,428	9,311	843	37,425	8,798	832	-3,997	-1,974**	6,662
SEUNS EERSTE TOETSING										
SEUNS TWEEDE TOETSING (N = 28)										
I	20	11,889	4,200	765	12,000	4,587	812	-0,111	-0,093	0,555
II	10	5,630	2,356	619	5,929	2,193	553	-0,299	-0,483	2,990
III	20	7,852	3,830	710	8,714	3,065	502	-0,862	-0,913	4,310
IV	10	5,852	1,791	270	6,786	2,500	723	-0,934	-1,578	9,340
Totaal	60	31,222	9,061	831	33,428	9,311	843	-2,206	-0,882	3,677
DOGTERS EERSTE TOETSING										
DOGTERS TWEEDE TOETSING (N = 87)										
I	20	12,140	3,926	727	13,483	4,014	766	-1,343	-2,231**	6,715
II	10	6,081	2,638	731	6,632	2,363	667	-0,551	-1,450	5,510
III	20	8,581	3,705	677	10,506	3,216	545	-1,925	-3,660*	9,625
IV	10	5,698	2,234	565	6,805	2,073	549	-1,107	-3,385*	11,070
Totaal	60	32,500	9,413	846	37,425	8,798	832	-4,925	-3,566*	8,208

+ = desimale tekens weggelaat

* = betekenisvol op 0,01-peil

** = betekenisvol op 0,05-peil

Dit is te betwyfel of die leerlinge kon gebaat het by 'n kennis van, of vertroudheid met, die betrokke toetsitems. Dit blyk onder ander uit die geringe verbetering in toetsprestasie by leerlinge wat na twee dae hertoets is: 0,6% in totaal terwyl Subtoets I 'n verswakking van 5,2% getoon het.

Die gewysigde antwoordblad mag tot 'n noukeuriger weergawe van antwoorde en akkurater nasien van die toets geleei het.

In die algemeen vertoon die toets bevredigende interne konsekwentheid (volgens KR-21) en tydstabiliteit (volgens toets-hertoets-korrelasiesyfers). Soos genoem (p.57), word eersgenoemde beïnvloed deur die homogeniteit van toetsitems. Van die vier subtoetse vertoon Subtoets I die grootste en Subtoets III die laagste homogeniteit tussen items.

- Vergelyking van toetsprestasie volgens geslag

Uit 'n vergelyking tussen die gemiddelde toetsprestasies van seuns en dogters (Tabel 5.11), blyk die volgende:

Met die uitsondering van Subtoets IV, eerste toetsing, is die dogters se toetsprestasie deurgaans beter as dié van die seuns. Hierdie verskil is egter slegs betekenisvol in die geval van Subtoets III (tweede toetsing) en die toetstotaal (tweede toetsing).

Die verskil tussen seuns en dogters se gemiddelde prestasie is, met die uitsondering van Subtoets IV, aansienlik groter tydens die hertoetsing as met die eerste toetsing, bv. 6,7% teenoor 2,1% wat die toetstotaal betref.

Die verskil in gemiddelde prestasie tussen eerste en tweede toetsing by seuns is deurgaans onbeduidend, terwyl dit by die dogters, met die uitsondering van Subtoets II, in alle gevalle beduidend is.

By sowel seuns as dogters, word die grootste verbetering by herhaling by Subtoets IV aangetref. By seuns is die kleinste verbetering by Subtoets I (0,6%) en by dogters by Subtoets II (5,5%).

TABEL 5.12

VERGELYKING VAN TOETSPRESTASIES VOLGENS OUDERDOM (N = 115)

Sub-toets	n	EERSTE TOETSING			TWEEDE TOETSING			r_{pm}	$\bar{X}_1 - \bar{X}_2$	z	Gemiddelde % verbetering
		\bar{X}_1	s_1	KR-21	X ₂	s_2	KR-21				
9 jaar (N = 22)											
I	20	10,917	3,525	633	12,182	4,272	778	+529**	-1,265	-1,047	6,325
II	10	5,667	2,599	707	6,045	2,535	698	514**	-0,378	-0,477	3,780
III	20	7,417	4,032	750	9,273	4,222	759	831*	-1,856	-1,457	9,280
IV	10	5,333	2,057	458	6,091	1,716	213	324	-0,758	-1,298	7,580
Totaal	60	29,333	8,535	808	33,591	9,364	845	759*	-4,258	-1,541	7,091
10 jaar (N = 33)											
I	20	11,514	4,053	740	12,424	4,338	789	825*	-0,910	-0,880	4,550
II	10	5,811	2,209	557	6,242	2,194	570	600*	-0,431	-0,795	4,310
III	20	7,649	3,474	641	9,758	3,269	560	269	-2,109	-2,538**	10,545
IV	10	5,243	2,114	491	5,939	2,318	612	643*	-0,696	-1,321	6,960
Totaal	60	30,216	8,879	823	34,364	8,813	825	770*	-4,148	-1,904	6,913
11 jaar (N = 31)											
I	20	12,310	4,150	763	13,258	4,379	809	729*	-0,948	-0,873	4,740
II	10	5,517	2,721	740	6,645	2,184	592	682*	-1,128	-1,799	11,280
III	20	8,000	3,474	634	10,258	2,756	360	369**	-2,258	-2,837*	11,290
IV	10	6,138	2,183	558	7,387	1,995	572	592*	-1,249	-2,352**	12,490
Totaal	60	31,965	9,179	837	37,548	8,469	818	758*	-5,583	-2,489**	9,305
12 jaar (N = 16)											
I	20	13,625	3,631	706	13,937	3,941	766	534**	-0,312	-0,225	1,560
II	10	7,500	2,708	827	6,812	2,762	795	687*	0,688	0,689	-6,880
III	20	11,312	2,960	462	10,062	2,977	459	124	1,250	1,153	-6,250
IV	10	6,375	1,893	395	7,375	2,306	707	637*	-1,000	-1,299	10,000
Totaal	60	38,812	8,408	820	38,187	9,404	857	729*	0,625	0,192	-1,042
13 jaar (N = 13)											
I	20	14,375	3,543	714	15,154	2,911	597	473	-0,779	-0,588	3,895
II	10	6,625	2,560	732	6,846	2,304	659	852*	-0,221	-0,222	2,210
III	20	10,500	2,976	460	11,769	2,555	272	522	-1,269	-1,121	6,345
IV	10	6,750	2,435	700	8,077	1,656	482	481	-1,327	-1,561	13,270
Totaal	60	38,250	8,276	811	41,846	7,872	809	921*	-3,596	-1,091	5,993

+ = desimale tekens weggelaat

* = betekenisvol op 0,01-peil

** betekenisvol op 0,05-peil

Toetsbetroubaarheid, volgens die KR-21 formule bereken, is bevredigend behalwe in die geval van Subtoets IV, eerste toetsing, by seuns. In die meeste gevalle toon Subtoets I die beste betroubaarheidsyfer van al die subtoetse, terwyl die toetstotaal deurgaans 'n hoogs bevredigende syfer (bokant 0,8) lewer.

- Vergelyking van toetsprestasie volgens ouderdom

Tabel 5.12 gee 'n vergelyking tussen die toetsprestasies van verskillende ouderdomsgroepe.

Toetsprestasie styg met styging in ouerdom. Hierdie styging is betreklik reëlmatig wat betref Subtoets I, terwyl daar 'n aansienlike styging gevind word vanaf elf- na twaalfjariges en 'n daling vanaf twaalf na dertien jaar wat die toetstotaal, Subtoets II en III betref.

Van al die ouerdomsgroepe toon elfjariges in die meeste gevalle by hertoetsing die grootste verbetering in toetsprestasie, terwyl twaalfjariges 'n verswakkings ten opsigte van Subtoets II en III en die toetstotaal toon (6,9% in die geval van Subtoets II). Die verskil tussen gemiddelde toetsprestasie van eerste teenoor tweede toetsing is slegs betekenisvol by Subtoets III in die geval van tienjariges en Subtoets III, IV en totaal by elfjariges. Van die verskillende subtoetse toon Subtoets IV by nege- en tienjariges en Subtoets IV by elf-, twaalf- en dertienjariges die grootste gemiddelde persentasieverbering.

Toetsbetroubaarheid (KR-21) is deurgaans bevredigend (bokant 0,5) wat Subtoets I en II en die toetstotaal betref, terwyl Subtoets III en IV se betroubaarheidsyfer in die meeste gevalle by elf-, twaalf- en dertienjariges onbevredigend is.

Die toets-hertoetsbetroubaarheidsyfer is in die meeste gevalle betekenisvol. Die uitsonderings is Subtoets III by tien-, twaalf- en dertienjariges en Subtoets IV by nege- en dertienjariges.

TABEL 5.13

VERGELYKING VAN TOETSPRESTASIES VOLGENS STANDERD (N = 115)

Sub-toets	n	\bar{X}_1	EERSTE TOETSING s_2	KR-21	\bar{X}_2	TWEEDE TOETSING s_2	KR-21	r_{pm}	$\bar{X}_1 - \bar{X}_2$	z	Gemiddelde % verbetering
Standerd 2 (N = 28)											
I	20	11,107	3,573	+645	12,214	3,947	+731	+565*	-1,133	-1,105	5,535
II	10	5,643	2,407	640	5,929	2,433	658	520*	-0,286	-0,435	2,860
III	20	7,571	4,032	748	9,536	4,041	731	651*	-1,965	-1,788	9,825
IV	10	5,250	1,974	400	6,000	1,764	254	436**	-0,750	-1,473	7,500
Totaal	60	29,571	8,399	801	33,678	8,415	805	712*	-4,107	-1,796	6,845
Standerd 3 (N = 30)											
I	20	11,333	4,310	774	12,067	4,608	815	824*	-0,734	-0,637	3,670
II	10	5,700	2,351	618	6,400	2,111	537	574*	-0,700	-1,125	7,000
III	20	7,467	3,501	651	9,567	3,277	563	367**	-2,100	-2,400**	10,500
IV	10	5,333	2,218	549	6,033	2,371	638	673*	-0,700	-1,180	7,000
Totaal	60	29,833	9,296	858	34,067	9,465	850	811*	-4,234	-1,748	7,057
Standerd 4 (N = 29)											
I	20	12,172	3,809	707	14,276	3,693	737	500*	-2,104	-2,098**	10,520
II	10	5,724	2,604	710	6,690	2,173	590	723*	-0,966	-1,507	9,660
III	20	8,276	2,825	425	10,103	2,833	397	367**	-1,827	-2,407**	9,135
IV	10	6,103	2,144	536	7,897	1,633	419	493*	-1,794	-3,525*	17,940
Totaal	60	32,276	7,919	775	38,965	7,159	746	708*	-6,689	-3,315*	11,148
Standerd 5 (N = 21)											
I	20	13,190	3,723	712	13,714	4,573	836	652*	-0,524	-0,397	2,620
II	10	7,048	2,854	827	6,762	2,879	818	762*	0,286	0,315	-2,860
III	20	10,234	4,024	728	10,857	2,575	265	332	-0,623	-0,583	3,115
IV	10	6,190	1,861	354	7,143	2,435	729	722*	-0,953	-0,843	9,530
Totaal	60	36,667	9,941	870	38,476	10,750	895	872*	-1,809	-0,553	3,015
Standerd 6 (N = 7)											
I	20	16,000	2,449	491	14,714	2,690	487	051	1,286	0,866	-6,430
II	10	6,714	2,563	738	7,000	1,826	411	890*	-0,286	-0,173	2,860
III	20	10,857	3,024	481	11,857	2,911	453	262	-1,000	-0,583	5,000
IV	10	6,857	2,734	791	7,714	1,799	506	532	-0,857	-0,471	8,570
Totaal	60	40,429	7,807	797	41,286	6,601	716	841**	-0,857	-0,205	1,428

+ = desimale tekens weggelaat

* = betekenisvol op 0,01-peil

** = betekenisvol op 0,05-peil

- Vergelyking van toetsprestasie volgens standerd

Tabel 5.13 gee 'n vergelyking tussen die toetsprestasies van verskillende standerdgroepe.

Daar is weinig verskil in toetsprestasie tussen standerd 2- en 3-leerlinge. Vanaf standerd 3 is daar egter 'n betreklik reëlmataige styging volgens styging in standerd. Prestasies in Subtoets IV, tweede toetsing, bereik 'n hoogtepunt by standerd 4-leerlinge en verswak dan by standerd 5.

Behalwe by Subtoets III toon standerd 4-leerlinge die grootste verbetering by herhaling. Die grootste enkele verbeterings by herhaling word gevind by standerd 4-leerlinge wat Subtoets IV en die toetstotaal betref (verskil tussen gemiddeldes betekenisvol op 0,01-peil). Die verskil in gemiddelde toetsprestasie by eerste en tweede toetsing is slegs betekenisvol by standerd 3-leerlinge, Subtoets III en standerd 4-leerlinge, Subtoetse I, III, IV en die totaal.

Met die uitsondering van standerd 6-leerlinge, is die toetsbetroubaarheid van Subtoetse I, II en die toetstotaal deurgaans bevredigend terwyl toets-her-toets-betroubaarheid, met die uitsondering van Subtoets III (standerd 5) en Subtoetse I, III en IV by standerd 6-leerlinge deurgaans betekenisvol is.

- Vergelyking van subtoetse onderling

Die volgende tabel vergelyk die gemiddelde toetsprestasie (uitgedruk as %) in die verskillende subtoetse en vergelyk dié resultate met Bentley (1966a, p.107) se bevindinge.

Toetsprestasie in Subtoets III is in alle gevalle opmerklik laer as in die ander subtoetse, terwyl Subtoets IV by die ondersoekgroep, tweede toetsing, en by Bentley se eie resultate, die hoogste gemiddelde prestasie aantoon. Op hierdie aspek word weer in die latere navorsingsprogram teruggekom (vide p. 191, 192).

TABEL 5.14

VERGELYKING TUSSEN SUBTOETSGEMIDDELDES, ONDERSOEKGROEP/BENTLEY

	N	EERSTE TOETSING				TWEDE TOETSING			
		I	II	III	IV	I	II	III	IV
		%	%	%	%	%	%	%	%
huidige ondersoek	115	60,565	60	42,045	57,570	65,61	64,61	50,35	68,00
Bentley	ca. 2000	53	56	37	64				

Die volgende interkorrelasietabel (5.15) gee 'n aanduiding van die mate waarin die subtoetse onderling korreleer asook die verband tussen elke subtoets en die toetstotaal.

Behalwe in die geval van die groep leerlinge wat twee dae later hertoets is, is die korrelasiesyfers feitlik deurgaans betekenisvol sodat dit wil voorkom of daar 'n redelike mate van oorvleueling tussen die subtoetse is en hul dus nie heeltemal onafhanklik is in funksie wat gemeet word nie. In die geval van die volle groep, varieer hierdie korrelasies vanaf 0,240 tot 0,563 - 'n gemeenskaplike variansie dus van tussen 5,8% en 31,7%. Die twee subtoetse wat in die meeste gevalle die laagste met mekaar korreleer, is Subtoets III en IV (akkoordanalise en ritmiese geheue).

Die feit dat elke subtoets hoog korreleer met die toetstotaal, gee 'n verdere aanduiding van die interafhanklikheid tussen die subtoetse. Die subtoetse wat in die meeste gevalle die laagste met die toetstotaal korreleer en dus die 'onafhanklikste' voorkom, is Subtoets IV (ritmiese geheue).

- Itemontleding

Algemene rasionaal

Die kwaliteit en waarde van 'n toets hang uiteindelik af van die meriete van die individuele items waaruit die toets saamgestel is. Die algemene doel met itemontleding is dus die evaluering van toetsitems in die lig van die oogmerk en rasional met die toets as metingsmiddel (Freeman, 1955, p.38). Itemontleding voorsien data met betrekking tot die effektiwiteit van die individue.

TABEL 5.15

INTERKORRELASIES TUSSEN SUBTOETSE EN TOETSTOTAAL

Hertoetsing (N = 115)						Hertoetsing na 2 dae (N = 21)				
	I	II	III	IV	Totaal	I	II	III	IV	Totaal
I	-	+563*	365*	473*	853*	-	310	034	145	693*
II	451*	-	403*	439*	769*	187	-	166	093	529**
III	335*	387*	-	240**	691*	-016	-167	-	279	633*
IV	286*	478*	287*	-	659*	130	399	022	-	559*
Totaal	752*	746*	714*	590*	-	571*	437**	585*	584*	-
Eerste toetsing						Eerste toetsing				
Hertoetsing na 2 maande (N = 46)						Hertoetsing na 4 maande (N = 48)				
I	-	555*	279	542*	839*	-	466*	411*	381*	843*
II	351**	-	458*	463*	781*	407*	-	325**	428*	724*
III	259	330**	-	226	673*	331**	406*	-	208	703*
IV	328**	316**	371**	-	688*	176	444*	233	-	602*
Totaal	738*	697*	729*	682*	-	729*	776*	728*	578*	-
Eerste toetsing						Eerste toetsing				

103

+ = desimale tekens weggelaat

* = betekenisvol op 0,01-peil

** = betekenisvol op 0,05-peil

duele items en help om items wat gebrekbaar is in een of meer opsigte te identifiseer. Vir die huidige doel gaan dit hier om die diagnose van defektiewe items en die verdere analise van die iteminhoud met die doel om moontlike oorsake vir die minder geslaagde funksionering van die item te soek.

Basies behels itemontleding twee hoofasette, naamlik die vasstelling van die moeilikheidswaarde en diskriminasievermoë van elke item. Laasgenoemde bied daarby ook 'n geldigheidsindeks deurdat dit aandui hoedanig die item meet of diskrimineer in ooreenstemming met die res van die toets (Guilford, 1954, p.417) – dit wil sê, inhoudsgeldigheid (vide p.56 hierbo). Itemontleding kan verder 'n aanduiding gee van die betroubaarheid van elke item asook hoe doeltreffend elke afleier funksioneer.

Moeilikheidswaarde (p)

Dit is die verhouding van die toetslinge wat 'n item korrek beantwoord het tot die aantal wat die item probeer het. Teoreties word dit meestal aanvaar dat die ideale verhouding-korrekt-waarde van 'n item 0,5 behoort te wees, dit wil sê dat dit deur 50% van die leerlinge korrek beantwoord behoort te word (Gullikson, 1950, p.374; Garrett, 1953, p.350); ander skrywers stel dit wyer, byvoorbeeld tussen 50% en 65% van die middel van die groep wat getoets word (Lindvall & Nitko, 1975, p.137). Ten einde voldoende variasiebreedte en dus diskriminasie oor die volle meetbestek te verkry, behoort items egter gebalanseerd te wees en te varieer van baie maklik tot baie moeilik, maar met 'n gemiddelde moeilikheidswaarde van ongeveer 0,5. Op hierdie peil is die diskriminasie tussen toetslinge op sy maksimum sodat die maksimum getal individuele verskille na vore gebring word.

'n Belangrike praktiese aspek wat hieruit voortvloeи is die volgorde waarin die individuele toetsitems aangebied word. Dit mag naamlik 'n invloed uitoefen op die toetsresultate veral deurdat dit verband hou met motivering, volgehoue belangstelling en algemene selfvertroue by die leerling. Veral wat die jonger leerlinge betref is dit byvoorbeeld wenslik dat met redelike 'maklike' items begin word – items wat hoogs waarskynlik deur feitlik al die leerlinge korrek beantwoord kan word – en geleidelik moeiliker items ingevoeg word. Dit kan so die jong leerling aanmoedig en selfvertroue bevorder.

Daar word dan gewoonlik afgesluit met enkele items waarvan dit twyfelagtig is of meer as net enkele toetslinge dit korrek sal kan beantwoord (Du Bois, 1965, p.407; Downie & Heath, 1965, p.255).

Uit Tabel 5.16 en Figure 5.1, 5.2, 5.3 en 5.4 is die algemene dalende tendens in die moeilikheidswaarde-indeks, dit wil sê 'n styging in die moeilikheidsgraad, dan ook duidelik. In die geval van Subtoets III is hierdie algemene dalende patroon nie so merkbaar as in die geval van die ander subtoetse nie.

Dit is selde moontlik om toetsitems presies aan die genoemde vereistes ('n gemiddelde moeilikheidswaarde van 0,5 en sentrering om hierdie waarde) te laat beantwoord. Die grense word dan ook dikwels so wyd soos tussen 0,2 en 0,8 gestel (Alberts, 1967, p.66; Colwell, 1969, p.70; Aiken, 1971, p.50) en die mening uitgespreek dat 'n toets wat bestaan uit items van 'n wye moeilikheidsgraadspektrum, moontlik psigologies selfs 'n beter meetinstrument sal wees (Garrett, loc. cit.). Tog kan aanvaar word dat toetsitems wat op die uiteindes van die skaal lê, dit wil sê deur feitlik almal of deur feitlik geen leerling korrek beantwoord kan word nie, nie veel diskriminasiewaarde het en dus nie van veel waarde gereken kan word nie. Items waarvan die moeilikheidswaarde nie binne die bestek van ongeveer 0,3 tot 0,7 val nie, verminder dan ook die diskriminasiewaarde van die toets en behoort ondersoek te word (vide Ahmann & Glock, 1971, p.191; Henrysson in Thorndike, 1971, p.144 e.v.). Daar behoort dus veel meer moeilikheidswaardes rondom die middelwaarde as by die uiteindes te wees, tensy die toetsinhoud baie homogeen (hoë iteminterkorrelasies) en die toetsgroep baie heterogeen is, in welke geval 'n wyer verspreiding wenslik is (Tinkelman in Thorndike, op. cit., p.64). Tabel 5.18 dien as illustrasie en verdere toeligting.

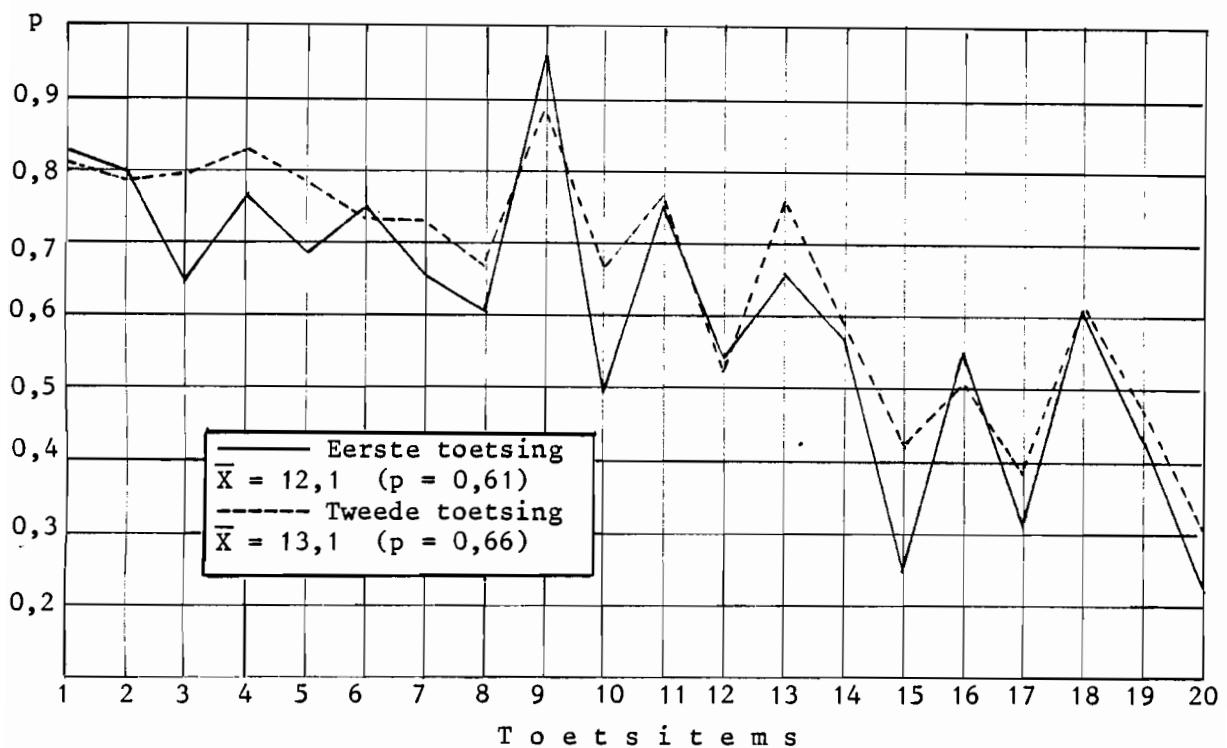
TABEL 5.16
MOEILIKHEIDSWAARDE VAN TOETSITEMS (N = 115)

Item	Subtoets I				Item	Subtoets III				
	Eerste toetsing		Tweede toetsing			Eerste toetsing		Tweede toetsing		
	Aantal korrek	% korrek	Aantal korrek	% korrek		Aantal korrek	% korrek	Aantal korrek	% korrek	
1	95	82,61	93	80,87	1	59	51,30	90	78,26	
2	92	80,00	91	79,13	2	23	20,00	14	12,17	
3	75	65,22	92	80,00	3	40	34,78	42	36,52	
4	88	76,52	96	83,48	4	63	54,78	67	58,26	
5	79	68,70	91	79,13	5	40	34,78	40	34,78	
6	87	75,65	85	73,91	6	34	29,57	47	40,87	
7	76	66,09	84	73,04	7	39	33,91	53	46,09	
8	70	60,87	78	67,83	8	66	57,39	72	62,61	
9	111	96,52	103	89,57	9	71	61,74	89	77,39	
10	57	49,57	77	66,96	10	50	43,48	69	60,00	
11	87	75,65	89	77,39	11	64	55,65	66	57,39	
12	58	50,44	60	52,17	12	66	57,39	77	66,96	
13	76	66,09	88	76,52	13	48	41,74	47	40,87	
14	66	57,39	68	59,13	14	44	38,26	57	49,57	
15	29	25,22	49	42,61	15	60	52,17	71	61,74	
16	64	55,65	59	51,30	16	52	45,22	72	62,61	
17	36	31,30	44	38,26	17	38	33,04	63	54,78	
18	71	61,74	72	62,61	18	28	24,35	23	20,00	
19	50	43,48	54	46,96	19	34	29,57	50	43,48	
20	26	22,61	36	31,30	20	49	42,61	49	42,61	
 Subtoets II										
1	68	59,13	98	85,22	1	106	92,17	99	86,09	
2	88	76,52	98	85,22	2	68	59,13	81	70,43	
3	87	75,65	78	67,83	3	68	59,13	77	66,96	
4	60	52,17	69	60,00	4	73	63,48	88	76,52	
5	80	69,57	63	54,78	5	66	57,39	85	73,91	
6	79	68,70	100	86,96	6	64	55,65	83	72,17	
7	78	67,83	84	73,04	7	54	46,96	70	60,87	
8	75	65,22	71	61,74	8	73	63,48	80	69,57	
9	46	40,00	59	51,30	9	64	55,65	83	72,17	
10	30	26,09	23	20,00	10	26	22,61	36	31,30	

Soos aangetoon in Tabel 5.17 en 5.18 word 'n wye verspreiding van moeilikhedswaardes in elke subtoets aangetref by hierdie ondersoeksgroep:

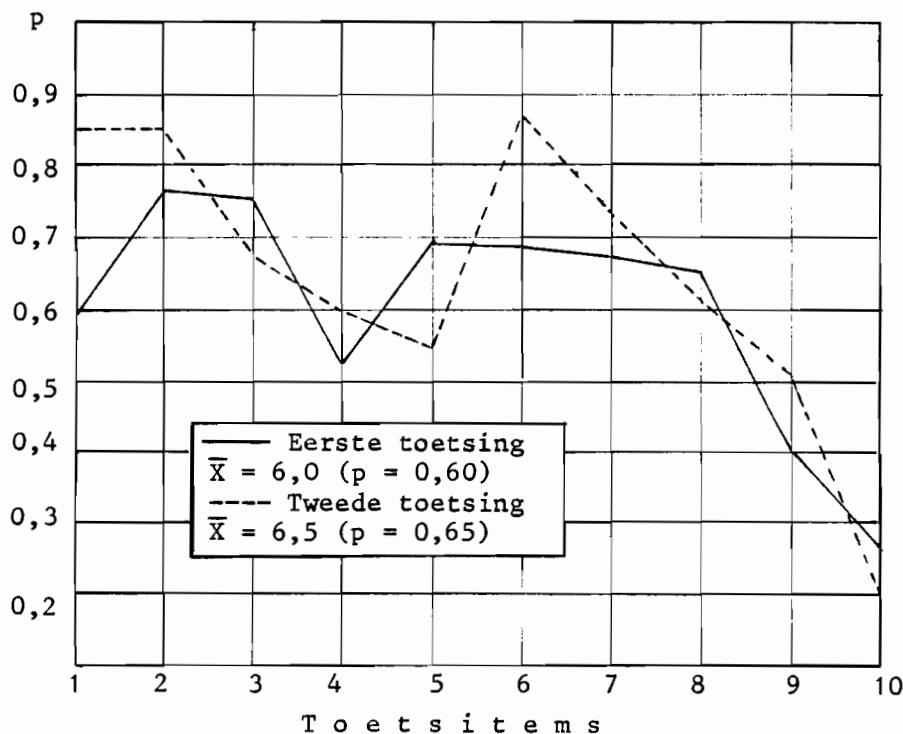
FIGUUR 5.1

BENTLEY-SUBTOETS I - MOEILIKHEIDSWAARDE (VERHOUDING KORREK) (N = 115)



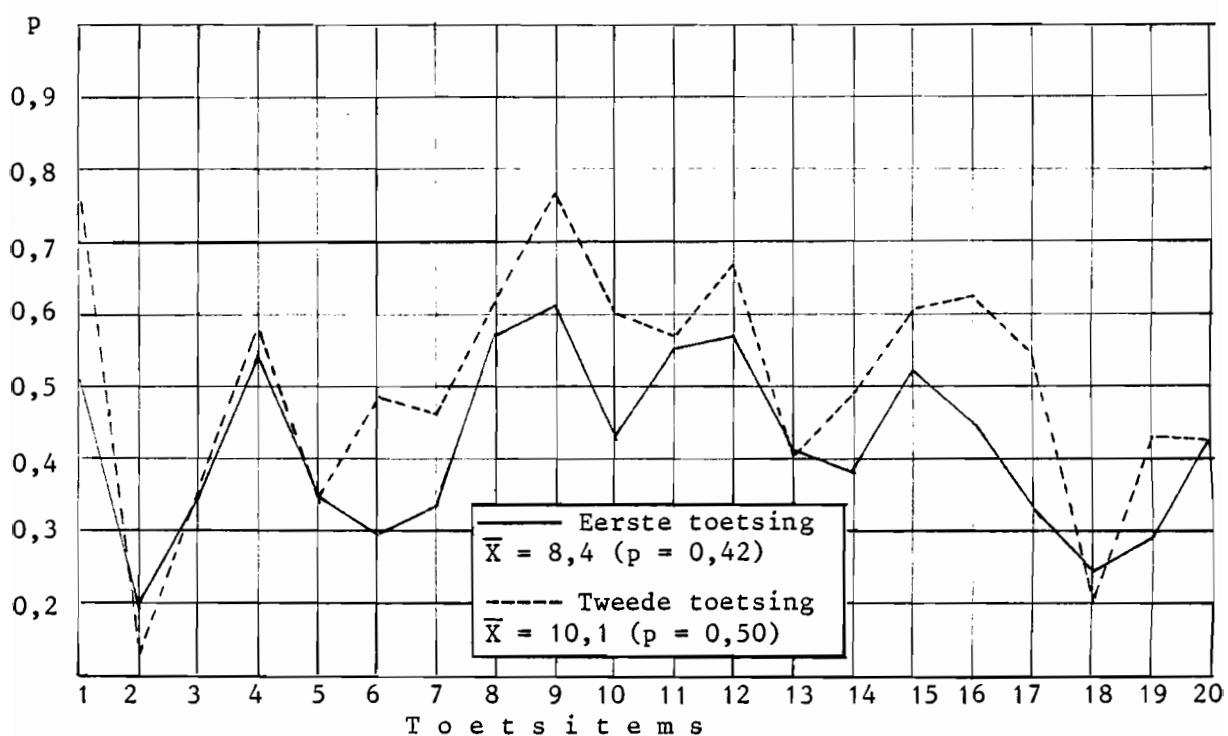
FIGUUR 5.2

BENTLEY-SUBTOETS II - MOEILIKHEIDSWAARDE (VERHOUDING KORREK) (N = 115)



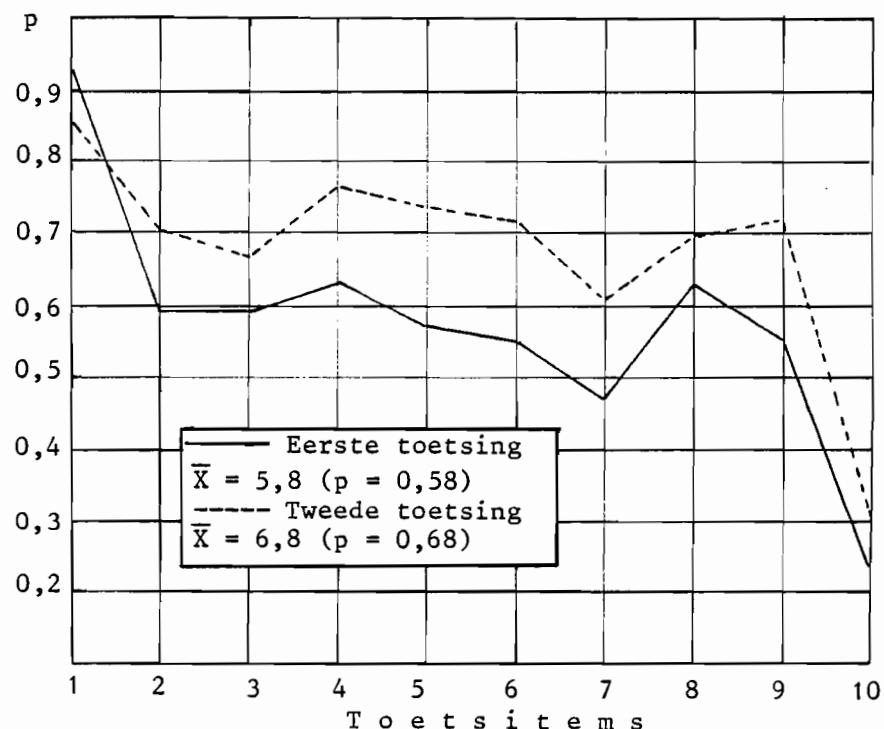
FIGUUR 5.3

BENTLEY-SUBTOETS III - MOEILIKHEIDSWAARDE (VERHOUDING KORREK) (N = 115)



FIGUUR 5.4

BENTLEY-SUBTOETS IV - MOEILIKHEIDSWAARDE (VERHOUDING KORREK) (N = 115)



TABEL 5.17

VERSPREIDING VAN MOEILIKHEIDSWAARDES: MINIMUM EN MAKSIMUM

	Eerste toetsing		Tweede toetsing	
	Minimum	Maksimum	Minimum	Maksimum
Subtoets I	0,23	0,97	0,31	0,90
II	0,26	0,77	0,20	0,87
III	0,20	0,62	0,12	0,78
IV	0,23	0,63	0,31	0,86

Enkele items is dan ook duidelik of baie maklik (bv. item 9, Subtoets I, item 1, Subtoets IV) of baie moeilik (byvoorbeeld item 2, Subtoets III) sodat die diskriminasiewaarde van hierdie items bevraagteken kan word.

Soms word die 'ideale' moeilikheidswaarde met die oog op die beste diskriminasiewaarde-indeks aan die hand van 'n model uitgewerk. 'n Voorbeeld hiervan word in die literatuur gevind (vide Nunnally, 1967, p.254), waar die volgende ongekorrigeerde p-waardes die mees diskriminerende items sou aandui:

Tweekeusige item: 0,85; driekeusig: 0,77; vierkeusig: 0,74; vyfkeusig: 0,69. Gekorrigeer volgens die formule $K_1 = K - \frac{F}{K-1}$, (waar K_1 = gekorrigeerde telling, K = aantal korrekte antwoorde, F = aantal foutiewe antwoorde en k = aantal moontlike keuses) sou hierdie waardes soos volg gewysig moet word: tweekeusig: 0,70; driekeusig: 65,5; vierkeusig: 65,3 en vyfkeusig: 61,3.

Ongekwalificeerd kan hierdie syfers egter nie sonder meer as kriterium geneem word nie. Geen korreksie van moeilikheidswaardes (vir die moontlike raai van antwoorde of ander toevalle faktore) is dan ook by die ondersoekgroepe se resultate uitgevoer nie, op grond van die volgende:

Voldoende tyd is toegelaat om al die antwoorde te voltooi en die persentasie toetslinge wat die toetsitems nie voltooi het nie, is baie gering. P-waardes is in hierdie ondersoek in elk geval gebaseer op die totale aantal toetslinge en nie op die aantal toetslinge wat die item probeer het nie.

TABEL 5.18

DISTRIBUSIE VAN TOETSITEMS VOLGENS MOEILIKHEIDSWAARDE-INDEKS

		Eerste toetsing				Tweede toetsing			
Subtoetse:		I	II	III	IV	I	II	III	IV
Mw:	98								
	96	9							
	94								
	92				1				
	90								
	88								
	86								
	84								
	82	1							
	80	2				1 2 3 5			
	78					11			
	76	4 6 11	2 3			13			
	74					6 7	7		
	72								4
	70		5						5
	68	5	6 7						6
	66	3 7 13	8						9
	64				4 8				
	62	18		9		18	8	8 15 16	
	60	8	1		2 3	14	4	10	
	58	14		8 12	5			4 11	
	56	16		11	6 9				
	54			4			5	17	
	52		4	1 15		12 16	9		
	50	10 12						14	
	48								
	46			16	7	19		7	
	44	19		10				19	
	42			13 20		15		20	
	40		9					6 13	
	48			14		17			
	36							3	
	34			3 5 7 17				5	
	32	17				20			
	30			6 19					10
	28								
	26	15	10	18					
	24								
	22	20			10				
	20			2			10	18	
	18								
	16								
	14								
	12							2	
	10								
	8								
	6								
	4								
	2								

Verskille in toetstellings tussen toetslinge mag sowel aan geneigdheid tot die raai van antwoorde as aan ander faktore soos die verskil in vermoë, gereflekteer in die toetsprestasies, begrip van toetsvereistes, aandag en konsentrasievermoë, ensovoort, te wyte wees.

Die geldigheid en betroubaarheid van 'n korreksieformule kan bevraagteken word omdat dit gebaseer is op die veronderstelling dat elke toetsling of die korrekte antwoord ken of raai en verder dat al die antwoordkeuses presies gelykwaardig is vir die toetsling wat die antwoord nie weet nie. Hierdie aanname is selde geregverdig (Thorndike & Hagen, 1969, p.123).

Die uitwerking van korreksie is gewoonlik gering, veral by veelkeusige items en items met 'n hoë moeilikheidswaarde-indeks, byvoorbeeld volgens Guilford (1954, p.421) is:

p ongekorriggeerd	aantal keuses			
	2	3	4	5
,90	= ,800	,850	,867	,875
,80	= ,600	,700	,733	,750
,70	= ,400	,550	,600	,625
,60	= ,200	,400	,467	,500
,50	= ,000	,250	,333	,375

(gekorriggeerde waardes)

Die relatiewe stand van die items word nie so radikaal beïnvloed sodat toe-passing van korreksieformules geregverdig is nie, veral waar 'n toets deeglik gekonstrueer, die tydsbeperking ruim gestel en die toets toepaslik is vir die betrokke ondersoekgroep (Downie & Heath, 1970, p.255). Uiteenlopende standpunte word in die literatuur aangetref met betrekking tot die waarde van die toepassing van 'n korreksieformule wat geldigheid en betroubaarheid betref en of dit 'n oor- of onderkorreksie tot gevolg het (Mehrens & Lehmann, 1973, p.316).

Waarskynlik die belangrikste betekenis van die moeilikheidswaarde-indeks is dat dit die kenmerke van die verspreiding van toetstellings beïnvloed: die simmetrie en standaardafwyking (variansie). Verder hou dit verband met die betroubaarheid van 'n toets: 'n toets wat bestaan uit items wat sentreer om die gemiddelde p-waarde, vertoon 'n laer betroubaarheid as waar die items meer versprei lê om die gemiddeld (Nunnally, 1967, p.251). Verskeie ondersoeke be-

vestig dit dat items met 'n matige moeilikheidswaarde (tussen ,3 en ,7) die hoogste korreleer met die totale toetstelling (Anstey, 1966, p.68).

Diskriminasie-indeks (r_{it})

Dit word gewoonlik aanvaar dat al die items in 'n betrokke subtoets homogeen is met betrekking tot die basiese funksie of eienskap (byvoorbeeld 'toonhoogte') wat gemeet word. Ten einde vas te stel of dit die geval is en elke subtoets 'n enkele aspek of funksie dek en dus 'n homogene karakter en samestelling het, word die interkorrelasies tussen items in die vorm van 'n punt-biserial-korrelasieberkening ondersoek. Dit is naamlik die korrelasie tussen elke item en die subtoetstotaal waarvan dit 'n deel uitmaak en is gedoen deur elke leerling se antwoord te digtomiseer in reg of verkeerd (1 of 0) (sien 'Statistiese tegnieke' hieronder, p.141). Die kovariansie tussen elke item en die totaal waarvan dit deel uitmaak, is 'n funksie van die itemvariansie en die kovariansie tussen die item en die ander items in die subtoets (Du Bois, 1965, p.408). Terselfdertyd gee die verkreeë resultate 'n indeks van die diskriminasiewaarde van elke item deurdat dit die mate aandui waarin elke item diskrimineer tussen leerlinge wat goed en die wat swak in die betrokke subtoets presteer het.

'n Baie lae r_{it} -waarde (byvoorbeeld onder 0,2) gee gewoonlik 'n aanduiding dat die betrokke item ñf baie maklik ñf baie moeilik is. As ander omstandighede gelyk is, kan gesê word dat hoe groter die diskriminasiewaarde van die items, hoe hoër is die betrouwbaarheid of konsekwentheid van die subtoets of toets as geheel. Redelik hoë korrelasies sou dus ideaal wees, dog té hoë korrelasiesyfers kan weer daarop dui dat elke item dieselfde, identiese klein onderdeel van die besondere kenmerk oor en oor meet en die toets nie voldoende differensieer wat die eienskap se meting betref nie. Korrelasies bokant 0,2 word gewoonlik as bevredigend beskou (Oosthuizen, 1972, p.48; Nunnally, 1967, p.142; Colwell, 1969, p.70) terwyl sommige skrywers 0,3 as minimum stel (Aiken, 1971, p.50). Die volgende uiteensetting (Furst, 1958, p.314) gee 'n nuttige algemene leidraad vir die interpretasie van die verkreeë korrelasies:

0,50 en hoër : 'hoog-positief'
0,30 tot 0,49 : 'matig-positief'
0,20 tot 0,29 : 'grenslyn positief'
0,00 tot 0,19 : 'zero tot laag-positief'
-0,01 en laer : 'zero tot negatief'

'n Ander basis vir interpretasie word gegee deur Ahmann & Glock (1971, p.191):

0,4 en hoër = baie goed
0,2 - 0,4 = bevredigend
0,0 - 0,2 = swak

of:

meer as 50% van die items behoort waardes hoër as 0,4,
minder as 40% van die items behoort waardes tussen 0,2 en 0,4,
minder as 10% van die items behoort waardes tussen 0,0 en 0,2 te hê
en geen item behoort negatiewe diskriminasiewaardes te bevat nie.

In soverre as wat die toets- (of subtoets) totaal geldig is (dit wil sê dat dit meet wat dit veronderstel is om te meet), gee die diskriminasie-indeks ook 'n aanduiding van die geldigheid van die individuele toetsitems (inhoudsgeldigheid). Die itemgeldigheid is egter baie minder stabiel en betroubaar as die moeilikheidswaarde-indeks en sal baie meer van een steekproef na 'n ander varieer. Daarteenoor kan beskou word dat die diskriminasie-indeks 'n belangriker maatstaf is by die beoordeling van die waarde van items. Die waarde en kwaliteit van items word grootliks bepaal deur hoe goed die items interkorreleer. Interkorrelasies tussen items hang weer nou saam met die mate waarin elke item met die totale toetstelling korreleer. Dit is dan ook moontlik dat 'n item met 'n 'ideale' p-waarde (digby 0,5) baie laag korreleer met die betrokke toetstotaal (Nunnally, 1967, p.250). Trouens, die diskriminasievermoë van 'n item word beïnvloed deur die relatiewe interkorrelasie van items. By lae interkorrelasies tussen toetsitems (heterogeniteit van items) word maksimum diskriminasie verkry by items met 0,5-moeilikheidswaarde, terwyl by hoë interkorrelasies beter diskriminasie verkry word as die moeilikheidswaardes wyd verspreid lê (Henrysson in Thorndike, 1971, p.152).

TABEL 5.19

DISKRIMINASIEWAARDE-INDEKS (PUNT-BISERIALE KORRELASIES) (N = 115)

Eerste toetsing				Tweede toetsing				Eerste toetsing				Tweede toetsing			
Item	r _{it}	Item	r _{it}	Item	r _{it}	Item	r _{it}	Item	r _{it}	Item	r _{it}	Item	r _{it}	Item	r _{it}
Subtoets I								Subtoets III							
1	+550	11	562	1	586	11	568	1	500	11	351	1	510	11	326
2	589	12	466	2	661	12	488	2	222	12	516	2	246	12	486
3	527	13	534	3	546	13	592	3	325	13	447	3	278	13	407
4	647	14	373	4	540	14	432	4	438	14	384	4	505	14	376
5	577	15	324	5	461	15	469	5	246	15	470	5	249	15	536
6	686	16	075	6	601	16	137	6	446	16	184	6	298	16	300
7	642	17	286	7	681	17	436	7	384	17	424	7	350	17	440
8	553	18	523	8	623	18	307	8	564	18	427	8	364	18	230
9	-007	19	104	9	256	19	385	9	393	19	342	9	371	19	252
10	484	20	122	10	507	20	200	10	259	20	315	10	232	20	143
Subtoets II								Subtoets IV							
1	591	6	612	1	536	6	422	1	302	6	483	1	427	6	571
2	661	7	527	2	484	7	671	2	388	7	559	2	548	7	437
3	486	8	445	3	456	8	487	3	421	8	534	3	506	8	419
4	671	9	621	4	713	9	717	4	372	9	549	4	451	9	499
5	453	10	389	5	459	10	360	5	572	10	326	5	477	10	547

+ = desimale tekens weggelaat

Uit die uiteenstellings in Tabel 5.19 en ontleding van items in Tabel 5.20 blyk die volgende:

Met die eerste toetsing toon 24 uit 60 items (40%) en met die tweede toetsing 22 uit 60 items (36,7%) 'n 'hoog-positiewe' korrelasie en dus baie goeie diskriminasiewaarde. Met die eerste toetsing toon 27 uit 60 items (45%) en met die tweede toetsing 28 uit 60 items (46,7%) 'n 'matig-positiewe' korrelasie en dus 'n redelik goeie diskriminasiewaarde. In totaal toon meer as 80% van die items dus 'n bevredigende tot baie goeie diskriminasiewaarde, volgens die kriteria wat Furst (op. cit.) stel, terwyl slegs sowat 16% van die items 'n beperkte tot twyfelagtige diskriminasiewaarde het. Dit kan dus aangeleid word dat die toets as geheel besondere gebruiksmoontlikhede het.

Verskeie items wat op die uiteindes van die moeilikhedswaardeskaal lê, toon ooreenstemmend 'n lae diskriminasiewaarde, byvoorbeeld items 9 en 20 by Subtoets I en items 2 en 18 by Subtoets III. Daarteenoor vertoon sommige items met 'n lae diskriminasiewaarde tot 'n bykans 'ideale' moeilikhedswaarde van ongeveer 0,5, byvoorbeeld item 16, Subtoets I en item 16, Subtoets III. Op hierdie aspek sal daar weer later teruggekom word met die verdere navorsingsprogram (vide p.260 e.v.).

Betroubaarheidsindeks ($r_{it} \times s_i$)

Hierdie statistiek gee 'n baie goeie aanduiding van die waarde van 'n item deurdat sowel die moeilikhedswaarde as die diskriminasiewaarde van 'n item betrek word (Alberts, 1967, p.69). Die betroubaarheidsindeks word naamlik bepaal deur die produk van die diskriminasiewaarde en die standaardafwyking van die betrokke item. Die standaardafwyking van die toetstotaal is weer gelijk aan die som van die itembetroubaarheidsindekse van al die toetsitems (Gulliksen, 1950, p.377). Die standaardafwyking van 'n item is aan die anderhand gelyk aan die vierkantswortel van die produk van die verhouding toetslinge wat 'n betrokke item reg het en die res van die toetslinge ($s_i = \sqrt{pq}$ waar $q = 1-p$) en hou dus direk verband met die moeilikhedswaardeindeks (p-waarde).

TABEL 5.20

DISTRIBUSIE VAN TOETSITEMS VOLGENS ITEMDISKRIMINASIEWAARDE-INDEKS

Dis= krimi= nasie= waar= des	Hoog-positief	Eerste toetsing				Tweede toetsing			
		SUBTOETSE							
		I	II	III	IV	I	II	III	IV
+98									
96									
94									
92									
90									
88									
86									
84									
82									
80									
78									
76									
74									
72									
70									
68	6	4				7			
66		2				2	7		
64	4 7								
62		6 9				8			
60		1				6 13			
58	2 5					1			
56	1 8 11			8	5	11			
54	13				7	3 4	1	15	6
52	3 18	7	12		9			1	2
50			1			10		4	10
									3 9
48	10	3			6	12	2 8	12	
46	12	5	15			5 15	3 5		4 5
44		8	4 6 13			14 17			7
42			17 18	3			6		1 8
40			9						
38	14	10	7 14	2 4		19	10	14	
36			11						7 8
34			19						9
32	15		3 20	10					11
30				1		18			6 16
Grenslyn- positief	28	17							3
	26			10					10 19
	24			5					2 5
	22			2					18
	20					20			
Zero tot laag-positief	18			16					
	16								
	14								
	12	20							
	10	19							
	8								
	6								
	4								
	2								
	-0								
	-0	16							
		9							

+ = desimale tekens weggelat

Die optimumwaarde van die betroubaarheidsindeks is gelyk aan 0,5 (waar die itemmoeilikheidswaarde 0,5 en die diskriminasiewaarde 1,0 is). Die waarde van die betroubaarheidsindeks daal in verhouding tot die afwyking van die itemmoeilikheidswaarde vanaf 0,5. 'n Betroubaarheidsindeks behoort, ideaal gesproke, nie laer as 0,2 te wees nie (Alberts, 1967, p.71).

Responsontleding

Hierdie metode van itemontleding bied waardevolle addisionele inligting omtrent die intrinsieke waarde van 'n item. Dit behels die vasstelling van die persentasie response wat elke afleier (antwoordkeuse-moontlikheid) trek en het ten doel die bepaling of al die afleiers behoorlik funksioneer. Hoewel 'n item 'n goeie diskriminasiewaarde mag hê, kan dit nog wees dat die trekkingskrag van die afleiers ongebalanseerd is.

Onder ideale omstandighede, met 'n moeilikheidswaarde van 0,5, behoort elke afleier 'n gelyke persentasie van die response te trek, byvoorbeeld in 'n vyfkeusige item: $50/4 = 12,5\%$ of in 'n driekeusige item: $50/2 = 25\%$. 'n Volledige responsontleding van 'n item kan dan byvoorbeeld so daar uitsien:

Item X: Afleiers (antwoordkeuses)

1 12,5%	2* 50,0%	3 12,5%	4 12,5%	5 12,5%
------------	-------------	------------	------------	------------

OF:

1 25%	2 25%	3* 50%
----------	----------	-----------

* = korrekte afleier

Dit is onwaarskynlik dat hierdie ideale verhouding ooit bereik sal word. 'n Meer realistiese kriterium is dat elke afleier minstens 5% van die response moet trek by 'n vyfkeusige item (Owen, 1970, p.62) of 10% by 'n driekeusige item. Indien 'n afleier te min response trek, kan dit daarop dui dat die moontlike antwoorde te opsigtelik verkeerd is, terwyl 'n afleier wat te veel response trek 'n aanwysing kan wees dat die moontlike antwoorde dubbelsinnig en vatbaar vir misverstand is.

'n Goeie afleier behoort nie slegs genoeg response uit te lok nie, maar differensiële trekkingskrag te hê, dit wil sê, dit moet meer aantreklik wees vir dié toetslinge wat matig tot swak in die toets as geheel presteer as vir dié wat relatief hoog presteer. Op hierdie wyse dra 'n afleier by tot die diskriminasievermoë van die toetsitem. Afleiers kan ook nie reg funksioneer indien die korrekte respons te vanselfsprekend is nie. Die gevolg sou 'n te hoë moeilikhedswaarde en 'n te lae diskriminasie-indeks vir die betrokke item wees.

Dit is duidelik dat 'n responsontleding 'n ryke bron van informasie bied omtrent die gehalte van toetsitems.

d-metode van Itemontleding

Hierdie metode word beskryf deur Anstey (1966, p.112-134). Dit vereis die berekening van 'n gemiddelde toetstelling vir elke groep toetslinge wat by 'n item die korrekte antwoord, en dié wat 'n ander antwoord gekies het, sowel as die verskil tussen hierdie stelle gemiddeldes. Anstey onderskei vir hierdie doel tussen 'reg', 'verkeerd', 'uitgelaat' en 'nie probeer' (of 'nie bereik'). Vir die huidige doel, waar voldoende tyd vir die voltooiing van al die toetsitems gegee is, word slegs onderskei tussen 'reg', 'verkeerd' en 'nie gedoen'.

Hierdie vorm van ontleding gee 'n indeks waarvolgens die relatiewe waarde van alle antwoordmoontlikhede onderling vergelyk kan word, sowel as in vergelyking met die waarde van die korrekte respons. Ook kan die relatiewe waarde van verskillende items vergelyk word. Indien die toetsgemiddelde van toetslinge wat een van die verkeerde antwoorde gekies het byvoorbeeld hoër is as die toetsgemiddelde van diegene wat die korrekte antwoord gekies het, moet die betrokke afleier en die item as geheel ondersoek word (vide Henrysson in Thorndike, 1971, p.143). Die toetsgemiddelde kan so 'n belangrike maatstaf wees byvoorbeeld waar meer toetslinge verkeerde antwoorde as die korrekte een gekies het. Indien die toetsgemiddelde van die toetslinge wat die korrekte antwoord verstrek het egter beduidend hoër is as die toetsgemiddelde van die toetslinge wat 'n verkeerde antwoord verstrek het, diskrimineer die betrokke afleier in so 'n geval wel goed en hoef dit dus nie bevraagteken te word nie.

Hierdie metode identifiseer nie slegs onbevredigende items nie, maar bied, in vergelyking met ander metodes, vollediger insig in die moontlike oorsake waarom dit onbevredigend voorkom (Anstey, loc. cit.). Dit kan as 'n uitbreiding van die gebruiklike responsontleding gesien word.

- Besprekking van itemontledingsresultate

Algemene evaluering

Uit tabel 5.21, 5.22, 5.23 en 5.24, wat 'n samevatting van al die verskillende aspekte van itemontleding gee, en volgens kriteria hierbo gestel, blyk dit dat die meeste toetsitems goed funksioneer. Met die drie indekse, moeilikheidswaarde, diskriminasiewaarde en betroubaarheid as maatstawwe, blyk die volgende items egter minder bevredigend te wees, ten spyte daarvan dat twee van die items 'n bykans 'ideale' moeilikheidswaarde van 0,51* en 0,43* het:

Subtoets I:	Moeilikheids= waarde	Diskriminasie= waarde	Betroubaarheid
Item 9:	0,90	0,26	0,08
16:	0,51*	0,14	0,07
20:	0,31	0,20	0,09
Subtoets III			
Item 2:	0,12	0,25	0,08
18:	0,20	0,23	0,09
20:	0,43*	0,14	0,07

Responsanalise

Hierbo is as maatstaf gestel dat 'n afleier minstens 5% van die response by 'n vyfkeusige item en minstens 10% van die response by 'n driekeusige item moet trek. Volgens hierdie kriterium is daar etlike afleiers wat nie bevredigend funksioneer nie:

Subtoets I:	by 11 uit 20 items
II:	10 uit 10 items
III:	6 uit 20 items
IV:	9 uit 10 items

TABEL 5.21

ITEMONTLEDING: BENTLEY-SUBTOETS I : TOONHOOGTEONDERSKEIDING (N = 115 - TWEEDE TOETSING)

Item	p	r_{it}	s_i	$r_{it} \cdot s_i$	Responsanalise*)			% van totale aantal toetslinge			Gemiddeld (\bar{X}) vir subtoets			d	$D(\frac{d}{s_t})$
					S	H	L	reg	verkeerd	nie gedoen	reg	verkeerd	nie gedoen		
1	+8087	+586	+393	+230	0,87	18,26	80,87	80,87	19,13	0,00	14,31	8,09	0,00	6,22	1,491
2	7913	661	406	268	0,00	79,13	20,87	79,13	20,87	0,00	14,54	7,75	0,00	6,79	1,627
3	8000	546	400	218	5,22	80,00	14,78	80,00	20,00	0,00	14,26	8,57	0,00	5,70	1,366
4	8348	540	371	200	2,61	13,91	83,48	83,48	16,52	0,00	14,13	8,05	0,00	6,07	1,456
5	7913	461	406	187	4,35	79,13	16,52	79,13	20,87	0,00	14,11	9,38	0,00	4,74	1,135
6	7391	601	439	264	6,09	20,00	73,91	73,91	26,09	0,00	14,61	8,90	0,00	5,71	1,369
7	7304	681	444	302	7,83	19,13	73,04	73,04	26,96	0,00	14,85	8,45	0,00	6,40	1,534
8	6783	623	467	291	11,40	68,42	20,18	67,83	31,30	0,87	14,91	9,36	9,00	5,56	1,333
9	8957	256	306	078	89,57	6,09	4,35	89,57	10,43	0,00	13,49	10,00	0,00	3,49	0,836
10	6696	507	470	238	15,65	66,96	17,39	66,96	33,04	0,00	14,61	10,11	0,00	4,50	1,079
11	7739	568	418	238	2,61	20,00	77,39	77,39	22,61	0,00	14,40	8,73	0,00	5,67	1,359
12	5217	488	500	244	31,58	52,63	15,79	52,17	46,96	0,87	15,07	11,02	10,00	4,07	0,976
13	7652	592	424	251	6,09	17,39	76,52	76,52	23,48	0,00	14,49	8,67	0,00	5,82	1,395
14	5913	432	492	212	13,91	26,96	59,13	59,13	40,87	0,00	14,62	10,96	0,00	3,66	0,877
15	4261	469	495	232	37,72	42,98	19,30	42,61	56,52	0,87	15,39	11,37	16,00	3,95	0,947
16	5130	137	500	068	51,30	22,61	26,09	51,30	48,70	0,00	13,68	12,54	0,00	1,14	0,273
17	3826	436	486	212	46,96	38,26	14,78	38,26	61,74	0,00	15,43	11,69	0,00	3,74	0,897
18	6261	307	484	149	12,17	25,22	62,61	62,61	37,39	0,00	14,11	11,47	0,00	2,64	0,633
19	4696	385	499	192	31,30	21,74	46,96	46,96	53,04	0,00	14,83	11,61	0,00	3,22	0,772
20	3130	200	464	093	61,74	31,30	6,96	31,30	68,70	0,00	14,36	12,56	0,00	1,80	0,432

N = 115; $\bar{X} = 13, 122$; $s_t = 4,171$; KR-20 = 0,812; $s_m = 1,808$;

Verklaring:

- p : moeilikhedswaarde
 s_i : itemstandaardafwyking
 d : verskil tussen \bar{X} reg en \bar{X} verkeerd
 s_t : toetsstandaardafwyking
 *) : korrekte respons is onderstreep
 H : hoër
 L : laer
 S : selfde
 r_{it} : diskriminasiewaarde
 $r_{it} \cdot s_i$: betroubaarheid
 KR-20 : toetsbetroubaarheid volgens formule
 s_m : standaardmetingsfout
 + : desimale tekens weggelaat
 N : aantal leerlinge
 \bar{X} : gemiddelde telling

TABEL 5.22

ITEMONTLEDING: BENTLEY-SUBTOETS II : MELODIESE GEHEUE (N = 115 - TWEDE TOETSING)

Item	p	r_{it}	s_i	$r_{it} \cdot s_i$	Responsanalise*: % van aantal gedaan						% van totale aantal toetslinge			Gemiddeld (\bar{x}) vir subtoets			d	$D(\frac{d}{s_t})$
					S	1	2	3	4	5	reg	ver= keerd	nie gedoen	reg	ver= keerd	nie gedoen		
1	+8522	+536	+355	+190	4,35	0,00	<u>85,22</u>	6,09	1,74	2,61	85,22	14,78	0,00	6,98	3,47	0,00	3,51	1,511
2	8522	484	355	172	4,35	0,87	<u>0,87</u>	4,35	4,35	<u>85,22</u>	85,22	14,78	0,00	6,93	3,76	0,00	3,17	1,365
3	6783	456	467	213	8,70	2,61	<u>11,30</u>	<u>67,83</u>	9,57	<u>0,00</u>	67,83	32,17	0,00	7,19	4,92	0,00	2,27	0,977
4	6000	713	490	349	39,43	<u>60,00</u>	1,74	1,74	4,35	1,74	60,00	40,00	0,00	7,81	4,43	0,00	3,38	1,455
5	5478	459	498	229	26,09	5,22	1,74	3,48	<u>54,78</u>	8,70	54,78	45,22	0,00	7,43	5,29	0,00	2,14	0,921
6	8696	422	337	142	2,61	1,74	4,35	2,61	<u>1,74</u>	<u>86,96</u>	86,96	13,04	0,00	6,84	3,93	0,00	2,91	1,253
7	7304	671	444	298	14,78	0,87	<u>73,04</u>	6,09	3,48	<u>1,74</u>	73,04	26,96	0,00	7,48	3,97	0,00	3,51	1,511
8	6174	487	486	237	22,61	2,61	<u>0,87</u>	10,43	<u>61,74</u>	1,74	61,74	38,26	0,00	7,35	5,02	0,00	2,33	1,003
9	5130	717	500	359	36,52	<u>51,30</u>	4,35	4,35	<u>2,61</u>	0,87	51,30	48,70	0,00	8,08	4,75	0,00	3,33	1,433
10	2000	360	400	144	60,00	<u>1,74</u>	8,70	<u>20,00</u>	6,96	2,61	20,00	80,00	0,00	8,13	6,04	0,00	2,09	0,900

$$N = 115 \quad \bar{x} = 6,461 \quad s_t = 2,323$$

$$KR-20 = 0,717 \quad s_m = 1,236$$

121

Verklaring:

- p : moeilikheidswaarde
 r_{it} : diskriminasiewaarde
 s_i : itemstandaardafwyking
 $r_{it} \cdot s_i$: betroubaarheid
 d : verskil tussen \bar{x} reg en \bar{x} verkeerd
 $KR-20$: toetsbetroubaarheid volgens formule
 s_t : toetsstandaardafwyking
 s_m : standaardmetingsfout
 S : selfde
 $+$: desimale tekens weggelaat
 N : aantal leerlinge
 \bar{x} : gemiddelde telling
*) : korrekte respons is onderstreep

TABEL 5.23

ITEMONTLEDING: BENTLEY-SUBTOETS III: AKKOORDANALISE (N = 115 - TWEEDE TOETSING)

Item	p	r_{it}	s_i	$r_{it} \cdot s_i$	Responsanalise*)			% van totale aantal toetslinge			Gemiddeld (\bar{X}) vir subtoets			d	$D(\frac{d}{s_t})$
					% van aantal gedoen			reg	ver=keerd	nie gedoen	reg	ver=keerd	nie gedoen		
					1	2	3								
1	+7826	+510	+412	+210	78,26	13,04	8,70	78,26	21,74	0,00	10,94	6,92	0,00	4,02	1,238
2	1217	246	327	080	12,17	55,65	32,17	12,17	87,83	0,00	12,21	9,77	0,00	2,44	0,751
3	3652	278	481	134	8,70	36,52	54,78	36,52	63,48	0,00	11,26	9,38	0,00	1,88	0,579
4	5826	505	493	249	58,26	22,61	19,13	58,26	41,74	0,00	11,46	8,13	0,00	3,33	1,025
5	3478	249	476	119	7,83	34,78	57,39	34,78	65,22	0,00	11,18	9,48	0,00	1,70	0,523
6	4087	298	492	147	40,87	48,70	10,43	40,87	59,13	0,00	11,23	9,26	0,00	1,97	0,607
7	4609	350	498	174	9,57	46,09	44,35	46,09	53,91	0,00	11,30	9,02	0,00	2,28	0,702
8	6261	364	484	176	63,16	22,81	14,04	62,61	36,52	0,87	10,99	8,55	8,00	2,45	0,754
9	7739	371	418	155	6,09	16,52	77,39	77,39	22,61	0,00	10,65	7,77	0,00	2,88	0,887
10	6000	232	490	114	12,17	60,00	27,83	60,00	40,00	0,00	10,61	9,07	0,00	1,54	0,474
11	5739	326	495	161	19,13	57,39	23,48	57,39	42,61	0,00	10,98	8,84	0,00	2,14	0,659
12	6696	486	470	229	66,96	21,74	11,30	66,96	33,04	0,00	11,18	7,82	0,00	3,36	1,034
13	4087	407	492	200	40,87	37,39	21,74	40,87	59,13	0,00	11,66	8,97	0,00	2,69	0,828
14	4957	376	500	188	11,30	49,57	39,13	49,57	50,43	0,00	11,30	8,86	0,00	2,44	0,751
15	6174	536	486	260	61,74	25,22	13,04	61,74	38,26	0,00	11,44	7,86	0,00	3,58	1,102
16	6261	300	484	145	4,35	33,04	62,61	62,61	37,39	0,00	10,82	8,81	0,00	2,01	0,619
17	5478	440	498	219	54,78	30,43	14,78	54,78	45,22	0,00	11,37	8,50	0,00	2,87	0,884
18	2000	230	400	092	20,00	63,48	16,52	20,00	80,00	0,00	11,57	9,70	0,00	1,87	0,575
19	4348	252	496	125	15,65	43,48	40,87	43,48	56,52	0,00	11,00	9,35	0,00	1,65	0,508
20	4261	143	495	071	16,52	42,61	40,87	42,61	57,39	0,00	10,61	9,67	0,00	0,94	0,289

N = 115; $\bar{X} = 10,070$; $s_t = 3,248$; KR-20 = 0,609; $s_m = 2,031$;

Verklaring:

p = moeilikheidswaarde

r_{it} = diskriminasiewaarde

s_i = itemstandaardafwyking

$r_{it} \cdot s_i$ = betroubaarheid (produk)

d = verskil tussen \bar{X} reg en \bar{X} verkeerd

KR-20 = toetsbetroubaarheid volgens KR-20-formule bereken

s_t = standaardafwyking van die subtoets

s_m = standaardmetingsfout

*) = korrekte respons is onderstreep

+ = desimale tekens weggelaat

N = aantal leerlinge

\bar{X} = gemiddelde telling

TABEL 5.24

ITEMONTLEDING: BENTLEY-SUBTOETS IV: RITMIESE GEHEUE (N = 115 - TWEEDE TOETSING)

Item	p	r_{it}	s_i	$r_{it} \cdot s_i$	Responsanalise *)					% van totale aantal toetslinge			Gemiddeld (\bar{X}) vir subtoets			d	$D(\frac{d}{s_t})$
					% van aantal gedoen					reg	ver= keerd	nie gedoen	reg	ver= keerd	nie gedoen		
1	+8609	+427	+346	+148	0,87	6,09	86,09	6,96	0,00	86,09	13,91	0,00	7,17	4,50	0,00	2,67	1,234
2	7043	548	456	250	15,65	1,74	2,61	9,57	70,43	70,43	29,57	0,00	7,57	4,97	0,00	2,60	1,201
3	6696	506	470	238	0,87	1,74	7,83	66,96	22,61	66,96	33,04	0,00	7,57	5,24	0,00	2,33	1,077
4	7652	451	424	191	76,52	2,61	4,35	8,70	7,83	76,52	23,48	0,00	7,34	5,04	0,00	2,30	1,063
5	7391	477	439	209	12,17	73,91	8,70	5,22	0,00	73,91	26,09	0,00	7,45	5,10	0,00	2,35	1,086
6	7217	571	448	256	19,13	0,00	3,48	72,17	5,22	72,17	27,83	0,00	7,57	4,81	0,00	2,76	1,275
7	6087	437	488	213	23,48	4,35	6,09	5,22	60,87	60,87	39,13	0,00	7,56	5,62	0,00	1,94	0,896
8	6957	419	460	193	69,57	1,74	9,57	7,83	11,30	69,57	30,43	0,00	7,40	5,43	0,00	1,97	0,910
9	7217	499	448	224	6,96	72,17	13,91	5,22	1,74	72,17	27,83	0,00	7,47	5,06	0,00	2,41	1,114
10	3130	547	464	254	20,87	14,78	31,30	26,09	6,96	31,30	68,70	0,00	8,58	6,03	0,00	2,55	0,849

N = 115

 $\bar{X} = 6,800$ $s_t = 2,164$ $s_m = 1,299$

KR-20 = 0,640

123

Verklaring:

p = moeilikheidswaarde

 r_{it} = diskriminasiewaarde s_i = itemstandaardafwyking $r_{it} \cdot s_i$ = betroubaarheid (produk)d = verskil tussen \bar{X} reg en \bar{X} verkeerd

KR-20 = toetsbetrouwbaarheid volgens KR-20-formule bereken

 s_t = standaardafwyking van die subtoets s_m = standaardmetingsfout

*) = korrekte respons is onderstreep

+ = desimale tekens weggelaat

N = aantal leerlinge

 \bar{X} = gemiddelde telling

Sommige afleiers het selfs geen response getrek nie. Die andersoortigheid van hierdie tipe toets moet hierby in ag geneem word. Die leerling moet naamlik op 'n akoestiese in plaas van 'n visuele prikkel reageer. Daarby het die toetsopsteller feitlik geen wyse waarop hy 'n afleier kan wysig soos by ander vergelykbare toetse van die veelkeusige-antwoordtipe nie. Die feit dat die ondersoekgroep uit vrywilligers, met 'n klaarblyklike sterk belangstelling in musiek, bestaan, mag 'n verdere moontlike verklaring vir bostaande bevindings bied. By die latere navorsing word daar verder hierop ingegaan (vide p.210 e.v., 257 e.v.).

Die d-metode van Itemontleding

Soos hierbo gestel (vide p.118), word volgens hierdie metode die totale subtoetstelling van die leerlinge wat 'n korrekte antwoord gegee het met die totale toetstelling van dié wat 'n verkeerde antwoord verstrek het, vergelyk.

Uit bostaande itemontledingstabellie blyk dit dat in geen enkele item se geval leerlinge wat 'n verkeerde antwoord verstrek het, 'n hoër gemiddelde totale subtoetstelling as leerlinge wat die korrekte antwoord gegee het, behaal het nie. Items vertoon dus deurgaans bevredigend, diskriminerend en betroubaar. Dié beginsel van vergelyking kan egter verder gevoer word.

Die verskil tussen die gemelde twee tellings (\bar{X} reg minus \bar{X} verkeerd = d-waarde), gee naamlik 'n verdere aanduiding van die relatiewe waarde van elke item in vergelyking met die ander toetsitems. Hierdie verskiltelling kan gedeel word deur die standaardafwyking van die subtoets as geheel: $\frac{d}{S} = D$ -waarde. Aan die hand van 'n klassifikasiesisteem wat Anstey (1966, p.133) voorstel, kan items volgens D-waarde ongeveer soos volg geëvalueer word:

- 1,00 en hoër : goed
- 0,75 - 0,99 : redelik
- 0,50 - 0,74 : twyfelagtig
- 0,25 - 0,49 : swak
- minder as 0,25 : baie swak

Tabel 5.25 klassifiseer die toetsitems volgens hierdie kriteria. Die aantal items wat as 'redelik' tot 'goed' geklassifiseer kan word, is dus soos volg:

Subtoets I : 17 uit 20 items = 85%
 II : 10 uit 10 items = 100%
 III : 10 uit 20 items = 59%
 IV : 10 uit 10 items = 100%

Totaal: 47 uit 60 items = 78,3% terwyl in totaal slegs 4 uit 60 items (6,7%) as 'swak' bestempel kan word. Geen 'baie swak' items kom voor nie.

TABEL 5.25 EVALUERING VAN TOETSITEMS VOLGENS D-waarde

	Goed		Redelik		Twyfelagtig		Swak		Baie swak		Totaal
	Aantal items	%	Aantal items	%	Aantal items	%	Aantal items	%	Aantal items		
Subtoets I	11	55	6	30	1	5	2	10	0		20
II	7	70	3	30	0	0	0	0	0		10
III	4	20	6	30	8	40	2	10	0		20
IV	7	70	3	30	0	0	0	0	0		10
Totaal	29	48,3	18	30	9	15	4	6,7	0		60

- Samevattung

Met betrekking tot die diskriminasiewaarde en konsekwentheid van die toetsitems, kan dus op grond van die voorafgaande ontledings, die gevolgtrekking gemaak word dat (met die uitsondering van Subtoets III) die meeste toetsitems goed funksioneer en die toets as geheel betroubaar voorkom. Vir die doel van 'n itemontleding beantwoord die ondersoekgroep ($N = 115$) egter nie heeltemal aan die vereistes nie (veral wat grootte en samestelling betref - vide Nunnally, 1967, p.242), sodat bestaande gevolgtrekkings as 'voorlopig' gereken moet word. Genoemde aspekte word tydens die verdere ondersoeke weer onder die loep geneem en die resultate vergelyk voor daar finale konklusies gemaak word.

- Norms-metingskale

Ten slotte is ondersoek of die metingskaal wat Bentley verskaf van toepassing is op Suid-Afrikaanse leerlinge en hoedanig afwykings, indien enige, daar uitsien.

Bentley verskaf geen persentiel- of standaardtellingnorme in sy handleiding (1966a) nie, maar gee slegs 'n graderingstabel van routellings volgens ouderdom. Ooreenstemmend met Wing (1970, p.71) word roupunttellings dan ook in vyf klasintervalle verdeel en merk hy op dat dit vir praktiese doeleinades voldoende is en persentielrange of ander kleiner verdelings of onderskeidings nie geregverdig is of enige voordeel inhoud nie (Bentley, op. cit., p.102). Verder gebruik hy slegs die totaal telling vir sy 'norm'-gradering.

In die volgende uiteenstellings word volledige persentiel- en standaardtelling-norme (volgens die C-skaal) vir sowel die subtoetse as vir die toetstotaal gegee. Hierdie metode is in ooreenstemming met die heersende psigometriese praktyk.

Rasional

Standaardtellings, met die standaardafwyking vanaf die rekenkundige gemiddelde as verwysingspunt, bied besondere voordele. Standaardtellings maak direkte vergelyking tussen toetse en subtoetse moontlik - iets wat selde met roupunttellings (absolute tellings) moontlik is. Sodanige getransformeerde tellings het standaardbetekenis, met 'n gemeenskaplike graderingsskaal as verwysingsraamwerk. Teenoor roupunttellings bied dit 'n standaard wat meer uniform, objektief en konstant vir vergelykingsdoeleindes is (Van der Walt, 1970, p.41) (vide Bylae A, p.560).

Persentielrange. 'n Persentiel is die punt in die verdeling van tellings waaronder 'n bepaalde persentasie van die gevalle lê, soos aangedui deur die nommer van die gegewe persentiel. Die sestigste persentiel (P_{60}) is byvoorbeeld die toetstelling waaronder sestig persent van die toetslinge val. 'n Persentielrang is dus die rangorde wat in terme van 'n persentasie uitgedruk word.

Besondere voordele van persentielnorme (bo roupunttellings) is dat dit maklik interpreteerbaar is en 'n vergelyking tussen toetse moontlik maak. Die belangrikste nadeel van die persentielsskaal is dat, aangesien dit 'n ordinale of rangordeskaal is, persentiele nie gelyke eenhede in individuele verskillende verteenwoordig nie en dus nie uniform in grootte is nie. 'n Standaardtelling-

skaal hef hierdie probleem op deur dat in konstante eenhede, gereken vanuit die rekenkundige gemiddelde, gemeet word.

Om op te som: soos by standaardtellings word by persentielsskale tellings in verskillende veranderlikes tot gemeenskaplike eenhede herlei. Verskille in persentiele is egter nie proporsioneel tot verskille in die oorspronklike rou=tellingeenhede nie. Standaardtellings is, met verwysing na die gemiddelde en standaardafwyking, maklik interpreteerbaar.

Bentley se 'norm'-gradering hou verband met 'n persentielsskaal in dié opsig dat hy die ondersoeksgroep in vyf klasse of groepe indeel volgens persentiel=range: P_{90} skei klas A van klas B, P_{70} B van C, P_{30} C van D en P_{10} D van E. In terme van standaardafwyking is dié verdelingspunte $+1,2816\sigma$, $+0,5244\sigma$, $-0,5244\sigma$ en $-1,2816\sigma$ (benaderd $\pm 1,3\sigma$ en $\pm 0,5\sigma$). Met die standaardafwyking as uitgangs=punt, sou sodanige vyf klasse ook onderskei kan word met $\pm 1,5\sigma$ en $\pm 0,5\sigma$ as verdelingspunte. Dit is ongeveer gelyk aan P_{93} , P_{69} , P_{31} en P_7 (vide Du Bois, 1965, p.297, 298; Bailey, 1971, p.81, 82). 'n Ander moontlikheid word hieron=der voorgestel by die behandeling van die C-skaal.

Standaardtellingskale

Die z-skaal. Dit is die eenvoudigste tipe standaardtellingsskaal en het 'n ge=middeld van 0 en standaardafwyking van 1,0. 'n Moontlike beperking op die ge=bruik van hierdie tipe skaal, is dat aanvaar moet word by vergelyking (bv. tussen verskillende toetse) dat die verkreeë verspreiding dieselfde is as die ware, normale populasieverspreiding en dat die werklike gemiddeldes, standaard=afwyking en distribusie uniform is wat betref die eienskappe wat gemeet word. Waar geen bewyse van sodanige normaliteit bestaan nie, moet interpretasie ge=baseer op die gebruik van standaardtellings met omsigtigheid gedoen word. z-tel=lings het verder die nadeel dat die helfte van die tellings 'n negatiewe waarde het, wat berekenings bemoeilik. Nog 'n nadeel is die groot eenheid (1σ) waar=mee gewerk word.

Die T-skaal het 'n gemiddeld van 50 en 'n standaardafwyking van 10. Dit is 'n genormaliseerde standaardskaal wat die besware teen die z-skaal ophef. Een tien=

de van 'n standaardafwyking word as eenheid geneem, sodat 'n gewone verspreiding van 'n groot toetspopulasie minstens 50 tot 60 heeltal T-tellings lewer. 'n Besondere voordeel is die uitbreiding buite die distribusie-omvang van 'n enkele populasie. Die normalisering vind plaas deur die herleiding na sentielrange en daarna na korresponderende afwykings vanaf die gemiddelde in 'n normale verspreiding. 'n Moontlike beswaar teen die T-skaal, is dat dit met 0,1σ as eenheid 'n oorverfynde skaal en valse idee van metingsakkuraatheid mag gee. Dit is verderstrydig met die standpunt dat 'n tellingpunt relatief en benaderd is en toegelaat moet word vir 'n standaardmetingsfout (dieselfde beswaar geld teen die persentielkskaal). Die C-skaal bied 'n oplossing vir hierdie probleem.

Die C-skaal. In plaas daarvan om in 50 of 60 eenhede te werk, word basies elf eenhede, van 0 tot 10 gebruik. Die C-skaal het 'n gemiddeld van 5 en 'n standaardafwyking van 2. In vergelyking met die T-skaal bied dit 'n diskriminasieakkuraatheid wat meer in ooreenstemming is met die mate van akkuraatheid wat van 'n toets verwag kan word. Met 'n standaardafwyking van 2 sou dié skaal op die uiteindes uitgebrei kan word tot -1 en 11. Dit is egter standaardpraktyk om slegs die genoemde elf eenhede te gebruik. Die C-skaal bied ook 'n nuttige gradering van resultate in vyf klasse (vide p.127 hierbo), byvoorbeeld: (vide Karmel, 1970, p.148).

TABEL 5.26

GRADERING VOLGENS C-SKAAL

	A Baie goed	B Goed	C Gemiddeld	D Swak	E Baie swak
C-telling (11)	10 9	8 7	6 5 4	3 2	1 0 (-1)
% in elke eenheid (0,24)	0,92 2,79	6,55 12,10	17,47 19,74 17,47	12,10 6,55	2,79 0,92 (0,24)
% in elke klas	3,71	18,65	54,68	18,65	3,71
Benaderde %	4	19	55	19	4

Die Stanegge-skaal. Dit is 'n variant van die C-skaal deurdat die twee eenhede aan beide die uiteindes gekombineer word sodat 'n negepunt- in plaas van 'n elfpuntskaal gevorm word, met 'n gemiddeld van 5 en 'n standaardafwyking van 1,96 ('n weinig laer as dié van die C-skaal, weens die sametrekking op die eindpunte).

Die verhouding tussen die verskillende normskale word in Bylae A prakties geïllustreer (vide p.560).

- Norms - eksperimentele groep

Die huidige eksperimentele groep is te klein ($N = 115$) om betekenisvolle standerd- of ouderdomsnorme op grond van die toetsresultate op te stel.

Tabel 5.27 gee egter 'n normskaal vir die groep as geheel (standerd 2 tot 6). Soos in die weergawe van resultate van die voorondersoek en hoofondersoek wat later volg, word norme hier in die vorm van persentielrange en C-skaaltelings weergegee.

TABEL 5.27 .

BENTLEY-TOETSTOTAALNORME: PERSENTIELRANGE EN C-SKAALTELLINGS
($N = 115$, TWEDE TOETSING)

Roupunte	Frekwen=sie	Persentiel=rang	C-skaal	Roupunte	Frekwen=sie	Persentiel=rang	C-skaal
53	1	99	(11)	35	3	43	5
52	4	99	10	34	8	41	5
50	2	96	8	33	3	34	4
49	4	94	8	32	2	31	4
48	1	90	8	31	5	30	4
47	6	90	8	30	3	25	4
46	3	84	7	29	4	23	3
45	2	82	7	28	1	19	3
44	5	80	7	27	2	18	3
43	5	76	6	26	3	17	3
42	2	71	6	24	2	13	3
41	6	70	6	23	5	12	3
40	6	64	6	22	3	8	2
39	4	59	5	21	1	5	2
38	4	56	5	20	1	4	2
37	5	52	5	18	3	3	1
36	5	48	5	14	1	1	(-1)

- Vergelyking: eksperimentele groep/Bentley se metingskaal

Tabel 5.28 gee 'n vergelyking van die metingskaal wat Bentley voorsien met die toetsresultate van die huidige eksperimentele groep (vide p.126) hierbo. Vir vergelykingsdoeleindes word die toetsresultate in dieselfde vorm as wat

Bentley (1966a, p.102) dit doen, weergegee.

'n Redelike ooreenkoms in bree uitleg is opvallend, hoewel 'n matige verskuiwing in gradering duidelik blyk. Hier moet die volgende in ag geneem word.

Bentley maak van 'n ongeselekteerde groep leerlinge gebruik terwyl die huidige ondersoekgroep, soos gemeld, vrywilligers is en dus, wat musikale belangstelling en moontlike musikale aanleg betref, nie so heterogeen soos in die geval van 'n algemene steekproef sou wees nie.

Verder is die ondersoekgroep relatief klein sodat afleidings nie sonder meer gemaak kan word nie. In die voorondersoek, waar met baie groter groepe leerlinge se resultate gewerk word, sal 'n verdere vergelyking gemaak en resultate ontleed word (vide p.197, 198).

TABEL 5.28

VERGELYKING IN GRADERINGSTELLINGS: BENTLEY/EKSPERIMENTELE GROEP

Ouderdom in jare	Klas A boonste 10%		Klas B 20%		Klas C middelste 40%		Klas D 20%		Klas E onderste 10%	
	minstens									
9	*39	42	*38-31	41-38	*30-21	37-24	*20-15	23-20	*14	19
10	41	43	40-34	42-38	33-23	37-28	22-18	27-19	17	18
11	44	46	43-38	45-41	37-29	40-30	28-22	29-22	21	21
12	47	49	46-42	48-46	41-31	45-32	30-23	31-27	22	26
13	48	50	47-44	49-46	43-35	45-35	34-25	34-28	24	27

* = Eerste kolom: Bentley se skaal ($N = ca. 2000$);
tweede kolom: eksperimentele groep ($N = 115$).

5.2.3 Eksperimentele groep III

5.2.3.1 Doelstellings

Om die geldigheid van die Bentleytoets eksperimenteel te probeer vasstel (vide p. 56) hierbo.

Rasional.

Voordat 'n toets met vertroue gebruik kan word, is dit nodig om, benewens betroubaarheid en gepaste verspreiding van moeilikhedswaardes, objektiewe bewys van die toets se geldigheid vas te stel (Anstey, 1966, p.194). Bentley beskryf sy toets as "Measures of Musical Abilities". Die probleem is egter of die toets wel as maatstaf vir musikale aanleg (of 'vermoëns') kan dien, dit wil sê of dit wel 'meet wat dit veronderstel is om te meet' (Van der Walt, 1970, p.71) en dit dus aan sy basiese funksie voldoen.

Vir hierdie probleem is daar geen maklike en absolute oplossing nie. Soos wat daar geen eksakte omskrywing van musikale aanleg moontlik is nie (vide hierbo, p. 37 e.v.), kan daar ook geen absolute geldigheidsindeks vir 'n musikale aanlegtoets verkry word nie. Die toetskomponente toonhoogteonderskeiding, melodiee en ritmiese geheue en akkoordanalise vorm verder slegs sommige belangrike aspekte van die totale spektrum van musikale aanleg.

Soms word in die literatuur onderskeid tussen 'konvergente validering' en 'diskriminante validering' as belangrike komponente van konstruktgeldigheid (vide Guilford, 1978, p.437). Hierbo (p.86 - 88) is leerlinge se toetsprestasies vergelyk met hul IK-tellings en skoolvakprestasies en is vasgestel dat die toetstellings nie betekenisvol korreleer met hierdie genoemde veranderlikes wat verskil van dit wat die toets teoreties behoort te meet nie. Diskriminante validering is dus hier toegepas. By die onderlinge vergelyking van verskillende standaard-musikale aanlegtoetse hieronder (vide p.161 e.v.), word 'n ander aspek van die bepaling van konstruktgeldigheid betrek, naamlik die vasstelling of daar hoë korrelasies bestaan tussen toetstellings in die onderskeie toetse en subtoetse - veranderlikes waarmee 'n musikale aanlegtoets teoreties behoort te korreleer. Dit staan bekend as konvergente validering (vide Gouws et al, 1979, p.56, 158, 160).

'n Ander tipe konvergente validering word hieronder toegepas en bespreek, te same met algemene empiriese en voorspellingsgeldigheid.

5.2.3.2 Die ondersoekgroep

Die groep het bestaan uit 23 onderwysstudente wat musiek as spesialiseringsvak neem.

5.2.3.3 Die instrument

Die Bentleytoets soos reeds hierbo beskryf (vide p. 83).

5.2.3.4 Metode

'n Paneel van sewe ervare musiekdosente is gevra om die musikale aanleg van die studente (heeltemal onbekend aan die dosente) aan die begin van hul eerste studiejaar te evalueer. Dit is gedoen op grond van die studente se onvoorbereide voordrag van 'n werk (toets van bladleesvermoë) asook voordrag van twee voorbereide werke na eie keuse. Die klavier was in alle gevalle die medium en beoordeling is op 'n sespuntskaal gelewer:

- 0 = uiters swak
- 1 = swak
- 2 = middelmatig
- 3 = redelik goed
- 4 = baie goed
- 5 = uitstekend

Die gemiddeld van die sewe beoordelaars se puntetoekennings is as onafhanklike kriterium geneem en gekorreleer met die ondersoekgroep se tellings op die Bentleytoets (betroubaarheid, bereken volgens die KR-2 α -formule, was 0,520 vir die beoordelaarspaneelevaluering en 0,608 vir die Bentleytoetsresultate).

Verder is die ondersoekgroep se toetstellings gekorreleer met hul gemiddelde eksamenprestasie in musiek aan die einde van hul eerste, tweede en derde studiejare ten einde die toets se voorspellingsgeldigheid te probeer bepaal.

5.2.3.5 Resultate en gevolgtrekkings

Empiriese geldigheid, dit wil sê hoe naverwant die toetstellings aan die kriteria is (evaluering deur 'n beoordelaarspaneel enersyds en eksamenpunte in musiek andersyds) word in die vorm van geldigheidskoëffisiënte in Tabel 5.29 weergegee.

Wat die toetstotaal betref, word in alle gevalle 'n beduidende geldigheidskoëffisiënt aangetref. Konvergente validering, met die beoordelaarspaneelvaluering as kriterium, is beduidend op die 0,01-peil wat Subtoets III (akkoord-analise) en die toetstotaal betref en beduidend op die 0,05-peil wat Subtoets I (toonhoogteonderskeiding) betref.

Voorspellingsgeldigheid, met eksamenpunte in musiek as kriterium, is meestal beduidend op die 0,05-peil wat Subtoets II (melodiese geheue) asook die toets-totaal betref. Dit is verder opvallend dat die geldigheidskoëffisiënt matig styg met die verloop van tyd (eerste-, tweede- en derdejaars se eksamenpunte as kriterium) - vanaf 0,457 (eksamenpunte aan die einde van die eerstejaar) tot 0,558 (eksamenpunte aan die einde van die derdejaar).

Hoewel die ondersoekgroep te klein is vir afdoende gevolgtrekkings, wil dit tog voorkom of die Bentleytoets bevredigende empiriese geldigheid vertoon en voorspellingswaarde het.

TABEL 5.29

BENTLEY-GELDIGHEIDSKOËFFISIËNTE: PANEELEVALUERING EN EKSAMENPRESTASIES IN MUSIEK (N = 23)

	Subtoets I Toonhoogte= onderskeiding	Subtoets II Melodiese geheue	Subtoets III Akkoord= analise	Subtoets IV Ritmiese geheue	Toets= totaal
Beoordelaars= paneelevaluatie= ring	+418**	228	529*	220	636*
Eksamenspunte eerstejaar	215	415**	277	319	457**
Eksamenspunte tweedejaar	395	303	298	198	508**
Eksamenspunte derdejaar	323	479**	316	334	558*

+ = desimale tekens weggelaat * = betekenisvol op 0,01-peil
** = betekenisvol op 0,05-peil

Opmerking

'n Aantal verdere eksperimente is met groepe skoliere en studente uitgevoer ten einde te probeer vasstel in hoe 'n mate verskillende musikale aanlegtoetse verband hou met mekaar. Omdat hierdie aspek ten nouste saamhang met die finale keuse van toetse vir die hoofondersoek, word hierdie eksperimente (weliswaar deel van die voorbereidende ondersoek) nie hier gerapporteer nie maar later, tydens die bespreking van die algemene metodologie van die hoofondersoek (vide p.162 e.v.).

5.2.4 Samevatting van die belangrikste voorlopige bevindings

- **Betrouwbaarheid**

Wat tydstabiliteit (volgens toets-hertoets-metode bereken) en interne konsekwentheid (volgens KR-21-formule) betref toon die Bentleytoets bevredigende en betekenisvolle toetsbetrouwbaarheid, veral in die lig van die feit dat die hele toets uit slegs 60 items bestaan.

- **Geldigheid**

Empiriese geldigheid, met beoordelaarsevaluering en eksamenprestasies as onafhanklike kriteria, blyk beduidend te wees, veral ten opsigte van die toetstaal. Verder vertoon diskriminante validering, wat korrelasie tussen leerlinge se toetsprestasie, IK-tellings en skoolvakprestasies behels, ook betekenisvol.

- **Diskriminasiewaarde van toetsitems**

Die meerderheid toetsitems toon 'n goeie tot baie goeie diskriminasiewaarde en funksioneer hoogs bevredigend, volgens die verkreeë resultate.

- **Moeilikheidswaarde-indeks**

'n Wye verskeidenheid moeilikheidswaardes word by al die subtoetse aangetref, met 'n heel bevredigende itemvolgorde.

- Vergelyking tussen subtoetsgemiddeldes

Subtoets III toon in alle gevalle die laagste gemiddelde telling terwyl Subtoets IV in een en Subtoets I in 'n ander geval die hoogste telling lewer.

- Interkorrelasies tussen subtoetstellings onderling, subtoetse en die toets-totaal

Subtoetse toon in die meeste gevalle 'n beduidende interkorrelasie en 'n hoogs beduidende korrelasie met die toetstotaal, waaruit dit wil voorkom of die subtoetse in 'n redelike mate oorvleuel ten opsigte van die funksies wat gemitte word. Op hierdie aspek sal in die latere navorsingsprogram teruggekom word.

- Geslag en toetstellings

Feitlik geen betekenisvolle verskil kon tussen die gemiddelde toetsprestasies van seuns en dogters aangetoon word nie, hoewel die dogters meestal 'n beter toetsprestasie gelewer het as die seuns.

- Ouderdom en standerd

'n Betreklik reëlmatige styging in toetsprestasie met styging in ouderdom en standerd is deurgaans opvallend.

5.3 DIE VOORONDERSOEK

5.3.1 Doelstellings

Aangesien die vorige eksperimente met relatief klein en onverteenwoordigende groepe leerlinge uitgevoer is, kon daar kwalik afdoende afleidings en betekenisvolle gevoltrekkings gemaak word. In die lig hiervan is besluit om op 'n omvangryke skaal leerlinge by die volgende ondersoek te betrek. Die belangrikste doel was om die Bentleytoets onder 'n wye verskeidenheid omstandighede uit te toets en 'n beeld te probeer kry van sy praktiese bruikbaarheid, toepasbaarheid en beperkings (sien ook Metodologiese oorwegings hieronder, p.145 e.v.).

'n Belangrike doel van die voorondersoek was dus om te bepaal hoedanig die Bentleytoets as meetinstrument onder Suid-Afrikaanse omstandighede funksioneer en om verder plaaslike resultate te vergelyk met die resultate wat Bentley (1966a) in Brittanje verkry het. Ander doelstellings was die volgende:

- Om ondersoek in te stel na die verskil in toetsprestasie volgens ouerdom, standerd, geslag en taal (skoolmedium).
- Om 'n itemontleding uit te voer ten einde die subtoetssamestelling en afsonderlike toetsitems intensief te ondersoek en dié bevindings te vergelyk met die resultate van die voorondersoek.
- Om voorlopige persentiel- en standaardtellingnormskale vir verskillende ouerdoms- en standerdgroepe te verkry.

5.3.2 Die ondersoekgroep

Tabel 5.30 gee die samestelling van die ondersoekgroep: getal skole (primêr en sekondêr), getal leerlinge, verdeling volgens taal, geslag en skoolgrootte terwyl Tabel 5.31 die verdeling van die groep volgens standerd en ouerdom aandui.

Om die toetsgroep so verteenwoordigend moontlik te maak, is leerlinge uit vyftig primêre skole en twintig sekondêre skole, waarby twee beroepskole ingesluit is, betrek. By primêre skole is alle beskikbare leerlinge vanaf graad 2 tot en met standerd 5 getoets; by sekondêre skole is alle beskikbare standerd 6- en 7-leerlinge, benewens enkele standerd 8-leerlinge, ingesluit. Die aantal leerlinge per skool het gewissel van 55 tot 751.

Volgens sosio-ekonomiese indeks (vide hieronder, p.148, 149) verteenwoordig die groep stedelik, laag en medium en nie-stedelik, laag, medium en hoog.

Figuur 5.6 toon die ligging van die skole in die ondersoekgroep aan. Uit praktiese en ekonomiese oorwegings is die ondersoek naamlik beperk tot Suid- en Wes-Transvaal.

TABEL 5.30

SAMESTELLING VAN ONDERSOEKSGROEP: VOORONDERSOEK

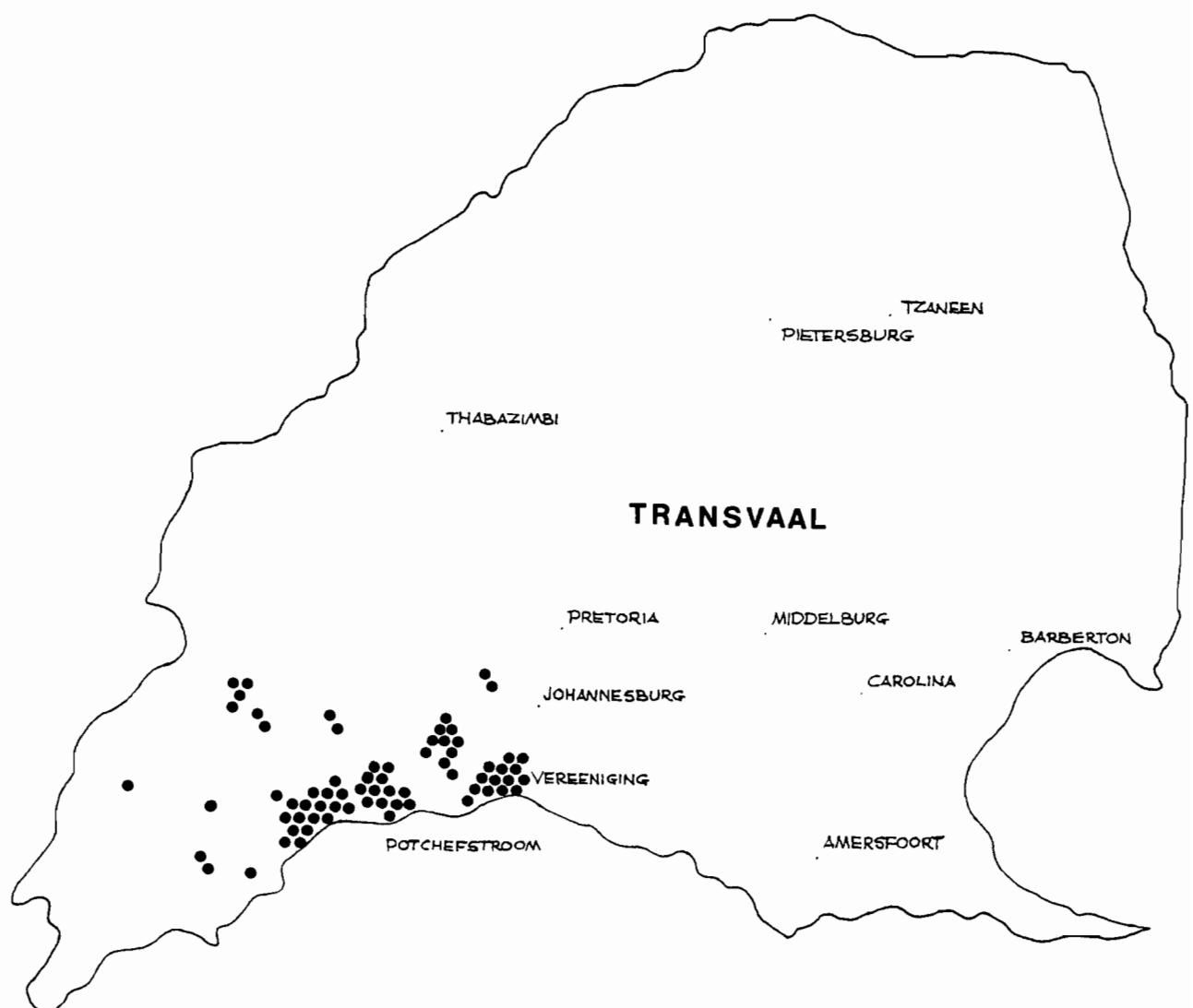
Skooltype	Getal skole						Getal leerlinge					
	Volgens skoolmedium		Volgens leerlingtaal			Totaal	Volgens geslag		Volgens skoolmedium		Totaal	
	Afrikaans	Engels	600 en meer	250-599	minder as 250		Seuns	Dogters	Afrikaans	Engels		
Primêre skole	41	9	23	24	3	50	11398	10946	18649	3695	22344	
Sekondêre skole	15	5	12	7	1	20	3838	3645	6082	1401	7483	
Totaal	56	14	35	31	4	70	15236	14591	24731	5096	29827	

TABEL 5.31

VERDELING VAN ONDERSOEKSGROEP VOLGENS STANDERD EN OUDERDOM

Skool= standerd	Gr. 2	St. 1	St. 2	St. 3	St. 4	St. 5	St. 6	St. 7	St. 8		
N	3755	4044	3758	3672	3891	3224	3919	3528	36		
Ouderdom	6 jaar	7 jaar	8 jaar	9 jaar	10 jaar	11 jaar	12 jaar	13 jaar	14 jaar	15 jaar	16 jaar
N	789	3060	3805	3797	3843	3883	3712	3536	2571	708	123

FIGUUR 5.5: LIGGING VAN SKOLE IN ONDERSOEKSGROEP BY VOORONDERSOEK



Volgens die Chi-kwadraattoets verteenwoordig die groep nie 'n normale verspreiding nie (op sowel die 0,05- as 0,01-peil). Hieronder word die Kolmogorov-Smirnovtoets vir normaalverdeling op die toetstellings toegepas en sal weer na hierdie aspek verwys word (vide p.172).

5.3.3 Die instrument

Bentley (1966): Measures of Musical Abilities (vide hierbo p. 83, 92).

Die toetsmateriaal bestaan uit 'n plaatopname wat die toetsinstruksies sowel as die toetsvrae bevat, 'n nasiensleutel, antwoordblaie en 'n kort brosjure as toetshandleiding. Sy boek (1966a) kan egter as 'n volledige toetshandleiding beskou word.

Die unieke eienskap van hierdie toets, is dat dit die enigste beskikbare toets is wat ontwerp is met die primêre doel om kinders so jonk soos sewe jaar te toets. Die toets se onderskeie subtoetse is dan ook so kort dat dit nie die aandagsvermoë van selfs die sewearige ooreis nie. Elke subtoets duur, saam met die instruksies, ongeveer vyf minute en die totale toets ongeveer 21 minute.

'n Ander besondere kenmerk van die toets is dat Bentley benewens 'n ossillator (Subtoets I, toonhoogte) van 'n pyporrel gebruik maak, teenoor die klavier of strykinstrumente by die meeste vergelykbare toetse. Die orrel bied die voordeel dat dit, teenoor die klavier, 'n meer reëlmatige aangehoue toon kan voorsien en die verskillende tone in 'n akkoord meer akkuraat van gelyke volume kan wees (aanslag speel 'n belangrike rol by die klavier en absolute gelyke toonsterkte is nie altyd moontlik nie).

Die subtoetse, in die bepaalde voorgeskrewe volgorde, bied voldoende kontras en verskeidenheid vir volgehoute belangstelling en aandag, deurdat verskillende tipes oordeel opeenvolgend vereis word: toonhoogteonderskeiding, toongeheue, akkoordanalise en ritmiese geheue. Dié bepaalde volgorde bied dus 'n redelike mate van afwisseling en kan so 'n uitwerking hê op die leerlinge en bydra tot beter motivering, belangstelling en selfvertroue (Van der Walt, 1970, p.166).

5.3.4 Statistiese tegnieke

- Vir die bepaling van verband of kovariansie (korrelasiekoëffisiëntberekening volgens die produk-momentmetode) is die volgende standaard-statistiese formule gebruik:

$$r = \sqrt{\frac{N\sum XY - (\sum X)(\sum Y)}{[\sum X^2 - (\sum X)^2][\sum Y^2 - (\sum Y)^2]}}$$

waar r = korrelasiekoëffisiënt

N = aantal gevalle (proefpersone) of aantal tellings binne elke verspreiding

Σ = som van

X, Y = 2 veranderlikes: X = telling van 'n persoon op 1 veranderlike

Y = telling van dieselfde persoon op die ander veranderlike

$\sum X^2$ = som van die kwadrate van die X -veranderlike

$\sum Y^2$ = som van die kwadrate van die Y -veranderlike

$\sum XY$ = som van die produkte van die twee veranderlikes

(vide Peatman, 1963, p.93; Guilford & Fruchter, 1978, p.83).

Om die beduidendheid van die verkreeë korrelasiekoëffisiënte te bepaal, is gebruik gemaak van 'n standaard-statistiese tabel (vide Du Toit, 1979b, p.65, 66; Roscoe, 1975, p.438).

Vir steekproefgroottes (N -waardes) wat nie in die betrokke tabel voorkom nie, is die beduidendheidswaardes soos volg bepaal: uit die 'verspreidingstabel van t vir gegewe waarskynlikheidsvlakke' (vide Du Toit, op. cit. p.61; Roscoe, op. cit., p.429) is die kritiese t -waardes op 0,01- en 0,05-peil vasgestel en met behulp van die volgende formule na die kritiese r -waardes herleid:

$$r = \sqrt{\frac{t^2}{N - 2 + t^2}}$$

waar r = korrelasiekoëffisiënt

t = teoretiese verspreidingsverhouding

N = aantal tellings/proefpersone

(vide Roscoe op. cit., p.265).

- Vir die berekening van die punt-biseriale-korrelasie is die volgende standaard-formule gebruik:

$$r_{pbi} = \frac{(\bar{X}_1 - \bar{X}_0)}{s_t} \sqrt{pq}$$

waar \bar{X}_1 = gemiddelde toetstelling van leerlinge wat die item reg beantwoord het

\bar{X}_0 = gemiddelde toetstelling van leerlinge wat die item verkeerd beantwoord het

s_t = standaardafwyking van die toets-/subtoetstellings

p = die verhouding van die groep wat die item reg beantwoord het

q = $1-p$ (verhouding van groep wat item verkeerd beantwoord het).

(vide Helmstadter, 1964, p.164; Roscoe, op. cit., p.114; Magnusson, 1967, p.200.)

- Standaardafwyking is bereken volgens die formule (vir ongegroepeerde data)

$$s = \sqrt{\frac{\sum X^2}{N-1}} - \bar{X}^2 \text{ wat 'n makliker berekeningsformule is as die}$$

'definisieformule' : $\sqrt{\frac{\sum x^2}{N-1}}$ deurdat direk vanaf verkreeë waardes bereken kan

word.

$\sum X^2$ = som van die kwadrate van die tellings

\bar{X}^2 = gekwadreerde gemiddelde van die tellings

N = aantal tellings/proefpersone

$\sum x^2$ = som van die gekwadreerde deviasietellings (afwykings vanaf die monstergemiddelde: $x = X - \bar{X}$)

(vide: Van der Walt, 1970, p.262; Horowitz, 1974, p.40; Du Bois, 1965, p.107).

- Betrouwbaarheid is volgens die volgende Kuder-Richardsonformules bereken:

* $N-1$ gee 'n skattting van die populasievariansie en word konsekwent by klein ondersoekgroepes gebruik terwyl daar by groot groepes dikwels direk met N (aantal proefpersone) gewerk word en die verskil in berekening weinig verskil in resultaat lewer (vide Guilford en Fruchter, op. cit., p.65; Roscoe, op. cit., p.69).

$$r_{tt} = \left(\frac{n}{n-1} \right) \left[1 - \frac{\sum pq}{s_t^2} \right] \quad (\text{KR-20})$$

$$r_{tt} = \left(\frac{n}{n-1} \right) \left[1 - \frac{\bar{X}(n-\bar{X})}{n s_t^2} \right] \quad (\text{KR-21})$$

waar n = aantal toetsitems

s_t^2 = toetsvariansie

\bar{X} = toetsgemiddelde

p = verhouding korrekte response

q = 1-p

pq = itemvariansie

(vide Nunnally, 1967, p.196; Downie & Heath, 1965, p.246; Mehrens & Lehmann, 1973, p.113).

Die verskil tussen die twee formules is dat KR-21 gelyke moeilikheidswaardes vir alle items veronderstel, dus dat p-waarde konstant is. Indien hierdie aanname nie waar is nie, sal KR-21 'n effens laer beraming van die betroubaarheid gee as die KR-20-formule.

- Vir die berekening van die standaardmetingsfout geld die volgende formule:

$$s_m = s_t \sqrt{1-r_{tt}}$$

waar s_t = standaardafwyking van die toets

r_{tt} = toetsbetroubaarheid

(vide Karmel, 1970, p.112; Guilford & Fruchter, op. cit., p.413).

- Standaardformules vir Skeefheid of asimmetrie ($\sqrt{\beta_1}$ of g_1) en vir Kurtose of spitsheid (β_2 of g_2):

$$g_1 = \sqrt{\frac{\sum x^3}{\left(\frac{\sum x^2}{N}\right)^3}} \text{ of } \frac{\sum (X - \bar{X})^3}{Ns^3}$$

$$g_2 = \frac{\sum x^4}{\left(\frac{\sum x^2}{N}\right)^2} - 3 \text{ of } \frac{\sum (X - \bar{X})^4}{Ns^4} - 3$$

waar $\sum x^3 / \sum x^4$ = som van die deviasietellings tot die derde-/vierdemag

$\sum x^2$ = som van die gekwadreerde deviasietellings

s^3 / s^4 = standaardafwyking tot die derde-/vierdemag

X = toetstelling

\bar{X} = gemiddelde telling

N = aantal tellings

(vide Peatman, op. cit., p.73-75; Du Bois, op. cit., p. 290, 293; Tate, 1965, p.85).

- Formule vir die berekening van χ^2 in die χ^2 -toets vir normaalverdeling :

$$\chi^2 = \sum \frac{(f - f^1)^2}{f^1}$$

waar f = empiriese frekwensie

f^1 = teoretiese frekwensie

(vide Guilford & Fruchter, op. cit., p.197; Siegel, 1956, p.43).

- Formule vir z-toets vir enkelvoudige vergelyking van gemiddeldes:

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Beduidendheid van verskil:

$$z \geq 1,960 \text{ (0,05-peil)}$$

$$z \geq 2,576 \text{ (0,01-peil)}$$

\bar{X}_1 / \bar{X}_2 = gemiddelde telling van eerste/tweede groep
 s_1^2 / s_2^2 = variansie van eerste/tweede groep
 n_1 / n_2 = aantal tellings/persone in eerste/tweede groep

(vide Peatman, op. cit., p.297; Roscoe, op. cit., p.173).

• Meervoudige vergelyking van gemiddeldes (eenrigting-variensie-analise):

F-toets:

$$F = \frac{GSB}{GSF} = \frac{\frac{SSB}{k-1}}{\frac{SSF}{N-k}}$$

Berekeningsformule:

$$F = \frac{\frac{\sum \frac{(\Sigma X)^2}{n} - \frac{(\Sigma \Sigma X)^2}{N}}{k-1}}{\frac{\Sigma \Sigma X^2 - \sum \frac{(\Sigma X)^2}{n}}{N-k}}$$

waar GSB = Gemiddelde som, van kwadrate tussen groepe (variensie tussen groepe)

GSF = Gemiddelde som van kwadrate binne groepe (gemiddelde variensie binne groepe)

SSB = Som van kwadrate tussen groepe

SSF = Som van kwadrate binne groepe

k-1 = vryheidsgrade vir GSB

N-k = vryheidsgrade vir GSF

$(\Sigma X)^2$ = kwadraat van die som van die tellings

$(\Sigma \Sigma X)^2$ = kwadraat van die som van die groepe se totaaltellings

$\Sigma \Sigma X^2$ = totaal van die som van die groepe se gekwadreerde tellings

k = aantal groepe

N = totale aantal tellings (groepe saam)

n = aantal tellings in elke groep

(vide Roscoe, op. cit., p.296-299; Snedecor & Cochran, 1976, p.265 e.v.; Guilford & Fruchter, op. cit., p.229 e.v.).

Die beduidendheid van die verkreeë F-waardes is bepaal volgens die standaard-statistiese tabelle van F-verspreiding (vide Du Toit, op. cit., p.68-72; Roscoe, op. cit., p.460, 461; Snedecor & Cochran, op. cit., p.560-563).

- Kolmogorov-Smirnov-normaalverdelingtoets:

$$D = \text{maks. } |F(x) - \hat{F}(x)|$$

waar $F(x)$ = teoretiese relatiewe kumulatiewe frekwensieverspreiding

$\hat{F}(x)$ = empiriese relatiewe kumulatiewe frekwensieverspreiding

D = maksimumverskil tussen die twee frekwensieverspreidings.

Soos by die χ^2 -toets, kan die teoretiese frekwensie op verskillende maniere bepaal word. Een so 'n metode is die volgende:

Die verkreeë toetstellings word in gepaste klasintervalle verdeel en die kumulatiewe frekwensie (proporsie) bepaal. Met gebruikmaking van die gemiddelde (\bar{x}) en standaardafwyking (s) van die toetstellings en die formule

$$z = \frac{(x - \bar{x})}{s}$$

word bepaal watter proporsie van 'n normale verspreiding tussen die grense van elke klasinterval lê (met verwysing na 'n standaard-statistiese tabel, byvoorbeeld, Du Toit, op. cit., p.45, 46). Hierdie proporsie word met die empiriese kumulatiewe frekwensie vergelyk en die grootste verskil of afwyking aangeteken. Die beduidendheidspeil vir die maksimumafwyking (D -waarde) word volgens die toepaslike standaard-statistiese tabel (vide Du Toit, op. cit., p.75) bepaal.

(Vide Korin, 1975, p.296-298, 383; Siegel, op. cit., p.47-51; Roscoe, op. cit., p.271-274).

5.3.5 Metodologiese oorwegings

- Die Bentley-antwoordblad vereis van die leerling om 'n syfer of letter langs elke vraagnommer neer te skryf. Uit praktiese oorwegings en veral ter vergemakliking van die oordra van antwoorde, is die uitleg verander deur al die moont=

like responsposisies by elke vraag aan te gee, sodat die leerling nou slegs 'n kruisie of ander merk in die betrokke blokkie hoef te maak om sy antwoord aan te dui. Die besondere voordele hiervan, is dat die jonger leerling veral sy antwoorde gouer en akkurater kan neerskryf. Dan is verwarring by die nasien, byvoorbeeld tussen 'S' en '5', uitgeskakel en is nasien en kontrole met behulp van 'n masker moontlik gemaak. Hierdie prosedure is dan ook in ooreenstemming met feitlik alle resente vergelykbare psigometriese toetse. Die volgende dien ter illustrasie:

Oorspronklike Bentley-
antwoordblad
R I T M E

1	
2	

ens.

Gewysigde Bentley-
antwoordblad
R I T M E

1	S	1	2	3	4
2	S	1	2	3	4

ens.

- Die Afrikaanse klankbaan is soos volg verkry. Die oorspronklike Engelse aanswings is so noukeurig moontlik in Afrikaans vertaal om dieselfde konnotasie, assosiasie en gevoelswaarde as die oorspronklike te bevat. Na afloop van die eerste eksperimente, is dit weer gekontroleer en wysings ter verbetering aangebring (vide p.93 hierbo). Die gewysigde teks is finaal deur 'n hoogs ervare voordragkunstenaar voorgedra en op band vasgelê waarna dit sorgvuldig op die oorspronklike klankbaan ingelas is met inagneming van die algemene tydsverloop, pousses tussen items, ensovoort. Die opnames is deur 'n ervare klanktegnikus in 'n goed toegeruste klankateljee gemaak na die nodige redigering en algemene kontrole.

- In aansluiting by die algemene doelstelling (vide p.135, hierbo) om die toets onder 'n wye verskeidenheid omstandighede uit te toets, is verskillende toetsomstandighede op die proef gestel. So is enkele klasgroepe in 'n gewone klaskamer geneem; dan is klasgroepe gekombineer en in 'n groter lokaal soos 'n biblioteek of skoolsaal getoets. Verskillende tyds- en weersomstandighede is beproef, byvoorbeeld vroeg in dieoggend, laat in die skooldag, in die winter, somer en lente met verskillende variasies in temperatuur en vogtigheidstoestande.

Verskillende praktiese en organisatoriese aangeleenthede soos die gebruik van verskillende soorte banke, tafels en stoele, verskillende plasings van die

luidsprekers en toetsapparaat, die hantering van skryfapparaat, toetsmateriaal, ensovoort, is nagegaan en deeglik uitgetoets.

- Met behulp van nasienmaskers is die antwoordblaaien nagesien en die nodige puntebewerkings gedoen. Die versamelde data is vanaf die antwoordblaaien op ponskaarte oorgedra en, nadat die nodige redigering en programmering gedoen is, het die rekenaarverwerkings gevolg.

5.4 DIE HOOFONDERSOEK

5.4.1 Algemene doel en beplanning

In die lig van die besondere doelstellings hierbo genoem (p.6), is alle beskikbare (veral resente) aanlegtoetse bestudeer om sodoende, vir hierdie deel van die navorsingsprogram, 'n eklektiese keuse van toetse te kan maak. Die algemene oogmerk was om hierdeur 'n insig in die ooreenstemmende en afwykende, eiesortige eienskappe van bestaande toetse te kon verkry (vide p.79 hierbo) en so 'n verteenwoordigende beeld moontlik, van die huidige stand en betekenis van musikale aanlegtoetsing op te bou.

Waar die voorondersoek tot 'n enkele toets beperk was, is met hierdie finale deel van die ondersoekprogram beoog om, ter wille van breër perspektief, verskillende toetse te gebruik wat in 'n mate oorvleuel, maar daarby elk 'n eie karakter het en musikale aanleg op 'n unieke wyse benader. Die program het dan ook ingesluit: 'n basiese aanlegtoets met die aksent veral op die jong leerling (Bentley); 'n aanlegtoets wat elemente van 'n prestasietoets bevat en ervaring en huislike agtergrond bykomstig insluit (Gaston); 'n prestasietoets met aanleg inbegrepe (Colwell); 'n belangstellungstoets wat smaak, voorkeur en musikale agtergrondkennis insluit (Skrywer).

Om verder insig in die rol van persoonlikheids- en belangstellingsfaktore te probeer bekom, is twee vraelyste waarvan een deur die leerling self en die ander deur die klasonderwyser voltooi moes word, by die program ingesluit.

Met die voorondersoek is 'n groot aantal beskikbare en geredelik bekombare skole

en skoolpopulasies betrek. Om volledige gevolgtrekkings en afleidings ten opsigte van korrelasies, verskille tussen gemiddeldes, norms en faktorontleding te kan maak, was dit wenslik om van 'n gestratifiseerde steekproef vir die hoofondersoek gebruik te maak, veral ook met die oog op standaardisasie van, en normbepaling vir, die betrokke toetse vir eventuele gebruik in Suid-Afrikaanse skole.

5.4.2 Die ondersoekgroep

Dit volg uit bostaande dat 'n verteenwoordigende gestratifiseerde steekproef van skole en leerlinge as essensieel beskou is. Die vereistes waaraan sodanige steekproef moet voldoen, is:

- dit moet verteenwoordigend van die betrokke bevolking en op 'n onpartydige basis saamgestel wees;
- dit moet groot genoeg wees sodat die statistieke wat daarop bereken word, betroubaar sal wees;
- dit moet klein genoeg wees sodat die dataversameling en daaropvolgende analise effektief en ekonomies uitgevoer kan word (Du Bois, 1965, p.328; Edwards, 1958, p.50, 51).

Stratifikasie behels die proses waardeur die betrokke populasie of universum in onderskeibare subpopulasies, kategorieë of strate ingedeel word om min of meer homogene groepe te verkry ten opsigte van die eienskap wat ondersoek moet word en, na alle waarskynlikheid, verwag word om 'n sterk liniêre verband daarmee te toon. Die betrokke strata vorm so demografiese veranderlikes waarin die navorsing geïnteresseerd is (Garrett, 1953, p.205; Guilford, 1956, p.117, 158). Gestratifiseerde steekproeftrekking dra daar toe by om die betrouwbaarheid en presisie van die norms te verhoog (Angoff in Thorndike, 1971, p.551).

By vergelykbare psigometriese navorsingswerk word gewoonlik 'n sosio-ekonomiese en geografiese indeling gebruik (Gordon, 1965, p.45; Colwell, 1969, p.44; Lindvall & Nitko, 1975, p.141, 142). Dit word dan ook normaalweg aanvaar dat skolastiese prestasie en verwante eienskappe 'n verband toon met die sosio-ekonomiese vlak van die leerlinge, sowel as met die huislike en algemene so-

siale milieu waarbinne 'n kind grootgeword het. Op grond hiervan, is die skoolpopulasie van Transvaal vir die ondersoek demografies ingedeel in stedelik en nie-stedelik en elk van hierdie strata onderverdeel in die volgende substrata: sosio-ekonomies hoog, laag en gemiddeld (gebaseer op die beroep en mediaan inkomste van die vader*).

Om verteenwoordigend te wees, is die aantal leerlinge in elke kategorie op 'n kwotabasis bepaal volgens die verhouding van elke substratum tot die universum (Green, 1970, p.288; Aiken, 1971, p.54). 'n Ander veranderlike wat in ag geneem is, is die verhouding van Afrikaans- tot Engelssprekende leerlinge in die populasie.

Vir die bepaling van die steekproefgrootte is as basis 400 leerlinge per skoolstanderd geneem op grond van formules en tabelle in die literatuur (NEA, 1960, p.99; Guilford, 1965, p.214; Krejcie & Morgan, 1970, p.607-610). Die totale skoolpopulasie (1971**) bestaan naamlik uit 453879 leerlinge in Transvaalse skole met totale leerlingtal per klas wat varieer vanaf ongeveer 42000 leerlinge in Graad 1 tot ongeveer 16500 leerlinge in Standerd 10***). Volgens genoemde bronne sou 'n steekproefgrootte van 384 leerlinge per standerd uit die betrokke populasie voldoende wees.

Hoofsaaklik om praktiese redes, is vyftien leerlinge per standerd as basis geneem omdat dit 'n redelik maklik hanteerbare groep leerlinge per skool lewer (byvoorbeeld 75 leerlinge per sekondêre skool met vyf standerds) en die getal skole in die steekproef (byvoorbeeld $400/15 =$ ongeveer 27 sekondêre- en 27 primêre skole) binne die tydsbestek van ongeveer drie tot vier maande afgehandel kon word.

Die volgende tabel dien ter illustrasie hoe die getal skole per stratum bepaal is en proporsionele toekenning per stratum geskied het - volgens ligging, sosio-ekonomiese indeks en taalmedium:

*) Volgens inligting verkry uit die 'Talentopname' (1965) deur die destydse Nasionale Buro vir Opvoedkundige en Maatskaplike Navorsing (tans RGN) en Departement van Statistiek.

**) Volgens Afdeling Statistiek, Transvaalse Onderwysdepartement.

***) Die presiese syfers vir die getal leerlinge per standerd in Privaatskole was nie beskikbaar nie en daarom word benaderde syfers verstrek.

TABEL 5.32

VOORBEELD VAN VERHOUING VAN GETAL LEERLINGE PER STRATUM TOT UNIVERSUM

Stratum 2*	Aantal leerlinge in universum	Aantal leerlinge in stratum	% van Universum in stratum
	Afrikaans Engels Totaal		
	158858	11764 15464 27228	17,14
% Verhouding Afrikaans/Engels		43,21 56,79	

* = sekondêre skole, stedelik, sosio-ekonomies gemiddeld

Die verhouding van die aantal leerlinge in die betrokke stratum tot die aantal leerlinge in die universum, is $27228/158858 = 17,14\%$. Die aantal leerlinge wat in hierdie stratum per standerd getoets moet word om ongeveer 400 leerlinge per standerd vir die totale steekproef te gee, is dus 17,14% van 400, dit wil sê 69. Die aantal skole in hierdie stratum sou dus $69 + 15 = 4,6$ wees, dit wil sê 5 tot die naaste heeltal (15 leerlinge per standerd per skool is hierbo as basis gekies). Volgens die verhouding Afrikaans- tot Engelssprekende leerlinge sou ongeveer twee Afrikaans- en drie Engelsmediumskole ingesluit moet word.

Die bepaling van die betrokke skole wat volgens bogenoemde uiteensetting ingesluit moes word, het geskied deur aan elke skool in die stratum 'n unieke, identifiserende reeksnommer toe te ken en die betrokke vereiste getal volgens 'n tabel van toevalsyfers te trek. Trekking geskied sonder terugplasing; met ander woorde, 'n bepaalde skool kan slegs een keer voorkom en elk het 'n gelyke kans om getrek te word.

Tabel 5.33 toon die samestelling van die steekproef wat betref die getal skole en leerlinge, verdeel volgens skooltipe, skoolgrootte, taal en geslag. Figuur 5.7 gee 'n aanduiding van die ligging van die skole in die ondersoeksgroep.

TABEL 5.33

SAMESTELLING VAN ONDERSOEKGROEP: HOOFONDERSOEK

GETAL SKOLE	Akademiese skole				Beroepskole		Volgens leerlingtal			Totaal
	Afrikaans	Engels	Afrikaans/Engels	Privaat Engels	Afrikaans	Afrikaans/Engels	600 en meer	250-599	minder as 250	
Primêre skole	15	7	2	2	-	-	13	8	5	26
Sekondêre skole	14	6	1	2	1	4	20	8	-	28
Totaal	29	13	3	4	1	4	33	16	5	54

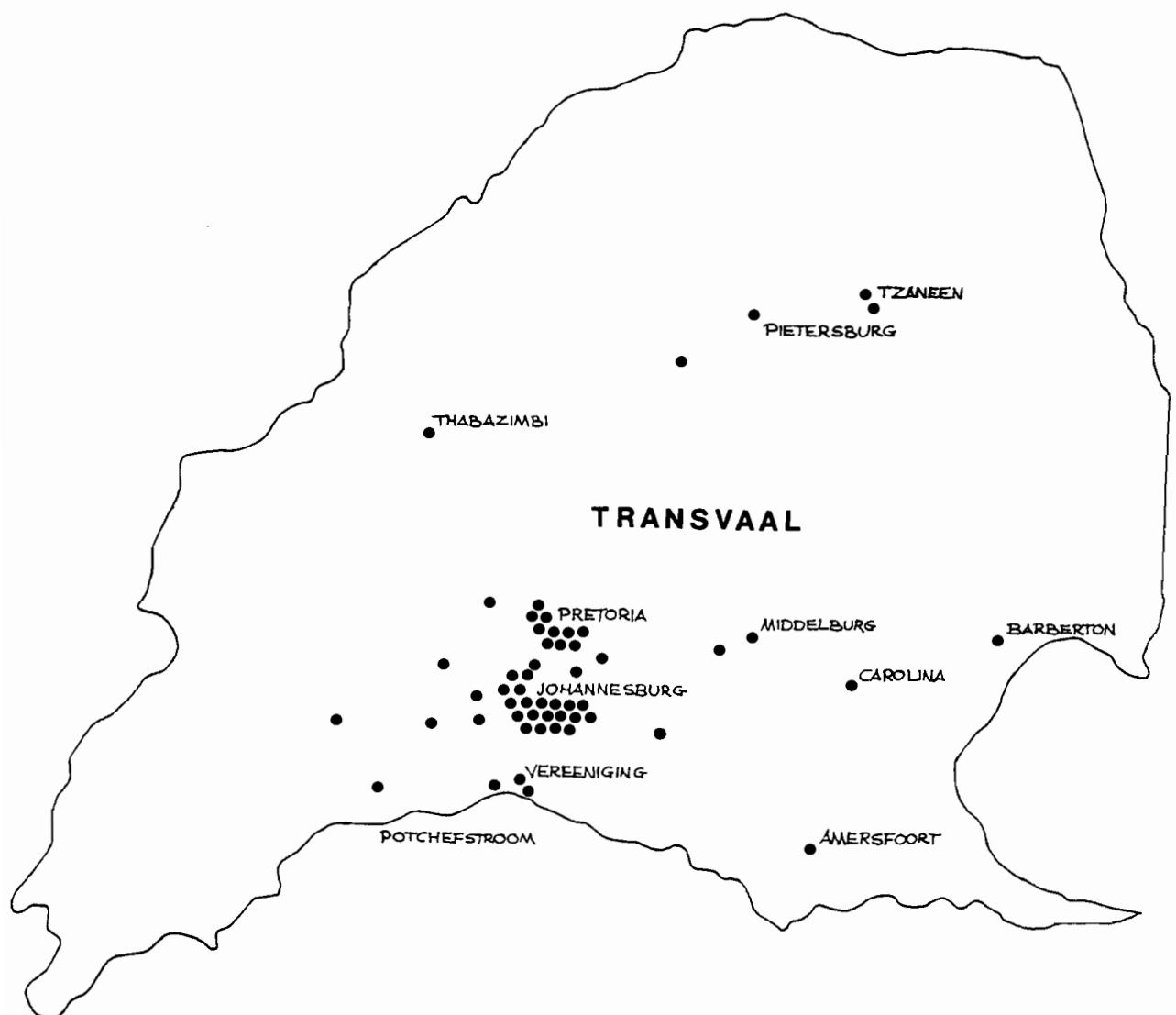
15

GETAL LEERLINGE	Volgens geslag seuns dogters	Volgens taal			Volgens leerlingtal			Totaal	
		Afrikaans	Engels	Ander	600 en meer	250-599	minder as 250		
Primêre skole	1099* 748**	1182 805	1531 1060	697 456	54 39	1143 778	788 537	351 240	2282 1555
Sekondêre skole	1067	1031	1374	629	94	1501	596	-	2097
Totaal	2166* 1815**	2213 1836	2905 2434	1326 1085	148 133	2644 2279	1384 1133	351 240	4379 3652

* = leerlinge vanaf graad 2 tot en met standerd 10

** = leerlinge vanaf standerd 2 tot en met standerd 10 (graad 2 en standerd 1 leerlinge het slegs die Bentleytoets afgelê terwyl standerd 2 tot standerd 10 leerlinge al die toetse gedoen het).

FIGUUR 5.6: LIGGING VAN SKOLE IN ONDERSOEKSGROEP BY HOOFONDERSOEK



Sistematiese steekproeftrekking is by elke skool toegepas om te bepaal watter leerlinge uit elke standerd die toetsprogram sou deurloop. Hier is van 'n alfabetiese naamlys van al die leerlinge in 'n betrokke standerd gebruik gemaak. As 'n betrokke standerd byvoorbeeld uit 60 leerlinge bestaan, is elke vierde leerling ($60 + 4$) vanaf 'n willekeurig gekose naam getrek om sodende die vereiste vyftien leerlinge aan te wys (vide Angoff in Thorndike, 1971, p.551).

5.4.3 Kontrolegroep

Vir algemene vergelykingsdoeleindes asook geldigheidsbepaling is van 'n kontrolegroep van leerlinge uit drie skole gebruik gemaak. Die groep het bestaan uit 89 primêre en 99 sekondêre skoolleerlinge wat met gespesialiseerde musiekopleiding besig was en 'n streng keuring *) vir toelating tot die betrokke skole deurloop het.

Tabel 5.34 toon die samestelling van die kontrolegroep, volgens ouderdom, standerd, geslag en taal.

TABEL 5.34

SAMESTELLING VAN DIE KONTROLEGROEP (HOOFONDERSOEK)

Ouderdom jaar	Getal leerlinge	Standerd		Geslag		Taal		
		Standerd	Getal leerlinge	Seuns	Dogters	Afrikaans	Engels	Ander
7	3	1	6	6	-	4	2	-
8	6	2	10	10	-	8	2	-
9	10	3	12	12	-	8	3	1
10	18	4	31	31	-	17	13	1
11	26	5	30	30	-	20	10	-
12	41	6	35	23	12	15	20	-
13	30	7	22	10	12	7	12	3
14	17	8	13	4	9	4	7	2
15	16	9	16	4	12	7	7	2
16	16	10	13	3	10	6	6	1
17	5							
Totaal	188	Totaal	188	133	55	96	82	10

*) Keuring geskied veral op grond van leerlinge se stemkwaliteit (wat insluit suiwer intonasie en stemproduksie), vaardigheid met betrekking tot instrumentale spel, gehoor- en bladleesvermoë en algemene evaluering van musikale aanleg deur hoogs ervare musiek personeel.

5.4.4 Die instrumente

5.4.4.1 Bentley (1966): Measures of Musical Abilities (MMA)

Dieselde toets as wat met die voorondersoek gebruik is, is ook met die hoof-ondersoek gebruik en word dus nie weer hier bespreek nie.

5.4.4.2 Colwell (1968): Music Achievement Tests I (MAI)

Die toetsmateriaal bestaan uit 'n plaatopname met aanwysings sowel as die toetsvrae, 'n nasienmasker, antwoordblaie en toetshandleiding.

Die toets bestaan uit drie subtoetse: Toonhoogteonderskeiding (25 items), Intervalonderskeiding (28 items) en Metrumonderskeiding (15 items); alle toets-items is driekeusig.

Die toets is die eerste uit 'n reeks van vier prestasietoetse*). Die volle toets duur slegs 18 minute en is interessant saamgestel veral deurdat in 'n ruim mate van redelik bekende volksliedjies en ander melodieë gebruik gemaak word asook die feit dat verskillende instrumente gebruik word: viool, tjello, klavier. In hierdie opsig verskil die toets aanmerklik van die Bentleytoets.

- Die toonhoogteonderskeidingsubtoets wyk af van die ooreenstemmende toetse van Seashore, Kwalwasser-Dykema en Bentley in die volgende opsigte. In plaas van 'n ossillator gebruik Colwell 'n viool en tjello. Die kleinste interval wat gebruik word, is die halftoon - teenoor $3/26$ van 'n halftoon (3 trillings per sekonde) by Bentley en $9/100$ van 'n halftoon (2 trillings per sekonde) by Seashore. Die totale toonomvang wat gebruik word, is egter veel groter as in ander vergelykbare toetse: 82,4 - 2489,2 trillings per sekonde (bykans vyf oktawe) teenoor 414 - 466 trillings per sekonde (heeltoon) by Bentley en 483 - 517 trillings per sekonde (halftoon) by Seashore.

*) Soos reeds genoem (vide p.66), is 'n musikale aanleg- en prestasietoets nie altyd duidelik onderskeibaar nie. Hierdie toets bevat duidelik elemente van 'n aanlegtoets.

In die eerste deel van hierdie subtoets moet die leerling telkens aandui watter van twee tone die hoogste is (of dieselfde) in plaas van of die tweede toon hoër, laer of dieselfde is. In die tweede deel word drie tone voorsien en moet aangedui word watter een van die drie die laagste is. Hierdie tipe onderskeiding kom by geen ander toets voor nie en is bedoel om 'n onderskeiding te vereis in 'n verband nader aan die werklike musikale situasie. Benewens eenvoudige vergelyking, moet leerlinge nou ook vorige toonhoogtes kan behou (toongeheue om die nodige vergelyking te kan maak) - dieselfde vermoë as wat by 'n musikale voordrag of aktiewe musiekbeluistering vereis word en nodig is om akkurate intonasie te verkry.

- Die intervalonderskeidingsubtoets bestaan, soos die vorige, uit twee dele. Die eerste deel vereis van die leerling om trapsgewyse beweging te onderskei van beweging in spronge en bevat tien drietoonitems. Die tweede deel vereis 'n soortgelyke onderskeiding maar bestaan uit agtien volledige frases van volks- en ander melodieë. Werklike intervalidentifisering is eintlik nie hier ter sprake nie omdat dit slegs 'n onderskeiding vra tussen sekundes (trapsgewyse beweging) en groter melodiese intervalle (spronge).
- Die derde subtoets, metrumonderskeiding, bestaan uit vyftien frases, weer eens geneem uit die volks- en ander liedliteratuur. Die leerling moet onderskei tussen basiese twee- en driemaatslag. Deur die gebruik van talryke bekende melodieë verseker die toetsoueur dat die leerling se aandag en belangstelling deeglik behou word.

5.4.4.3 Gaston (1957): A Test of Musicality (TOM)

Gaston se toetsmateriaal bestaan soos die vorige toets uit 'n plaatopname waarop die toetsitems sowel as volledige toetsinstruksies voorkom. Verder word 'n kort handleiding, waarin onder andere ook die volledige instruksies verskyn, 'n nasienmasker en antwoordblad, voorsien.

'n Besondere eienskap van die toets is dat dit 'n selfbeoordelingsvraelys insluit waarin die leerling vrae moet beantwoord met betrekking tot sy voorkeure vir verskillende musikale aktiwiteite en die betekenis van musiek in

sy huislike milieu. Verder bevat die antwoordblad 'n ruimte waarop die onderwyser die leerling se gesondheid en fisieke en persoonlikheidseienskappe kan beoordeel.

Colwell (1969, p.9) beskou die Gastontoets primêr as 'n prestasietoets deurdat gehoorskerpte en tonale geheue, wat hy as aangeleerde vaardighede eerder as eenlegseienskappe sien, getoets word. Die toets bevat wel elemente van 'n prestasietoets (byvoorbeeld audio-visuele onderskeiding) maar, soos reeds genoem (vide p. 61, 66, 76, 154) kan die onderskeiding tussen 'n musikale aanleg-en prestasietoets nie altyd gemaklik en eksak gemaak word nie. Verder is toetsse vir gehoorskerpte en toongeheue belangrike bestanddele van verskeie standaard-musikale aanlegtoetse, byvoorbeeld Seashore, Drake en Wing (vide p. 79). Gaston se toets is verder nie bloot op sensoriese waarneming gebaseer nie, maar op persepsie in 'n musikale verband. In dié oopsig is sy benadering nader aan die van Wing en Mursell as aan Seashore.

Volgens Gaston se eie siening is die musikaliteit van die gemiddelde kind nie direk afhanklik van aparte sensoriese vermoëns nie, maar wel van sy interaksie met musikale invloedsfaktore in sy omgewing (1957, p.1). Hy beklemtoon verder dat die hooffaktor in musikaliteit, die begrip van toonverwantskap is en sien melodie as die primêre musikale fenomeen. Melodiese bewussyn en respons is vir hom dus van hoë diagnostiese betekenis.

'n Praktiese voordeel van die toets, is dat dit betreklik kort is, as die vraelysgedeelte buite rekening gelaat word en binne 'n halfuur afgehandel kan word (21 minute werklike toetstyd). Daarby bied die subtoetse voldoende afwisseling, wat bevorderlik is vir volgehoud aandag en belangstelling. Aan die anderkant bestaan die eerste drie subtoetse uit slegs vyf items elk met slegs twee moontlike antwoorde by Subtoets 1 en 3, sodat die betrouwbaarheid van hierdie kort toetse bevraagteken kan word.

Na die vraelysdeel volg die subtoetse in die volgende volgorde: akkoordonder-skeiding - voorkoms van 'n enkeltoon in 'n akkoord, al dan nie (soortgelyk aan Wing); audio-visuele onderskeiding - klank met notasie vergelyk ten opsigte van toonhoogte en ritme (soortgelyk aan Kwalwasser-Dykema en Colwell, Toets II);

melodievoltooiing - ontbrekende toon as hoër of laer as die voorafgaande een voorstel (soortgelyk aan Kwalwasser-Dykema); tonaal-ritmiese geheue - vergelyk herhaling van melodie, verskil aandui ten opsigte van toonhoogte of ritme (soortgelyk aan Drake en Kwalwasser).

5.4.4.4 Skrywer (1971): Musikale belangstellings- en voorkeurtoets (MIP)

Hierdie toets is saamgestel met die doel om 'n algemene idee te verkry van leerlinge se smaak en voorkeur ten opsigte van musiekinstrumente en musieksoorte en die vermoë om instrumente en standaardwerke te kan identifiseer. Ook is beoog om 'n moontlike verband hiervan met musikale aanleg te probeer aantoon. Belangstelling, voorkeur en houding hou naamlik verband met wat Guilford (1959, p.205, 250) as 'motiverings- of hormetiese trekke' van die persoonlikheid sien en wat moontlik met aanleg kan korreleer.

- In die eerste deel van die toets moes leerlinge eerstens hul persoonlike voorkeur weergee en vyf name uit 'n lys van 24 instrumente kies en in voorkeurorde plaas. Tweedens moes tien instrumente geïdentifiseer word na aanleiding van kort voorgespeelde musiekuittreksels. Vier moontlike name is aangegee asook 'n 'Geeneen'-keuse, byvoorbeeld:

A Fluit B Klarinet C Trompet D Klavier E Geeneen.

Derdens moet leerlinge na twaalf musiekuittreksels luister waarin telkens 'n ander instrument voorgestel is. Eie voorkeur ten opsigte van die bepaalde instrument se klank, moes aangedui word op 'n vyfpuntskaal:

Hou niks van Hou effens van Hou van Hou baie van Hou besonder baie van
 1 2 3 4 5

In die tweede deel van die toets moes leerlinge eerstens na 12 uittreksels van verskillende musieksoorte luister en soos by die vorige subtoets, persoonlike voorkeur aandui volgens dieselfde waardeskaal. Die tipes was die volgende: lig-klassiek; populêre mars; opera; 'boeremusiek'; klassieke klaviermusiek; jazz; populêre wals; klassieke simfonie; operette; 'popmusiek'; gewyd; volkswysie.

Tweedens moes die leerlinge probeer om 12 musiekuittreksels te identifiseer. 'n Keuse moes gemaak word uit vier name plus 'n moontlike 'geeneen'-keuse, byvoorbeeld:

A	B	C	D	E
Mars, Aida (Verdi)	Soldatekoor, Faust (Gounod)	Jagterskoor, Freischütz (Weber)	Gaudeamus Igitur	Geeneen

Hier is gepoog om 'n wye verskeidenheid musiektipes in te sluit asook vir verskillende ouderdomsgroepe voorsiening te maak, byvoorbeeld deur die insluiting van 'n kleuterliedjie soos 'Swartskaap' (Baa-baa Black Sheep), 'n kinderliedjie soos Brahms se wiegeliedjie, populêre klassieke melodieë soos Tchaikovsky se Blommewals, populêre wysies soos 'Vaya Con Dios' en gewilde klassieke temas soos die begin van Beethoven se vyfde simfonie.

- Derdens is 16 verskillende uittreksels voorgespeel en moes leerlinge die algemene stemming en karakter probeer onderskei as vrolik of treurig, mars of wals, kalm of onrustig, grappig of ernstig. Op soortgelyke wyse moes 16 uittreksels onderskei word wat betref tempo en intensiteit: stadig, middelmatig of vinnig en sag, matig of hard.

Vir die laaste twee afdelings is norme gekonroleer deur groepe primêre en sekondêre leerlinge sowel as musiekstudente die items te laat beoordeel. Wat die intensiteitsdeel betref, is die klankvoorbeeld verder deur 'n pylregistreerder in desibels gemeet (laboratoriumtoets) om sodoende 'n meer eksakte normskaal te verkry.

5.4.4.5 Vraelyste

- Die eerste vraelys, deur leerlinge self voltooi, bestaan uit tien items met voorkeurvrae oor skoolvakke, klasmusiekaktiwiteite en musiektipes, byvoorbeeld:

Van watter drie skoolvakke hou jy die meeste (1e - 3e keuse)?

Van watter klasmusiekaktiwiteite hou jy die meeste/minste (eerste 3 keuses in elke geval uit lys van twaalf)?

Wat is jou geliefkooste soort musiek (1e - 3e keuse uit 'n lys van 13)?

Dan volg enkele vrae oor beroepskeuse, beroep van vader of voog, gesinsgrootte en voorkeur ten opsigte van vokale teenoor instrumentale musiek.

Die vrae is in die toetshandleiding (vide p.169) gestel en leerlinge moes die antwoorde op 'n antwoordkaart invul.

In aansluiting by die doelstellings hierbo genoem by die toets van Skrywer (vide p.157), was die algemene doel om 'n verdere insig te probeer verkry in die leerling se algemene musieksmaak en houding ten opsigte van musiek in sy daaglikslewe asook om die moontlike verband tussen vak- en beroepskeuse en toetsstellings te probeer aantoon.

- Die tweede vraelys is deur die skool (onderwyser of skoolhoof) voltooi en bestaan uit 8 vrae, benewens ruimtes vir algemene opmerkings, en is op 'n antwoordkaart uiteengesit.

Die vrae dek die volgende onderwerpe: algemene gesondheid en aanpassing, skoolbesoek, huislike agtergrond, IK (waar beskikbaar), sangstem en musicaliteit, en dit moes op 'n driepuntskaal (goed, gemiddeld, swak) of vyfpuntskaal (uitstekend, goed, gemiddeld, swak, baie swak) beantwoord word. Besondere skoolprestasies (kultureel, sosiaal, akademies) moes, waar van toepassing, vermeld word en toepaslike 'karakertrekke' uit die aangegewe lys gemerk word, byvoorbeeld: presies, agtelosig, hardwerkend, lui, senuweeagtig, ontspanne, vol selfvertroue, gebrek aan selfvertroue, ensovoort.

Die algemene doel met die vraelys was om die moontlike verband tussen die genoemde aspekte en die toetsresultate te probeer aantoon. Die evaluering van leerlinge se sangstem en musicaliteit was bedoel om te korreleer met die verskillende toetse en so by te dra tot die geldigheidsbepaling van die onderskeie toetse.

Die twee vraelyste is gedeeltelik gebaseer op vraelyste van Gaston (op sy toetsantwoordblad), die biografiese vraelys uit die 1965 Talentopname van die Nasionale Buro vir Opvoedkundige en Maatskaplike Navorsing en 'n belangstellingsvrae-

lys van Margaret Siemens (1969, p.272).

Vir die skoolvakkeuse (Vraelys 1) is die volgende indeling gebruik:

- 0 : Moedertaal
- 1 : Ander tale
- 2 : Geestes-/sosiale wetenskappe bv. Geskiedenis
- 3 : Natuurwetenskappe, bv. Biologie
- 4 : Wiskunde, Rekenkunde
- 5 : Handelsvakke, bv. Boekhou, Handelsrekenkunde
- 6 : Praktiese/Tegniese/Handvaardigheidsvakke
- 7 : Kunsvakke bv. Kuns, Drama, Ballet
- 8 : Musiekvakke
- 9 : Nie-eksamenvakke bv. Godsdiensonderrig, Liggaamlike opvoeding, Boekopvoeding

Vir die beroepskeuse was die indeling soos volg: (vide Greenleaf, 1955, p.568, 569).

- 0 : Professioneel, Tegnies en verwante werkers
- 1 : Administratiewe en bestuursberoep
- 2 : Klerklike en verwante werkers
- 3 : Verkoopswerkers
- 4 : Diens-/bedieningswerkers, Beskermingsdienste
- 5 : Landbou, Bosbou, Visbedryf, Wildbewaring, ens.
- 6 : Produksie en verwante werkers, Transport-toerustingoperateurs, geskoonde en halfgeskoonde arbeiders
- 7 : Swak of ongeskoonde arbeiders
- 8 : Werkers nie klassifiseerbaar volgens beroep bv. werkloses
- 9 : Persone ekonomies onaktief bv. huisvrou, student, pensionaris, ens.

5.4.4.6 Ander toetse oorweeg*(met benaderde tydsduur - vide p.79)

Seashore Measures of Musical Talents (1960); ongeveer 60 minute.

*) Wegelin-Wolmaranstoetse (MUSAT-S en MUSAT-J) was nog nie uitgegee nie (1971).

Drake Musical Aptitude Tests (1957) 40 of 80 minute (beide vorms).

Wing Standardised Tests of Musical Intelligence (1970) 60 minute.

Gordon Musical Aptitude Profile (1965) 120 minute.

Omdat die oogmerk was om 'n verskeidenheid toetse op één dag by 'n skool af te neem, was die tydsduur en vermoeidheidsfaktor belangrike praktiese oorwegings. Die bogenoemde toetse word vervolgens bespreek en die belangrikste oorwegings nie nie-insluiting aangedui.

- Die Seashoretoets is veral vir die jong kind lank en weens die suwer akoesties-sensoriese aard van die toetsitems oninteressant en selfs vervelig. Volgehoue belangstelling en motivering bly belangrike faktore waarmee rekening gehou moes word. Die betroubaarste subtoetse is waarskynlik die toonhoogte- en toongeheuetoe (Lundin, 1967, p.238, p.241). Dié is egter soortgelyk aan die eerste twee subtoetse in die Bentleytoets en sou onnodige duplisering betekenis.

Van die ses subtoetse bevat slegs die toongeheuesubtoets meervoudigekeuse-items van drie, vier en vyf antwoordkeuses. Al die ander subtoetse, met 'n totaal van 230 uit die 260 items, bevat slegs tweekeusige vrae, sodat die algemene betrouwbaarheid en waarde van die Seashore-toets as meetmiddel, bevraagken kan word.

- Die Draketoets duur ook betreklik lank, veral as beide vorms gebruik word, soos wat Drake aanbeveel waar leerlinge met minder as vyf jaar musikale ervaring, getoets word (vide Drake, 1957, p.5). Die ritme-subtoets is wesentlik 'n tempotoets en, weens die 'onmusikale' aard (metronoomslae), geneig om een-tonig te raak en besondere konsentrasie te vereis sodat dit minder geskik is vir veral die jonger leerling. Die 'musikale geheue'-toets is soortgelyk aan die vierde deel van die Gastontoets sodat insluiting nie geregverdig geblyk het nie. (Een besondere afwykende eienskap verbonden aan die toepassing van die Draketoets is die negatiewe metode van puntetoekenning: toetstellings = aantal foute. Verder is dit vreemd dat norms vir die een subtoets in tabelvorm en vir die ander subtoets in die vorm van grafiese gegee word. Gaston maak in sy toetshandleiding ook van beide metodes gebruik; hierop word later weer teruggekom - vide p.217).

- Wing se toets is een van die standaard-musikale aanlegtoetse wat wye erkenning geniet en sterk oorweeg is vir insluiting in die huidige toetsprogram. Daarom is dit van 'n Afrikaanse klankbaan voorsien en uitvoerig vooraf uitgetoets - sommige van hierdie eksperimente word hieronder bespreek.

Die toets is ongelukkig weer eens taamlik lank en vermoeiend, veral vir die jonger leerling. So is tydens 'n eksperimentele toepassing van die toetse gevind dat van 'n groep van twintig leerlinge met ouderdom wat wissel van 8 tot 14 jaar, twee van die agtjariges nie die toets voltooi het nie, waarskynlik weens vermoeidheid of gebrek aan volgehoue konsentrasie ('n klavier word deurgaans gebruik).

Subtoets 1 is feitlik identies aan Bentley se derde subtoets (akkoordanalise). Daarby toon Subtoets 2 'n mate van ooreenkoms met Gaston se eerste subtoets (herkenning van 'n enkeltoon in 'n akkoord), terwyl Subtoets 3 (toongeheue) nou ooreenkom met Bentley se tweede subtoets. Die laaste vier subtoetse meet onder andere voorkeur en smaak en is vir die jong leerling veeleisend - ten spyte daarvan dat Wing die aantal items in latere uitgawes verminder het van 20 tot 14 elk (vanaf die 1957-uitgawe). Musikale ondervinding en opleiding speel waarskynlik by hierdie laaste toetse 'n aansienlike rol en maak dit dus minder geskik vir die huidige doel (vide p.164). Dit was al voorgestel dat die laaste 4 toetse, veral vir die toets van jonger leerlinge, weggelaat word (vide Wing, 1970, p.85). Die totale toetstyd word hierdeur van 60 tot 25 minute verkort.

5.4.4.7 Vergelyking van Wingtoets met toetse van Gaston, Colwell, Bentley en Skrywer

Ten einde vas te stel wat die moontlike verband is tussen Wing se toets en die ander toetse in die toetsprogram, is 'n aantal eksperimente tydens die voorbereidende ondersoek (vide p.134) met groepe onderwysstudente wat musiek as spesialiseringsvak neem uitgevoer en interkorrelasies bereken.

Eksperiment 1: Wing-, Bentley- en Gastontoets (N = 11)

Volgens Tabel 5.35 is korrelasies tussen die Wing- en Bentleytoets meestal belduidend wat die totaal- en subtotaaltellings betref en kan dit afgelei word dat

die toetse in 'n redelike mate oorvleuel wat betref die terrein wat gedek word. Van die subtoetse is dit egter slegs Subtoets 2 en 6 van Wing wat 'n redelike aantal beduidende korrelasies met die Bentleysubtoetse toon.

By die vraelysgedeelte van die Gastontoets word byna deurgaans 'n negatiewe korrelasie met die Wingtoets gevind, sodat dit wil voorkom of huislike agtergrond, voorkeur en belangstelling wat musiek betref, weinig verband hou met die Wingtoets. Verder kom etlike negatiewe korrelasies by die Gaston-subtoetse voor hoewel die totaal- en subtoetstotaalstellings meestal beduidend met die Wingtoets korreleer. Daar moet natuurlik in ag geneem word dat die eksperimentele groep baie klein was sodat afleidings kwalik afdoende kan wees.

TABEL 5.35

KORRELASIES TUSSEN WING- EN TOETSE VAN BENTLEY EN GASTON (N = 11)

WING	BENTLEY					Totaal I-IV	GASTON					Totaal IV
	I	II	III	IV	Vraelys		I	II	III	IV		
1	+301	103	499	289	219	-491	048	228	416	-008	313	
2	273	735*	660**	525	737*	246	-022	-146	718**	513	648**	
3	327	577	440	238	539	-005	271	-270	307	411	384	
Totaal 1-3	532	641**	716**	474	668**	-327	128	-089	649**	417	605**	
4	176	204	766*	519	548	-404	-063	-178	463	300	342	
5	-104	025	539	003	183	-082	-291	-044	137	-023	-041	
6	688**	534	404	816*	760*	103	-081	-405	492	555	431	
7	556	527	130	615**	564	-249	-100	-162	507	822*	754*	
Totaal 4-7	524	495	689**	775*	795*	-232	-171	-315	611**	628**	566	
Totaal 1-7	361	609**	754*	673**	788*	-299	-026	-219	676**	563	629**	

+ = desimale tekens weggelaat

* = beduidend op 0,01-peil ($r \geq 0,735$)

** = beduidend op 0,05-peil ($r \geq 0,602$)

Eksperiment 2: Wing-, Colwell- en Skrywertoets (N = 30)

Veral wat die eerste drie subtoetse en totale van die Wingtoets betref, is korrelasies met die Colwelltoets beduidend, volgens Tabel 5.36, sodat daar 'n duidelike ooreenkoms is in die funksies wat gemeet word deur hierdie toetse. Hieruit kan ook afgelei word dat die betrokke Colwelltoets nie eksklusief as prestasietoets beskou moet word nie maar, minstens gedeeltelik elemente van 'n aanlegtoets bevat. Dit bevestig die stelling wat voorheen gemaak is (vide p.156) dat 'n musi-

kale aanleg- en prestasietoets nie altyd duidelik onderskeibaar is nie.

Skrywer se toets toon, met die uitsondering van die eerste subtoets en die totaal, geen besondere verband met die Wingtoets nie. Die voorkeuraspek by Wing (Subtoets 4 tot 7) dek klaarblyklik 'n heel ander terrein as Skrywer se toets, volgens hierdie tabel van korrelasies. Wing se genoemde subtoetse vereis by uitstek besondere musikale sensitiwiteit, aanvoeling en smaak - dinge wat kwaliëk van veral die jong leerling verwag kan word (vide Lundin, 1967, p.254).

TABEL 5.36

KORRELASIES TUSSEN WING- EN TOETSE VAN COLWELL EN SKRYWER (N = 30)

WING	COLWELL					SKRYWER					Totaal
	I	II	III	I-III	Totaal	I	II	III	IV	V	
1	+399**	232	328	393**	544*	420**	218	184	058	579*	
2	371**	232	305	374**	260	-085	034	143	188	145	
3	441**	353	308	468*	153	220	103	198	-127	247	
Totaal 1-3	573*	389**	444**	585*	436**	246	161	248	055	440**	
4	074	004	-080	007	075	-121	304	-123	-159	018	
5	224	414**	541*	485*	451**	208	056	253	038	349	
6	087	358	117	262	239	137	101	-001	-168	178	
7	245	200	079	233	-049	150	245	180	-302	144	
Totaal 4-7	246	407**	265	402**	298	157	229	119	-236	280	
Totaal 1-7	508*	470*	432**	597*	445**	245	226	227	-086	438**	

+ = desimale tekens weggelaat

* = beduidend op 0,01-peil ($r \geq 0,463$)

** = beduidend op 0,05-peil ($r \geq 0,361$)

Vergelyking van Wing- en Colwelltoets en vakprestasies

Die Wing-toetstellings is ook gekorreleer met dieselfde studente se klaspunte in verskillende teoretiese en praktiese musiekvakke.

Uit Tabel 5.37 blyk dit dat slegs enkele beduidende korrelasies voorkom. Die Wingtoets is hiervolgens dan blykbaar minder geskik as prestasietoets. Ter vergelyking is dieselfde studente se prestasies in die Colwelltoets ook met hierdie vakpuntetellings gekorreleer. Hier word 'n groter getal beduidende korrelasies aangetref sodat aangeleid kan word dat ervaring en opleiding in musiek moontlik by die Colwelltoets 'n groter rol speel as in die geval van die Wingtoets,

veral in die lig van die feit dat die toetsgroep uit studente bestaan wat reeds minstens vyf jaar musiekopleiding gehad het.

TABEL 5.37

KORRELASIES TUSSEN WING- EN COLWELLTOETS EN VAKPRESTASIES

WING	Gehoor en bladlees	Klas= musiek	Algemene musiek= kennis	Blok= fluit	Instru= ment I	Instru= ment II	Totaal
1	+237	-128	-072	508*	145	287	227
2	217	195	020	201	154	318	262
3	104	179	068	276	015	154	182
Totaal 1-3	260	132	013	453**	145	357	317
4	146	-041	209	-062	312	-142	165
5	305	035	-009	446**	112	340	301
6	447**	465*	242	111	124	140	417**
7	129	037	236	262	365**	041	293
Totaal 4-7	433**	234	277	300	352	161	485*
Totaal 1-7	398**	209	153	456**	280	321	463*
COLWELL							
I	297	168	152	520*	492*	437**	506*
II	542*	172	229	305	493*	412**	575*
III	457**	258	248	533*	244	393**	545*
Totaal I-III	555*	243	264	548*	543*	525*	688*

+ = desimale tekens weggelaat

* = beduidend op 0,01-peil ($r \geq 0,463$) ** = beduidend op 0,05-peil ($r \geq 0,361$)

- Gordon se omvangryke toets geniet wye erkenning en is waarskynlik een van die besefdeerde musikale aanlegtoetse tans beskikbaar.

Die toets is van 'n Afrikaanse klankbaan voorsien en volledig by 'n primêre skool op 60 leerlinge van standerd 2 tot 5 toegepas met die oog op moontlike insluiting in die finale toetsprogram (vide 'Eksperiment 3' hieronder). Daar is egter bevind dat die toets te lank duur en te veeleisend is om saam met ander toetse op één dag by 'n skool toe te pas. Gordon beveel dan ook self aan dat die toets in drie sessies op drie verskillende dae afgeneem word (1965, p.26).

Die groot aantal items waaruit elke subtoets bestaan (30 tot 40), veroorsaak by die jonger leerling 'n verlies aan konsentrasie en volgehoue motivering. So

is gevind dat twee van die sestig leerlinge na ongeveer die helfte van die toets afgeneem is, gekla het van hoofpyn en onwillig was om met die toetsprogram voort te gaan, ten spyte van die feit dat voldoende rusponses tussen die verskillende subtoetse gegee is.

Verdere moontlike besware teen die Gordontoets, is dat al die subtoetse te soortgelyk in opset is met te min variasie. Deurgaans word van snaartoon (viool en tjello) gebruik gemaak met die antwoorde stereotiep as 'dieselfde/verskil=lend' (160 items) of 'n keuse tussen '1 of 2' (90 items). Afgesien van die moontlike '?'-'respons is daar deurgaans slegs twee moontlike antwoordkeuses. Dit laat die vraag ontstaan of die 'raai'-moontlikheid nie 'n negatiewe uitwerking op die betrouwbaarheid en geldigheid van die toets uitoefen nie (dieselde beswaar is hierbo teen die Seashoretoets genoem - vide p.161).

Eksperiment 3: Gordon, Bentley, Colwell, Gaston en Skrywer se toetse (N = 60)

Om die moontlike verband tussen die Gordontoets en die ander toetse in die toetsprogram te probeer vasstel, is dié toets saam met die Bentley-, Colwell-, Gaston- en Skrywer se toets op gemelde groep van 60 leerlinge toegepas en korrelasies tussen toetstellings bereken. Die resultate word in Tabel 5.38 weer gegee.

Wat die toetstotale betref, word deurgaans hoogs beduidende korrelasies tussen Gordon se toets en die ander toetse gevind. Dit wissel tussen 0,408 en 0,807. Die korrelasies tussen die Gordontoets se totaal en die ander toetse se subtoetse, is deurgaans beduidend tot hoogs beduidend en wissel vanaf 0,271 tot 0,749. Onbeduidende korrelasies kom in enkele gevalle voor en dan meestal by Gordon se sewende subtoets ('Styl'). Gaston se tweede subtoets (audio-visuele onderskeiding) is die enkele subtoets uit die volle toetsreeks wat weinig beduidende korrelasies met die Gordontoets toon.

Uit die voorafgaande bespreking kan dit afgelei word dat daar 'n duidelike verband tussen die Gordontoets en die ander toetse bestaan, met 'n redelik sterk mate van oorvleueling in die funksies wat gemeet word. Hierdie feit, tesame met die bovenoemde punte van kritiek, was die belangrikste rede waarom die toets

nie in die finale toetsprogram ingesluit is nie, hoewel die nodige antwoordblaais en uitgebreide Afrikaanse klankbaan volledig voorberei is.

TABEL 5.38

KORRELASIES TUSSEN GORDON EN TOETSE VAN BENTLEY, COLWELL, GASTON EN SKRYWER
(N = 60)

GORDON	TOETSTOTALE				SUBTOETSE						
	Bentley	Colwell	Gaston	Skrywer	I	II	III	IV	I	II	III
Totaal	+716*	743*	761*	807*	594*	591*	530*	568*	627*	420*	517*
Totaal T	642*	669*	683*	730*	493*	597*	459*	495*	587*	352*	454*
Totaal R	643*	724*	669*	759*	546*	499*	509*	469*	558*	447*	543*
Totaal S	652*	598*	727*	675*	570*	501*	448*	583*	533*	322**	390*
++T1	663*	676*	703*	671*	518*	614*	512*	488*	628*	318**	442*
T2	456*	494*	496*	629*	337*	427*	278**	380*	387*	315**	357*
R1	657*	693*	668*	762*	537*	499*	521*	489*	554*	402*	513*
R2	540*	652*	570*	646*	474*	430*	430*	388*	490*	426*	489*
S1	517*	544*	604*	550*	461*	330*	403*	430*	470*	251	401*
S2	624*	483*	660*	579*	569*	562*	312**	554*	462*	303**	249
S3	411*	408*	465*	486*	326**	284**	382*	383*	361*	208	277**
SUBTOETSE											
GORDON	Vrae=	GASTON				SKRYWER					
	lys	I	II	III	IV	I	II	III	IV	V	
Totaal	+507*	384*	271**	499*	749*	579*	547*	534*	577*	455*	
Totaal T	456*	422*	248	492*	615*	558*	510*	464*	490*	403*	
Totaal R	522*	318**	245	364*	691*	537*	506*	476*	592*	447*	
Totaal S	365*	298**	278**	486*	738*	471*	452*	492*	443*	373*	
++T1	460*	415*	226	530*	649*	487*	506*	413*	454*	373*	
T2	350*	327**	223	322**	427*	513*	391*	412*	429*	348*	
R1	534*	277**	289**	324**	703*	573*	473*	512*	605*	396*	
R2	433*	303**	163	358*	579*	423*	460*	374*	492*	436*	
S1	317**	219	278**	346*	625*	442*	307**	417*	344*	300**	
S2	368*	338*	234	474*	633*	383*	389*	435*	368*	334*	
S3	185	157	142	329**	501*	297**	402*	320**	344*	250	

+ = desimale tekens weggelaat * = beduidend op 0,01-peil ($r \geq 0,330$)
** = beduidend op 0,05-peil ($r \geq 0,254$)

++ = T: Toonvoorstelling	T1 : Melodie	T2 : Harmonie
R: Ritmevoorstelling	R1 : Tempo	R2 : Metrum
S: Musikale sensitiwiteit	S1 : Frasering	S2 : Balans
S3 : Styl		

5.4.5 Statistiese tegnieke (vide hierbo, p.140 e.v.)

Die volgende is die belangrikste tipe statistiese bewerkings wat uitgevoer is: korrelasiekoëffisiëntberekening, itemontleding, frekwensieberekening, normbepaling en faktorontleding. Die rekenaarprogramme wat uitgevoer is, behels die volgende:

- Korrelasies

Produkmoment-korrelasiekoëffisiënte tussen tellings van die veranderlikes sowel as die gemiddeldes en standaardafwykings is bereken.

- Itemontleding

Benewens die gemiddeldes, standaardafwyking en metingsfout, is betroubaarheid volgens die Kuder-Richardsonformule-20 bereken. Dan is vir elke toetsitem 'n responsanalise uitgevoer, dit wil sê, vir elke moontlike respons (afleier) is die persentasie leerlinge wat die afleier gekies het, die toetsgemiddeld van die leerlinge wat die afleier gekies het, die toetsgemiddeld van die leerlinge wat die item korrek en die wat die item verkeerd beantwoord het, bepaal. Hieruit is dan die punt-biseriale-korrelasie met die toetstotaal bepaal. Ook is die produk van die itemstandaardafwyking en die diskriminasiekoëffisiënt bereken.

- Frekwensies

Hier is die frekwensieverdeling, persentasiefrekvensie, kumulatiewe frekwensie, kumulatiewe persentasiefrekvensie, gemiddelde, standaardafwyking, skeefheid en kurtose bereken.

- Norms

Benewens die berekenings onder 'Frekwensies' genoem, is normpunte bepaal volgens die C-skaal (vide p.128 hierbo), betroubaarheid volgens die Kuder-Richardsonformule-21 bereken en die Kolmogorov-Smirnovtoets vir normaalverdeling uitgevoer.

- Faktorontleding

Vanuit 'n gegewe korrelasiematriks is die minimum aantal faktore wat onderliggend aan die veranderlikes is, bepaal volgens die 'herhaalde hooffaktoranalise'-metode. Aanvangskommunaliteit is geskat deur die gebruik van gekwadreerde meer-voudige korrelasies van elke veranderlike tot elke ander veranderlike. Ongeroeteerde faktorladings, met die variansie soos verklaar deur elke faktor, is bereken en hierdie faktorpatroon volgens Kaiser se Varimaxmetode na 'n eenvoudige struktuur geroteer*).

5.4.6 Metodologiese oorwegings (vide p.146 hierbo)

Op soortgelyke wyse as by die Bentleytoets, is die Colwell- en Gastontoets geredigeer en van 'n Afrikaanse klankbaan voorsien sodat die toetse in die taalmedium van die betrokke skool toegepas kon word. In die geval van paralelmediumskole is die taal wat oorwegend deur die leerlinge gepraat word (volgens die personeel se oordeel), as medium gebruik.

Met die oog op die dataverwerkingsprosedures is nuwe toetsantwoordblaais vir al die toetse ontwerp en gedruk sodat alle versamelde inligting deur 'n IBM 1230 optiese merkleser nagesien en masjinaal op kaarte gepons kon word. Die vraelysgegewens is vanaf die antwoordkaarte op datavorms oorgedra en van daar deur die 1230 soos hierbo, gehanteer. Die jongste leerlinge (graad 2 en standerd 1) het hul antwoorde op die Bentleytoets op dieselfde antwoordblaais geskryf as wat met die voorondersoek gebruik is. Dit was naamlik ietwat twyfelagtig of hierdie jong leerlinge die koderingswyse op die IBM-toetsantwoordblaais in die beperkte tyd sou kon baasraak. Hierdie toetsdata is op 'n latere stadium op die laasgenoemde antwoordblaais oorgedra. Die biografiese gegevens is deur die betrokke klasonderwysers(esse) op hierdie vorms ingevul.

'n Toetshandleiding met volledige instruksies wat aanwysings vir die korrekte koderingswyse en algemene voltooiing van die antwoordblaais bevat sowel as instruksies vir die toetse van Gaston en Skrywer, is opgestel. Ook die vraelys wat die leerlinge op 'n antwoordkaart moes beantwoord, is daarin opgeneem (vide p.158).

*) Die algemene beginsels van faktorontleding word in die volgende hoofstuk bespreek (vide p.517), waar dit in direkte verband staan met, en voortvloeit uit die resultate van die ondersoek (veral wat betref interkorrelasies tussen toetse en subtoetse).

Eenvormigheid in toetsprosedure en -omstandighede is sover moontlik probeer verkry, onder andere deurdat alle toetse deur een toetsafnemer (skrywer) toegepas is. Die algemene toetssituasie, aanwysings ensovoort is dus so gekontroleerd en konstant moontlik gehou met besondere aandag aan die skep van 'n aangename atmosfeer, algemene motivering en aanmoediging. Verder is dieselfde klankapparaat en klankopnames by alle skole gebruik. Met die uitsondering van twee skole, is alle toetssessies tussen ongeveer 08h30 en 12h00 gehou. In feitlik alle gevalle is een of twee personeellede beskikbaar gestel om met die organisasie en algemene toesig behulpsaam te wees.

Om die moontlike probleem van oordrag en beïnvloeding tussen toetse sover moontlik te neutraliseer, is die volgorde waarin die toetse afgeneem is, op 'n vaste rotasiebasis voortdurend van skool na skool gevarieer. Om moontlike vermoeidheid en gebrek aan konsentrasie sover as moontlik uit te skakel, is voldoende pauses tussen toetse gegee. Verder is die andersoortige toets van die skrywer, waarin verskillende musiekuittreksels voorkom, in gedeeltes tussen die ander toetse afgeneem om verdere afwisseling en afleiding te bied. Hierdeur is daarin geslaag om die leerlinge se belangstelling en volle aandag end-uit te behou.

Toets-hertoets-betrouwbaarheid is tydens die voorbereidende ondersoek intensief nagegaan. 'n Verdere toepassing van hierdie metode is hier uitgevoer deur vergelyking van die toetsprestasies van 127 leerlinge wat tydens die voorondersoek sowel as die hoofondersoek deel van die ondersoekgroep uitgemaak het en dus twee keer die Bentleytoets afgelê het.

5.4.7 Samevatting

Hierdie hoofstuk het die fokus geplaas op die verskillende stappe by die beplanning en uitvoering van die navorsingsprojek. Waar die eerste vier hoofstukke te doen gehad het met die algemene identifisering, definisie en afbakening van die navorsingsterrein en -probleem, is in hierdie hoofstuk die ontwerp van die navorsingsprojek behandel, die keuse van toepaslike statistiese en metodologiese prosedures en eksperimentele manipulasies bespreek en enkele probleemgebiede ten opsigte van die weergawe van die navorsingsresultate aangesny ter voorbereiding van die verslag van resultate en gevolgtrekkings wat hierna volg.