

Psigometriese eienskappe van die ESSI-lees- en speltoets

K.G.F. Esterhuysen,* R.B.I. Beukes en P.M. Heyns

Departement Sielkunde, Universiteit van die Vrystaat, Posbus 339, Bloemfontein, 9300 Suid-Afrika
esterkg@hum.uovs.ac.za

* Aan wie korrespondensie gerig moet word

In drawing up the ESSI reading and spelling test, four phases were followed. In an earlier article, the development work that took place during the first three phases was discussed. This article focuses on the last phase during which the psychometric characteristics of the ESSI reading and spelling test were investigated. Firstly, the skewedness of the test scores on all the tests (vocabulary lists) was investigated. The results showed that there was a reasonable measure of symmetrical distribution of scores on all the Afrikaans and the English tests. The kurtosis of the scores was also investigated, and with some exceptions, it was clear that the scores had a normal distribution. Attention was also devoted to the standard error of measurement for the tests. Not one of the standard errors of measurement values for the tests was larger than 2. Information on the reliabilities of the tests is also included. Various reliability indices were calculated and all of them exhibited high levels of reliability. Concerning the validity of the tests, attention was devoted to content and criterion-related validity of the tests. It was possible to ensure that the requirements of content validity were met by requesting experienced persons to identify the words for the test. Information on the predictive validity of the tests indicated that all the tests have sound predictive validity.

Inleiding

Die ESSI-lees- en speltoets is saamgestel uit 'n reeks van 28 woorde-lyste vir Afrikaans- en Engelsprekende leerders in grade 1 tot 7. Daar het 'n behoefte ontstaan vir die ontwikkeling van so 'n meetmiddel om die destydse UK-lees- en speltoets wat reeds in 1944 verskyn het en verouderd was, te vervang. Die fases wat gevolg is ten einde die ESSI-lees- en speltoets te ontwikkel is soos volg:

Fase 1: Identifisering van voorlopige woordelyste;

Fase 2: Itemontleding en -seleksie;

Fase 3: Normbepaling;

Fase 4: Betroubaarheid- en geldigheidstudies.

In 'n vorige artikel is die ontwikkelingswerk wat gedurende fases 1, 2 en 3 plaasgevind het aangetoon en bespreek. In hierdie artikel val die fokus op fase 4 waartydens die psigometriese eienskappe van die ESSI-lees- en speltoets bespreek sal word. Benewens die betroubaarheid en geldigheid van die toetse sal ook na die skeefheid, kurtose en standaardmetingsfoute van die verskillende grade se toetse (woorde-lyste) verwys word.

Die skeefheid, kurtose en standaardmetingsfoute word vervolgens vir die Afrikaanse en Engelse lees- en speltoetse afsonderlik in Tabel 1 en Tabel 2 gegee. Hierdie inligting het betrekking op die eerste toepassing van die finale toetse gedurende die eerste kwartaal van 1996.

Skeefheid en kurtose

Skeefheid is 'n statistiek wat benodig word om die mate te bepaal waarin die distribusie van 'n groep toetstellings 'n normaalkurwe benader, aangesien dit die afwyking vanaf simmetrie meet. Die waardes varieër gewoonlik tussen -3 en $+3$. Indien die waarde gelyk aan nul is, is die distribusie van tellings simmetries om die gemiddelde. 'n Positiewe waarde dui daarop dat die meerderheid van toetslinge 'n telling kleiner as die gemiddelde en 'n negatiewe waarde dat die meerderheid 'n telling groter as die gemiddelde behaal het. Wanneer die waarde hoog en positief is, word die toets as moeilik beskou. 'n Hoë, negatiewe waarde dui weer op 'n maklike toets.

Ten opsigte van die Afrikaanse toetse (kyk Tabel 1) kan die afleiding gemaak word dat die tellings op die graad 2-leestoets, graad 3-speltoets, graad 4-lees- en speltoets en graad 7-speltoets 'n simmetriese distribusie toon. Al hierdie toetse se skeefheidskoeffisiënte is ongeveer nul. Vanuit hierdie tabel blyk verder dat, met die uitsondering van graad 2 se speltoets (0.22), al die ander toetse klein negatiewe skeefheidskoeffisiënte (wissel tussen -0.18 en -0.58) het. Dit wil dus voorkom of daar 'n redelike mate van simmetriese distribusie van tellings in al die Afrikaanse toetse voorkom.

Wat die Engelse toetse betref (kyk Tabel 2), toon graad 1 en

graad 3 se leestoetse, asook graad 6 se speltoets, 'n simmetriese distribusie, aangesien hierdie drie toetse se skeefheidskoeffisiënte ongeveer nul is. Beide die lees- en speltoets van graad 2, asook graad 3 en graad 7 se speltoetse, toon positiewe koeffisiënte. Hierdie skeefheidskoeffisiënte wissel tussen 0.31 en 0.81. Al die ander toetse toon klein negatiewe (wissel tussen -0.14 en -0.53) koeffisiënte. Ook hier blyk dit dat daar 'n redelike mate van simmetriese distribusie van tellings by die Engelse toetse voorkom. In vergelyking met die ander toetse wil dit tog voorkom of graad 2 se lees- en speltoetse effens aan die moei-like kant is. Die routellings is egter vir al die toetse genormaliseer voordat die norms bereken is.

Kurtose handel oor die relatiewe platheid of spitsheid van die verspreidingskurwe. 'n Normaalkurwe sal 'n kurtose van nul hê, terwyl 'n positiewe kurtosewaarde op 'n spitser kurwe dui. 'n Negatiewe kurtosewaarde dui weer daarop dat die distribusie platter as die normaalkurwe is.

Al die Afrikaanse sowel as Engelse toetse (kyk Tabelle 1 en 2) toon 'n negatiewe kurtosewaarde. Met die uitsondering van enkele Engelse toetse (graad 3-leestoets, graad 4 en 5 se speltoetse) is hierdie waardes klein en kan aanvaar word dat die tellings oor die algemeen normaal verdeel is.

Standaardmetingsfoute

Die standaardmetingsfoute van die Afrikaanse toetse (Tabel 1) wissel tussen 1.16 (graad 1-leestoets) en 1.89 (graad 7-leestoets). Dit beteken dat indien 'n graad 1-leerder 'n telling van 12 uit 20 op die leestoets sou behaal, met 'n waarskynlikheid van 0.95, gesê kan word dat die leerder se ware telling nie buite die omvang van $12 - 2.27 = 9.73$ en $12 + 2.27 = 14.27$ sal val nie. Die waarde 2.27 is verkry deur 1.16 met 1.96 te vermenigvuldig. Indien egter slegs met een standaardafwykingseenheid weerskante van die routelling gewerk word, dan beteken dit dat die grense vir dieselfde graad 1-leestoets $12 - 1.16 = 10.84$ en $12 + 1.16 = 13.16$ is, maar nou kan slegs met 'n waarskynlikheid van 0.68 gesê word dat die ware telling binne hierdie grense sal val.

Vir die Engelse toetse wissel die standaardmetingsfoute (Tabel 2) tussen 1.33 (graad 1-leestoets) en 1.96 (graad 4-speltoets). Dit beteken dat indien 'n graad 1-leerder 'n telling van 10 uit 20 op die leestoets sou behaal, met 'n waarskynlikheid van 0.95 gesê kan word dat die leerder se ware telling nie buite die omvang van $10 - 2.61 = 7.39$ en $10 + 2.61 = 12.61$ sal val nie. Die waarde 2.61 is verkry deur 1.33 met 1.96 te vermenigvuldig. Indien egter slegs met een standaardafwykingseenheid weerskante van die routelling gewerk word, dan beteken dit dat die grense vir dieselfde graad 1-leestoets $10 - 1.33 =$

Tabel 1 Statistiese eienskappe van die Afrikaanse lees- en speltoetse

Graad	Tipe toets	N	Getal woorde	Standaard-metingsfout	Skeefheid	Kurtose
Een	Spel	537	15	1.32	-0.23	-0.48
	Lees	534	15	1.16	-0.48	-0.62
Twee	Spel	456	20	1.56	0.22	-0.55
	Lees	450	20	1.72	-0.06	-0.62
Drie	Spel	466	20	1.69	0.06	-0.57
	Lees	470	20	1.68	-0.58	-0.06
Vier	Spel	467	20	1.77	0.025	-0.87
	Lees	454	20	1.56	-0.07	-0.30
Vyf	Spel	424	20	1.80	-0.18	-0.53
	Lees	392	20	1.69	-0.25	-0.61
Ses	Spel	434	20	1.66	-0.36	-0.52
	Lees	400	20	1.65	-0.29	-0.69
Sewe	Spel	433	20	1.55	0.06	-0.72
	Lees	394	20	1.89	-0.37	-0.72

Tabel 2 Statistiese eienskappe van die Engelse lees- en speltoetse

Graad	Tipe toets	N	Getal woorde	Standaard-metingsfout	Skeefheid	Kurtose
Een	Spel	284	15	1.36	-0.53	-0.12
	Lees	285	15	1.33	-0.08	-0.69
Twee	Spel	322	20	1.71	0.81	-0.05
	Lees	242	20	1.70	0.42	-0.62
Drie	Spel	300	20	1.70	0.35	-0.72
	Lees	261	20	1.78	-0.03	-1.22
Vier	Spel	307	20	1.96	-0.14	-1.04
	Lees	272	20	1.56	-0.41	-0.46
Vyf	Spel	316	20	1.74	-0.17	-1.05
	Lees	280	20	1.63	-0.34	-0.68
Ses	Spel	315	20	1.85	-0.01	-0.80
	Lees	280	20	1.78	-0.53	-0.38
Sewe	Spel	311	20	1.61	0.31	-0.77
	Lees	274	20	1.72	-0.30	-0.23

8.67 en $10 + 1.33 = 11.33$ is, maar nou kan slegs met 'n waarskynlikheid van 0.68 gesê word dat die ware telling binne hierdie grense sal val. Nie een van die toetse se standaardmetingsfoute is groter as 2 nie.

Betroubaarheid

Ten opsigte van die betroubaarheidsindekse is die toets-hertoetsbetroubaarheid vir elk van die woordelyste bereken. Verder is 'n skatting van die toetse se parallellevorm-betroubaarheid verkry deur twee verskillende prosedures te benut, naamlik die halveringsmetode en die α -koëffisiënte. 'n Bespreking van die betroubaarheidskoëffisiënte volg nou.

Parallellevorm-betroubaarheid

1. Alfa-koëffisiënt

Aangesien die lees- en speltoetse slegs reg of verkeerd gemerk kan word, word hier met binêre items gewerk. In hierdie geval is die alfa-koëffisiënt gelyk aan die Kuder-Richardson-formule 20 (Huysamen, 1996). Daar is met behulp van die SPSS-rekenaarprogrammatuur die alfa-koëffisiënt vir elke subtoets bereken en die inligting verskyn in Tabel 3.

By gestandaardiseerde toetse is 'n betroubaarheidskoëffisiënt van 0.85 en hoër verkieslik. Uit Tabel 3 blyk dit dat daar by uitsondering toetse met 'n koëffisiënt laer as 0.85 is. Hierdie koëffisiënte is almal egter nog hoër as 0.80, sodat daar met 'n redelike mate van vertroue aanvaar kan word dat die toetse intern konsekwent meet.

Tabel 3 Parallellevorm betroubaarheid van die Afrikaanse en Engelse lees- en speltoetse volgens die Kuder-Richardson 20-formule

Graad	Tipe toets	Afrikaanse toetse		Engelse toetse	
		N	(KR ₂₀) koëffisiënt	N	(KR ₂₀) koëffisiënt
Een	Spel	537	0.850*	284	0.819*
	Lees	534	0.908*	285	0.867*
Twee	Spel	456	0.840*	322	0.875*
	Lees	450	0.857*	242	0.862*
Drie	Spel	466	0.846*	300	0.873*
	Lees	470	0.852*	261	0.902*
Vier	Spel	467	0.849*	307	0.877*
	Lees	454	0.827*	272	0.890*
Vyf	Spel	424	0.808*	316	0.879*
	Lees	392	0.827*	280	0.868*
Ses	Spel	434	0.827*	315	0.866*
	Lees	400	0.871*	280	0.834*
Sewe	Spel	433	0.867*	311	0.871*
	Lees	394	0.821*	274	0.816*

* $p \leq 0.0001$

2. Halveringsmetode

Die tweede prosedure wat gebruik is om 'n skatting van die verskillende grade se parallellevorm-betroubaarheid te verkry, is die halveringsmetode. Met hierdie metode word elke toets in die helfte verdeel, sodat die items met gelyke nommers die een helfte en die items met ongelyke nommers die ander helfte vorm. Vir elke toetsling word 'n totaalstelling op die twee helftes bereken en gekorreleer. Aangesien die toetse op hierdie wyse met die helfte verkort word, lewer die Spearman-Brown-formule 'n beter skatting van die volledige toetse se betroubaarheid (Huysamen, 1996). Hierdie formule is benut en die resultate word in Tabel 4 verskaf.

Die betroubaarheidsresultate wat met behulp van hierdie metode bereken is, blyk aanvaarbaar tot redelik goed te wees.

Toets-hertoetsbetroubaarheid

Toets-hertoetsbetroubaarheid word ondersoek deur dieselfde toets (items) by twee verskillende geleenthede op dieselfde groep toetslinge wat verteenwoordigend is van die populasie waarvoor die toets opgestel is, toe te pas en die korrelasie tussen hierdie twee stelle tellings te bereken. In hierdie ondersoek is dieselfde leerders in die eerste en die vierde kwartaal van 1996 met dieselfde woordelyste getoets. (Die graad 1-leerders is egter in die derde en vierde kwartale getoets, sodat die tydsverloop in daardie geval korter was). Die verkreeë korrelasie staan bekend as die stabiliteitskoëffisiënt.

Aangesien betroubaarheid verwys na die konsekwentheid waarmee 'n toets meet, ongeag die geleentheid wanneer dit toegepas word, sou verwag kon word dat die attribute wat gemeet is, sou verander, aangesien ongeveer agt maande tussen die twee toepassings verloop het (graad 2 tot 7). Leerders se lees- en spelvermoe behoort oor 'n tydperk van agt maande te verbeter, sodat daar redelikerwys aanvaar kan word dat hulle oor die algemeen 'n beter punt tydens die tweede in ver-

Tabel 4 Halveringsbetroubaarheidsresultate van die Afrikaanse en Engelse lees- en speltoetse vir elke graad afsonderlik bereken

Graad	Tipe toets	Afrikaanse toetse		Engelse toetse	
		N	Koëffisiënt	N	Koëffisiënt
Een	Spel	536	0.871*	283	0.862*
	Lees	519	0.891*	283	0.904*
Twee	Spel	455	0.844*	329	0.872*
	Lees	448	0.863*	240	0.889*
Drie	Spel	465	0.853*	300	0.886*
	Lees	464	0.848*	261	0.937*
Vier	Spel	466	0.828*	307	0.919*
	Lees	443	0.767*	270	0.917*
Vyf	Spel	424	0.845*	316	0.918*
	Lees	387	0.855*	280	0.876*
Ses	Spel	434	0.836*	315	0.878*
	Lees	397	0.865*	279	0.845*
Sewe	Spel	433	0.858*	312	0.889*
	Lees	391	0.841*	272	0.815*

* $p \leq 0.0001$

gelyking met die eerste toepassing sou behaal. In so 'n geval dui dit nie daarop dat die meetinstrument onbetroubaar is nie, maar gaan dit oor die onstabiliteit van die attribuut wat gemeet word. Aangesien die twee toepassings lank uit mekaar plaasgevind het, kan daar met sekerheid aanvaar word dat oefening nie hier 'n rol sou speel nie. Die waarskynlikheid dat die tellings wat tydens die tweede toepassing behaal is, deur die eerste toepassing se tellings beïnvloed is, is dus uiters gering.

Ofskoon dit die geval is, is daar tog besluit om die leerders se tellings soos in die eerste kwartaal behaal met dié wat hulle in die vierde kwartaal behaal het, te korreleer en die resultate vir die Afrikaanse toetse verskyn in Tabel 5 en vir die Engelse toetse in Tabel 6.

Ofskoon daar 'n redelike lang tydsverloop (agt maande) tussen die twee toepassings was, is steeds bemoedigende betroubaarheidskoëffisiënte met hierdie metode verkry. Al die koëffisiënte is op die 0.01 peil beduidend.

Dit blyk dus vanuit die voorafgaande bespreking dat al die toetse oor genoegsame betroubaarheid beskik, sodat dit in die praktyk aangewend kan word.

Geldigheid

In hierdie studie word op die toegepaste gebied gewerk en sal gevolglik slegs aandag aan die inhoudsgeldigheid en kriteriumverwante geldigheid van die meetinstrumente aandag gegee word.

Inhoudsgeldigheid

Die inhoudsgeldigheid van 'n meetinstrument word nie in terme van 'n kwantitatiewe indeks uitgedruk nie en die items behoort deur kundiges op die spesifieke gebied op logiese gronde geëvalueer te word. Die belangrikste oorweging by die evaluering daarvan is of die toetsitems 'n bevredigende steekproef is van die spesifiek gedefinieerde universum van take, in hierdie geval die totaliteit van moontlike lees- en spelwoorde vir 'n gegewe graad.

Om inhoudsgeldigheid te verseker is ondersteuningsonderwysers, asook die hoofde van ondersteuningsonderwys by die verskillende kinderleidingklinieke by die identifikasie van woorde betrek. Hierdie persone was ervare, aangesien hulle minstens 8 jaar in die beroep moes wees. Hulle moes die woorde identifiseer op grond van die kennis wat hulle van die verskillende grade se sillabusse, asook hulle ervaring op die gebied van ondersteuningsonderwys het.

Tabel 5 Korrelasiekoëffisiënte vir Afrikaanse lees- en speltoetse soos bereken tussen die eerste en tweede toepassing

Graad	N	Eerste toepassing	Tweede toepassing	
			Spel	Lees
Een	508	Spel	0.762*	-
	509	Lees	-	0.670*
Twee	411	Spel	0.748*	-
	405	Lees	-	0.696*
Drie	427	Spel	0.792*	-
	431	Lees	-	0.765*
Vier	435	Spel	0.851*	-
	422	Lees	-	0.702*
Vyf	393	Spel	0.800	-
	362	Lees	-	0.775*
Ses	395	Spel	0.603*	-
	361	Lees	-	0.526*
Sewe	409	Spel	0.868*	-
	370	Lees	-	0.779*

* $p \leq 0.0001$

Tabel 6 Korrelasiekoëffisiënte vir Engelse lees- en speltoetse soos bereken tussen die eerste en tweede toepassing

Graad	N	Eerste toepassing	Tweede toepassing	
			Spel	Lees
Een	277	Spel	0.790*	-
	279	Lees	-	0.849*
Twee	304	Spel	0.749*	-
	227	Lees	-	0.773*
Drie	276	Spel	0.842*	-
	237	Lees	-	0.860*
Vier	296	Spel	0.909*	-
	254	Lees	-	0.853*
Vyf	301	Spel	0.902	-
	261	Lees	-	0.856*
Ses	298	Spel	0.885*	-
	259	Lees	-	0.860*
Sewe	267	Spel	0.913*	-
	261	Lees	-	0.830*

* $p \leq 0.0001$

Nadat die woorde ontvang is, is die woorde wat die meeste in die onderskeie kategorieë (maklik, gemiddeld, moeilik) voorgekom het, in die voorlopige woordelyste opgeneem waarna dit weer aan kundiges op die gebied voorgeleë is. Nadat daar konsensus bereik is, is die woordelyste gereproduseer en aan die skole versprei.

Kriteriumverwante geldigheid

Kriteriumverwante geldigheid sluit voorspellings- sowel as saamvallende geldigheid in. Saamvallende geldigheid is nie in hierdie studie ondersoek nie. Die voorspellingsgeldigheid van die lees- en speltoetse is wel ondersoek. Die wyse waarop dit gedoen is, is dat die leerders gedurende die eerste kwartaal van 1996 (in die geval van graad 1-leerders in die derde kwartaal) met die lees- en speltoetse getoets is

waarna inligting oor hulle prestasie in hul onderskeie skoolvakke aan die einde van 1996 by die skole bekom is. Elke leerder se totale lees- en speltoetstelling is met hul skoolvakpunt gekorreleer. Die skoolvakke se punte is almal na 'n persentasie omgeskakel alvorens die korrelasies gedoen is. Vir die graad 1-, 2- en 3-leerders is inligting vir slegs twee vakke, naamlik eerste taal en wiskunde bekom. Hierdie leerders verwerf slegs 'n simbool per vak. Hierdie simbole moes na 'n persentasiewaarde getransformeer word en inligting wat vanaf die Vrystaatse Onderwysdepartement verkry is, dui die volgende persentasie-intervalle vir die volgende simbole aan.

Die middelpunte van die onderskeie persentasie-intervalle is as die leerders se punte vir die onderskeie simbole geneem. Byvoorbeeld, indien 'n leerder 'n C-simbool in wiskunde behaal het, is 'n persentasiewaarde van 65 aan hom/haar toegeken. Vir die graad 4-, 5- en 6-leerders is punte op ses vakke, naamlik eerste taal, tweede taal, wiskunde, aardrykskunde, geskiedenis en wetenskap bekom. Die graad 7-leerders het 'n gesamentlike punt vir aardrykskunde en geskiedenis, sodat daar vir hulle ten opsigte van slegs vyf vakke punte bekom is. Die graad 4-leerders het slegs persentasies ten opsigte van hul eerste taal en wiskunde verwerf. Op die ander vier vakke is ook net simbole behaal. In hierdie geval is soos in Tabel 7 aangedui, te werk gegaan om 'n persentasiepunt te verkry.

Tabel 7 Persentasie-intervalle en middelpunt van intervalle vir die onderskeie simbole

Simbool	% - interval	Middelpunt van interval
A	80 – 100	90
B	70 – 79	75
C	60 – 69	65
D	50 – 59	55
E	40 – 49	45
F	30 – 39	35

Die leerders betrokke in die ondersoek se punte is, met die uitsondering van enkele skole, aan die begin van 1997 aan die navorser voorsien. Hierdie punte is by die oorspronklike dataset (hulle lees- en spel punte) gevoeg waarna korrelasies met behulp van die SAS-programmatuur bereken is. Hierdie resultate word in Tabelle 8 en 9 vir die twee taalgroepe afsonderlik verskaf.

Die resultate in Tabel 8 en Tabel 9 bevestig die rasionaliteit van die nuwe toetse wat gebaseer is op die veronderstelling dat leerders se vermoë om korrek te kan lees en spel 'n geldige voorspeller vir hulle prestasiepeil in skoolvakke en veral taalvakke is (Lerner, 1993; Mercer, 1997). Hoewel die berekende koëffisiënte dan ook die hoogste vir die taalvakke is, is die koëffisiënte vir al die ander vakke ook beduidend op die 0.01-peil. Hieruit volg baie duidelik dat 'n leerder wat beperkinge ten opsigte van lees- en spelling ervaar, baie beslis ook leerprobleme gaan ervaar.

Nie alleen is al hierdie korrelasiekoëffisiënte statisties beduidend groter as nul nie, maar hulle is met enkele uitsonderings almal hoër as 0.50. Volgens Cohen (1988) verteenwoordig 'n korrelasie van 0.50 'n groot effek en gevolglik is hierdie resultate van groot praktiese waarde.

Dit wil voorkom of lees- en spellingvaardighede ook al hoe belangriker word in die nuwe wiskunde, aangesien dit groter eise aan die leerder se leesvermoë (woordsomme) stel. In onderhoude met onderwysers het dit dan ook geblyk dat wanneer 'woordsomme' aan 'n leerder met leesprobleme, voorgelees word, hy/sy dit wel korrek kan bereken. Indien die leerder egter die som eers selfs moet lees, word probleme ervaar om dit korrek te bereken.

Uit die resultate in Tabelle 8 en 9 is dit duidelik dat die ESSI-lees- en speltoetse oor goeie voorspellingsgeldigheid beskik.

Slot

In hierdie artikel is besonderhede oor die ESSI-lees- en speltoetse se psigometriese eienskappe verskaf. Die vernaamste vereistes rakende

Tabel 8 Korrelasiekoëffisiënte tussen lees-/spel- en vakpunte vir die Afrikaanse groep

Graad	Tipe toets	Korrelasiekoëffisiënte						
		Afr.	Eng.	Wisk.	Aard.	Gesk.	Aard./Gesk.	Wet.
Een	Spel	0.575*		0.544*				
	N	498		498				
	Lees	0.556*		0.474*				
	N	479		479				
Twee	Spel	0.584*		0.529*				
	N	385		385				
	Lees	0.671*		0.555*				
	N	378		378				
Drie	Spel	0.639*		0.570*				
	N	404		404				
	Lees	0.583*		0.490*				
	N	405		405				
Vier	Spel	0.666*	0.554*	0.545*	0.494*	0.554*		0.407*
	N	409	380	409	380	380		380
	Lees	0.537*	0.503*	0.438*	0.391*	0.444*		0.347*
	N	391	362	391	362	362		362
Vyf	Spel	0.671*	0.618*	0.508*	0.559*	0.544*		0.535*
	N	370	365	370	364	363		364
	Lees	0.645*	0.638*	0.459*	0.472*	0.520*		0.534*
	N	337	333	337	333	332		333
Ses	Spel	0.465*	0.461*	0.345*	0.330*	0.337*		0.343*
	N	379	374	378	374	374		374
	Lees	0.419*	0.403*	0.286*	0.255*	0.274*		0.309*
	N	346	341	345	341	341		341
Sewe	Spel	0.652*	0.687*	0.549*			0.647*	0.590*
	N	384	379	379			330	378
	Lees	0.625*	0.640*	0.538*			0.542*	0.592*
	N	341	336	336			289	336

* $p \leq 0.0001$

Tabel 9 Korrelasiekoëffisiënte tussen lees-/spel- en vakpunte vir die Engelse groep

Graad	Tipe toets	Korrelasiekoëffisiënte						
		Afr.	Eng.	Wisk.	Aard.	Gesk.	Aard./Gesk.	Wet.
Een	Spel	0.636*		0.482*				
	N	254		254				
	Lees	0.631*		0.448*				
	N	253		253				
Twee	Spel	0.621*		0.527*				
	N	274		274				
	Lees	0.660*		0.505*				
	N	210		210				
Drie	Spel	0.617*		0.588*				
	N	266		266				
	Lees	0.624*		0.537*				
	N	226		226				
Vier	Spel	0.575*	0.691*	0.541*	0.693*	0.684*		0.627*
	N	239	276	276	255	255		254
	Lees	0.373*	0.662*	0.502*	0.580*	0.523*		0.456*
	N	202	238	238	217	217		216
Vyf	Spel	0.649*	0.734*	0.545*	0.560*	0.591*		0.580*
	N	268	283	283	268	268		268
	Lees	0.564*	0.648*	0.506*	0.447*	0.505*		0.524*
	N	233	248	248	233	233		233
Ses	Spel	0.628*	0.708*	0.533*	0.534*	0.538*		0.576*
	N	259	275	275	259	259		259
	Lees	0.546*	0.662*	0.432*	0.478*	0.489*		0.496*
	N	226	242	242	226	226		226
Sewe	Spel	0.495*	0.641*	0.556*			0.521*	0.595*
	N	251	268	268			251	251
	Lees	0.378*	0.554*	0.398*			0.439*	0.446*
	N	217	234	234			217	217

* $p \leq 0.0001$

die betroubaarheid en geldigheid van die toetse is ondersoek. Die gewens toon aan dat die toetse betroubare metings verskaf. Ten opsigte van die geldigheid van die toetse is gevind dat dit oor goeie voorspel-

lingsgeldigheid beskik. Leerders wat dus swak in die lees- en/of speltoetse aan die begin van die akademiesejaar vaar, sal geneig wees om ook swak in hul skoolvakke aan die einde van die jaar te presteer. Daar bestaan dus 'n beduidende positiewe verband tussen hulle lees- en spellingvermoë en hul uiteindelijke prestasie in die skoolvakke. Hierdie verband is ook ten opsigte van die vak wiskunde gevind. Die sogenaamde nuwe wiskunde stel hoë eise aan die leerder se leesvermoë en dit is belangrik dat in die skoolsituasie hieraan aandag gegee sal word.

Summary

The ESSI reading and spelling test consists of a series of 28 vocabulary lists for Afrikaans-speaking and English-speaking learners in Grades 1 to 7. A need arose for the development of such a measuring instrument to replace the dated UCT reading and spelling test, which had appeared as long ago as 1944. Various phases were followed in developing the ESSI reading and spelling test. In an earlier article, the development work that was done during phases 1, 2 and 3 was outlined and discussed. In this article, the focus is on the last phase during which the psychometric characteristics of the ESSI reading and spelling tests were examined. Apart from the reliability and validity of the test, we also refer to skewness, kurtosis and standard error of measurement of the tests (vocabulary) for the various grades.

Skewness is a statistic that is required to determine the extent that the distribution of a set of test scores approaches a normal curve because this statistic measures the deviation from symmetry. As far as the Afrikaans tests are concerned, we deduced that the scores on the Grade 2 reading test, the Grade 3 spelling test, the Grade 4 reading and spelling test, and the Grade 7 spelling test exhibited a symmetrical distribution. The skewness coefficients of all these tests should be approximately zero. With the exception of the Grade 2 spelling test (0.22), all the other tests exhibited small negative skewness coefficients (that varied between -0.18 and -0.58). As far as the English tests are concerned, the reading tests for Grade 1 and Grade 3, as well as the spelling test for Grade 6, displayed a symmetrical distribution. Both the reading and spelling test for Grade 2, as well as the spelling tests for Grades 3 and 7, exhibited positive coefficients. These skewness coefficients varied from 0.31 to 0.81. All the other tests displayed small negative coefficients (that varied between -0.14 and -0.53). It therefore appears that there was a reasonable measure of symmetrical distribution of scores in all the Afrikaans and the English tests.

Kurtosis deals with the relative flatness or peakedness of the distribution curve. A normal curve will have a kurtosis of zero, while a positive kurtosis value signifies a more peaked curve. A negative kurtosis value again signifies that the distribution is flatter than the normal curve. All the Afrikaans and the English tests exhibited a negative kurtosis value. With the exception of a few English tests (the Grade 3 reading test, Grades 4 and 5 spelling tests), these values were small and it can be accepted that the scores generally have a normal distribution.

The standard error of measurement for the Afrikaans tests varied from 1.16 (Grade 1 reading test) to 1.89 (Grade 7 reading test). This implies that if a Grade 1 learner obtains a score of 12 out of 20 on the reading test, with a probability of 0.95, the learner's true score will not fall outside the range of $12 - 2.27 = 9.73$ and $12 + 2.27 = 14.27$. For the English tests, the standard error of measurement varied between 1.33 (Grade 1 reading test) and 1.96 (Grade 4 reading test). Not one of the tests had a standard error of measurement value that exceeded 2.

Reliability

As far as these indices were concerned, the test-retest reliability for each of the tests was calculated. Furthermore, an estimate was made of the tests' parallel form reliability by means of two different pro-

cedures, namely, the split-half method and the alpha-coefficient. The alpha coefficient was calculated for each test and it was only in a few exceptional cases that a test was found with a coefficient lower than 0.85. However, all these coefficients were still higher than 0.80 so that one may accept with a reasonable measure of confidence that the tests are internally consistent. The second procedure that was used to obtain an estimate of the various grades' parallel form reliability was the split-half method. The reliability results that were calculated, using this method, showed coefficients for the Afrikaans tests that varied between 0.767 and 0.891, whilst the statistics for the English tests varied between 0.815 and 0.919.

Test-retest reliability

The test-retest reliability was investigated by testing the same learners in the first and the fourth quarters of 1996 using the same vocabulary lists. (However, the Grade 1 learners were tested in the third and fourth quarters.) Since the two applications (with the exception of Grade 1) occurred after a significant interim period, one may accept with certainty that practice could not have played a role. The reliability results that were calculated with the assistance of this method showed that the coefficients for the Afrikaans tests varied between 0.603 and 0.868, whilst these values varied between 0.749 and 0.913 for the English tests. All the coefficients fell within the limits of the 0.01 significance level.

Validity

Concerning the validity of these tests, it is important to note that in this case, we worked in the applied domain and for this reason, attention was devoted only to the content and criterion-related validity of the tests.

To ensure content validity, support teachers, as well as the heads of support education at various child-guidance clinics, were involved in the identification of words for the test. These persons were experienced, in fact, as individuals they had had experience of at least 8 years or longer in the profession. They had to identify the words on the basis of the knowledge they had of the syllabi for the various grades, as well as their experience in the field of support teaching.

Criterion-related validity includes both predictive and concurrent validity. Only the predictive validity of the reading and spelling tests was investigated. During the first quarter of 1996, the learners (in this case, Grade 1 learners in the third quarter) were tested, after which information on their performance (expressed as a percentage) in their various school subjects at the end of 1996, was obtained from the schools. Each learner's total reading and spelling test score was correlated with scores on school subjects.

The calculated coefficients, albeit highest for the language subjects, were also significant at the 0.01 level for the other subjects. From this follows that a learner, who experiences backlogs with regard to reading and spelling, will definitely also experience learning problems. Not only are these correlation coefficients larger than zero in a statistical sense, they are also — with few exceptions — all higher than 0.50. According to Cohen (1988), a correlation of 0.50 constitutes a large effect.

The findings indicated that the tests yield reliable measurements, and that they have a sound predictive validity.

Verwysings

- Cohen J 1988. *Statistical power analysis for the behavioral sciences*, 2nd edn. Hillsdale, NJ: Erlbaum.
- Huysamen GK 1996. *Sielkundige meting — 'n Inleiding*, 2de uitg. Pretoria: Van Schaik.
- Lerner J 1993. *Learning disabilities: Theories, diagnosis, and teaching strategies*, 6th edn. Boston: Houghton Mifflin.
- Mercer CD 1997. *Students with learning disabilities*, 5th edn. Columbus: Charles E. Merrill.