

Ten Geleide

Goed onderwijs voldoet zowel aan een kwaliteits- als aan een gelijkheids criterium, aldus Katrijn Denies in het eerste artikel van dit nummer. Het scheert langs hoge toppen zonder grote groepen leerlingen achterop te laten hinken. In haar analyse van data uit de European Survey on Language Competences (2011) laat Denies zien dat het voortgezet onderwijs Engels in Vlaanderen en Nederland vlot voldoet aan dat eerste criterium. Ook voor het tweede criterium doen beide landen het op het eerste gezicht relatief goed: de kloof tussen de meest en de minst vaardige leerlingen is er over het algemeen kleiner dan in andere Europese landen. De prestaties van de Vlaamse en Nederlandse leerlingen hangen echter systematisch samen met onder meer het socio-economische en culturele kapitaal van hun gezin. Vooral voor Vlaanderen is dit een aandachtspunt.

Femke van der Meulen doet verslag van een onderzoek waarin werd gekeken welke lesmethode leesvaardigheid Nederlands het meest verbetert. Vier verschillende manieren van werken aan leesvaardigheid werden met elkaar vergeleken: het zelfstandig oefenen met het beantwoorden van vragen bij een tekst; het werken met een sorteertaak; het werken met een *mental model*-taak; en het gebruiken van leesstrategieën. Leerlingen uit 4 vwo maakten een voortoets en werkten volgens één van genoemde methodes. Nagegaan is hoe de leerlingen na afloop van de lessen presteerden op een natoets. Uit de resultaten blijkt dat vooral zwakke lezers baat hebben bij het oefenen van leesvaardigheid. De manier waarop geoefend wordt, lijkt daarbij geen verschil te maken.

De afgelopen decennia heeft het debat zichtbaar zijn weg gevonden naar de lessen in het voortgezet onderwijs, waarbij debatteren ook wordt gezien als een effectief

didactisch middel voor vreemdetaalverwerking. Deze effectiviteit wordt voornamelijk toegeschreven aan het feit dat debat in een klassikale context de vier taalvaardigheden combineert en de leerlingen stimuleert om samen te werken en op een interactieve wijze te leren. In verschillende onderzoeken komt ook naar voren dat leerlingen debatteren plezierig vinden. Abid el Majidi, Rick de Graaff en Daniel Janssen onderzochten de factoren die ten grondslag liggen aan die positieve houding ten aanzien van debatteren bij het vak Engels. De volgende factoren blijken van invloed op de positieve houding ten aanzien van debatteren bij de ondervraagde leerlingen uit 3 havo en 5 havo: actieve deelname, uitdaging, samenwerken, plezier, kritisch denken, taalvaardigheid en debat vs. lesboek.

Tinka Keehnen, Martine Braaksma en Martien de Boer doen verslag van een praktijkonderzoek dat uitgevoerd is op de OSG West-Friesland in de klassen 3 vwo. Nagegaan is wat het effect is van één les observerend leren op de leesvaardigheid van de leerlingen. Observerend leren is een didactiek waarbij de leerling leert door te kijken en te reflecteren op het geobserveerde gedrag in plaats van door te doen; ze heeft haar effectiviteit getoond in voorgaand onderzoek. Ook al bestond hier de interventie slechts uit één les, de resultaten van voor- en nameting op beide groepen doen vermoeden dat observerend leren een uitermate waardevolle aanvulling kan zijn op de lespraktijk bij het aanleren van leesvaardigheid.

Namens de redactie,
HELGE BONSET

‘Equity’ in de Engelse taalvaardigheid van Vlaamse en Nederlandse jongeren in Europees perspectief

De lat hoog en iedereen erover?

KATRIJN DENIES

Goed onderwijs voldoet zowel aan een kwaliteits- als aan een gelijkheids criterium. Het scheert langs hoge toppen zonder grote groepen leerlingen achterop te laten hinken. De European Survey on Language Competences (2011) laat zien dat het voortgezet onderwijs Engels in Vlaanderen en Nederland vlot voldoet aan dat eerste criterium. Ook voor het tweede criterium doen beide landen het op het eerste gezicht relatief goed. De kloof tussen de meest en de minst vaardige leerlingen is er over het algemeen kleiner dan in andere Europese landen. De prestaties van de Vlaamse en Nederlandse leerlingen hangen echter systematisch samen met onder meer het socio-economische en culturele kapitaal van hun gezin. Vooral voor Vlaanderen is dit een aandachtspunt.

Bij het bepalen van de effectiviteit van een onderwijssysteem voor een schoolvak als het Engels kunnen twee concepten centraal staan: niet alleen kwaliteit (*quality*), maar ook gelijkheid (in de internationale literatuur veelbesproken onder de noemer *equity*, zie bijvoorbeeld OECD, 2012). Kwaliteit meten we aan de hand van gemiddelde prestaties. In welke mate doen de leerlingen in een

bepaalde school of in een bepaald land het beter voor Engels dan elders? Kwaliteit heeft dus betrekking op niveauverschillen tussen groepen van leerlingen.

Bij het meten van gelijkheid gaat het om verschillen binnen die groepen. Het is mogelijk dat er veel leerlingen zijn die bijzonder hoge scores halen, naast heel wat leerlingen met bijzonder lage scores. In andere situaties presteren bijna alle leerlingen op een niveau dat niet ver afwijkt van het groepsgemiddelde. De consensus bij onderzoekers is dat succesvol onderwijs erin slaagt om resultaten te behalen die beter dan gemiddeld zijn, zonder dat er sprake is van grote verschillen tussen leerlingen (Sammons, 2010). De grootte van onderlinge prestatieverschillen moet daarom bewaakt worden. Bovendien is het nodig om na te gaan met welke factoren de prestatieverschillen samenhangen. Als leerlingen het alleen minder goed doen dan anderen omdat ze hebben gekozen voor een studierichting waarin het zwaartepunt niet bij talen ligt, is dat niet noodzakelijk problematisch. Ook kunnen we bezwaarlijk verwachten dat leerlingen zonder grote talentknobbel even hoog scoren als leerlingen die veel talent hebben voor talen. Er zijn echter andere factoren

waarvan we minder makkelijk mogen aanvaarden dat ze samenhangen met wat leerlingen bereiken op school, na controle voor hun talent en inzet. Voorbeelden zijn de migratiestatus, het gezinsinkomen en het geslacht van de leerling. Toch toont de literatuur aan dat een zwakkere socio-economische status voor veel schoolvakken wel degelijk samenhangt met minder goede prestaties (Sirin, 2005). Nagaan of ons onderwijs Engels goed is, houdt daarom ook een controle in van de mate waarin dat onderwijs erin slaagt om negatieve invloeden van instroomkenmerken te beperken.

In dit artikel rapporteren we over ons onderzoek naar de gelijkheid in de Engelse lees-, luister- en schrijfpredaties van leerlingen in het secundair/voortgezet onderwijs in Vlaanderen en Nederland. Dat onderzoek is gebaseerd op data van de European Survey on Language Competences (European Commission, 2012a), waaraan de auteur meewerkte als coördinator van de Vlaamse tak van het onderzoek. In wat volgt bespreken we eerst deze gegevensbron en het onderzoeksdesign. Na een korte blik op het niveau van het onderwijs Engels gaan we na hoe groot de prestatiekloof in Vlaanderen en Nederland is, en tot slot gaan we in op factoren die samenhangen met die prestatiekloof.

Context

European Survey on Language Competences (ESLC) is een initiatief van de Europese Commissie. Die benadrukt in haar kaderstrategie rond meertaligheid dat Europese burgers idealiter twee vreemde talen beheersen, naast hun moedertaal (European Commission, 2008). Om na te gaan in welke mate dit doel gerealiseerd is, legden in het voorjaar van 2011 bij benadering 53.000 leerlingen lees-, luister- en schrijftoetsen af in een vreemde taal: Duits, Engels, Frans, Italiaans of

Spaans. De deelnemers aan ESLC kwamen uit 14 verschillende landen: Bulgarije, Engeland, Estland, Frankrijk, Griekenland, Kroatië, Malta, Nederland, Polen, Portugal, Slovenië, Spanje en Zweden. De drie taalgemeenschappen van België namen afzonderlijk deel. In 15 van de 16 landen of taalgemeenschappen die deelnamen was het Engels één van de twee getoetste talen.

Populatie en steekproef

Nagenoeg elk land dat deelnam aan ESLC toetste het Engels bij veertien- of vijftienjarigen die Engels in hun vakkenpakket hadden zitten. Concreet waren dat in Vlaanderen 1.813 leerlingen in het vierde leerjaar secundair onderwijs. In Nederland namen 1.633 leerlingen deel. De toetsen werden daar afgenomen in het vierde leerjaar van het vmbo en in het derde leerjaar van het havo en vwo.

Wie precies moest deelnemen aan de Engelse toetsen, werd vastgelegd in twee fasen. Eerst werd een gestratificeerde steekproef van scholen getrokken, en daarna werden uit elk van die scholen gemiddeld 25 leerlingen ingeloot. Na afloop van de dataverzameling werden er gewichten berekend die ervoor zorgen dat de leerlingen correct vertegenwoordigd zijn in de analyses, ook als er vanuit bepaalde schooltypes uiteindelijk minder leerlingen dan nodig hebben deelgenomen. Deze benadering zorgt ervoor dat de resultaten van ESLC volledig representatief zijn voor alle leerlingen die in de betrokken landen Engels volgen in het getoetste leerjaar.

Instrumenten en data

Elk land gebruikte identieke, centraal ontwikkelde instrumenten om deel te nemen aan ESLC. Er waren uitgebreide achtergrondvragenlijsten voor de leerlingen, hun leerkrach-

ten en hun schooldirecties, en de leerlingen maakten lees-, luister- en schrijftoetsen onder toezicht van een onafhankelijke toetsleider. De taaltoetsen bevatten items die taalvaardigheid op de niveaus A1 tot en met B2 van het Europees Referentiekader voor Talen (ERK) konden onderscheiden (Perlmann-Balme, 2013; Robinson, 2013). De lees- en luisteroetsen bestonden uit meerkeuzevragen, ja/nee-vragen, en vragen van het type 'verbind A met B' of 'vul de gaten'. Deze werden gescoord aan de hand van een sleutel. De schrijftoetsen waren open vragen met een indicatie van het aantal woorden dat verwacht werd. Ze werden in elk land beoordeeld door een eigen team beoordelaars dat centrale instructies opvolgde om de ingevulde schrijftaken te voorzien van een score voor 'taal' en een voor 'communicatie'. 'Taal' betreft een correct gebruik van woordenschat en grammatica, terwijl 'communicatie' slaat op het vermogen om het gegeven onderwerp goed uit te werken met een gepaste stijl. Uit een controle van de interbeoordelaarsbetrouwbaarheid bleek dat sommige landen aan even goede schrijfproducten systematisch lagere scores toekenden dan andere landen. Daarom werden de schrijfscores van elk land lineair getransformeerd in de richting van de scores die een centraal beoordelaarsteam in Engeland gaf (European Commission, 2012b).

Op deze manier leverde ESLC voor elke leerling een reeks deelscores op voor de verschillende vragen uit de toets. Deze scores werden niet eenvoudigweg opgeteld tot een totaalscore. Ze werden geanalyseerd met behulp van een *partial credit* model uit de item-responstheorie (Masters, 1982). Concreet houdt dit statistische model rekening met de moeilijkheidsgraad van de toetsvragen om voor elke leerling en per vaardigheid (lezen, luisteren of schrijven) te bepalen waar hij of zij zich situeert op een schaal van de laagst mogelijke tot de hoogst mogelijke vaardigheidsscore. Op basis van een taaltoets kun-

nen we echter nooit met absolute zekerheid zeggen hoe taalvaardig een leerling exact is. Daarom werken internationale studies zoals ESLC met een indicatie van de vaardigheid van de leerling: op de vaardigheidsschaal wordt geen precieze waarde aangeduid, maar wel een zone met waarden waarin de leerling zich hoogstwaarschijnlijk bevindt. In de ESLC-dataset wordt de eigenlijke vaardigheid van de leerlingen geschetst aan de hand van een set van vijf precieze waarden uit die zone (zie Wu, 2005 voor meer over dit systeem van waarschijnlijk vaardigheidswaarden).

Om de waarden meer betekenis te geven, koppelde een groep van 21 experts ze terug aan het ERK. De experts beslisten samen welke scores minimaal nodig zijn om te kunnen spreken van taalvaardigheid op het niveau A1, A2, B1 of B2. Dat deden ze voor de receptieve vaardigheden aan de hand van een variant op de Van der Schoot-methode (2009) die gebaseerd is op de *bookmark*-methode (Mitzel, Lewis, Patz, & Green, 2001). Voor schrijven hanteerden ze de *body of work*-methode (Kingston, Kahl, Sweeney, & Bay, 2001). Aangezien er geen items waren die taalvaardigheid op de C-niveaus konden vaststellen, is het correcter om elke leerling die de B2-cesuur haalt te omschrijven als een leerling die presteert op niveau 'B2 of meer'

Methode

De volgende paragrafen presenteren analyses die allemaal gebruik maken van de eerder vermelde gewichten; daardoor zijn de resultaten representatief voor de doelgroep. Ook werken alle analyses met de vijf waarschijnlijk vaardigheidswaarden per leerling en vaardigheid. Elk soort analyse werd vijf keer uitgevoerd – één analyse per vaardigheidswaarde – en de resultaten werden geaggregeerd tot een eindresultaat volgens de richtlijnen van Rubin (1987).

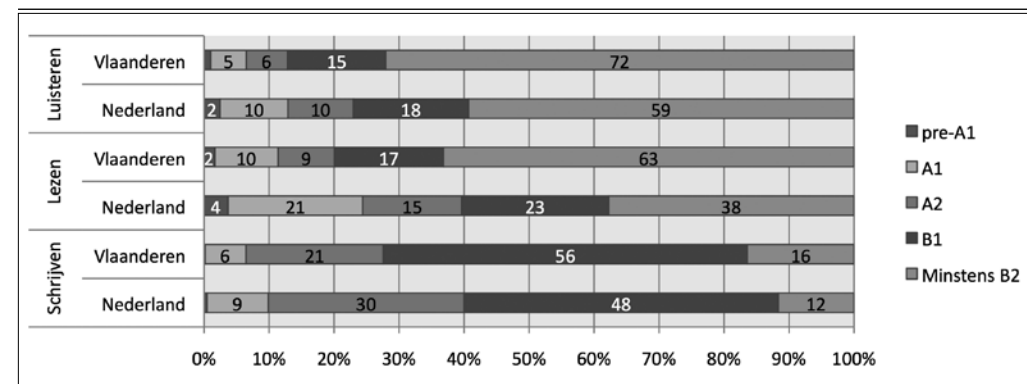
De kwaliteit van het onderwijs Engels staat niet centraal in dit artikel, maar we belichten dit aspect toch kort met een absolute en een relatieve beoordeling. De absolute beoordeling beschouwt de verdeling van de leerlingen over de ERK-niveaus. De relatieve beoordeling is gebaseerd op een rangschikking van de landen op basis van hun gemiddelde score en het betrouwbaarheidsinterval daaromheen.

De grootte van de prestatiekloof in elk land wordt onderzocht door verschillen tussen percentielen onder de loep te nemen. Hiertoe worden per land alle leerlingen gerangschikt volgens hun score en vervolgens in tien groepen van gelijke grootte verdeeld. Percentiel 50 is de mediaan, of de score van de leerling die precies in het midden van de reeks staat. Percentiel 10 komt overeen met de vaardigheidsscore van de leerling die door 90% van zijn of haar medeleerlingen overtroffen wordt, en percentiel 90 vertegenwoordigt de lat waar slechts 10% van de leerlingen overspringen. In deze analyses worden de absolute uitersten niet mee opgenomen. Percentiel nul en percentiel 100 staan immers onder sterke invloed van uitlopers. Eén uitzonderlijk begaafde leerling of één fout gelopen toets kunnen er voor zorgen dat het beeld van een bepaald land sterk vertekend raakt. Door de afstand tussen percentiel één en percentiel 99 in beschouwing te nemen, ontstaat er een betrouwbaarder beeld van de scorekloof tussen de zwakste en de sterkste leerlingen. Hoe dichter deze scores bij elkaar liggen, hoe minder onderlinge variatie er is in de scores van de leerlingen. De grafische weergave van de percentielscores werkt met gekleurde balkjes. De lengte van elk balkje geeft het aantal punten weer tussen twee percentielen. De totale lengte van alle balkjes samen kan onderling vergeleken worden: hoe langer de totale balk, hoe verder de leerlingen met de beste en de zwakste Engelse taalvaardigheid uit elkaar staan.

Tot slot identificeren multiniveau-regressieanalyses factoren waar de verschillen tussen leerlingen mee samenhangen. Een regressieanalyse voorspelt de score van de gemiddelde leerling op basis van een andere variabele. Dit soort analyse kan bijvoorbeeld onthullen of de score van een gemiddelde leerling beduidend vooruitgaat naarmate hij of zij meer lessen Engels volgt. De analyse kan drie mogelijke resultaten opleveren: de score van de gemiddelde leerling stijgt (een significant positieve coëfficiënt), de score verlaagt (een significant negatieve coëfficiënt) of de score verandert niet beduidend (een niet-significante coëfficiënt). Dit artikel concentreert zich niet op de samenhang van de scores met kenmerken van de lessen, maar wel op de samenhang met achtergrondkenmerken van de leerlingen. Wanneer dergelijke kenmerken geen cijfers zijn maar categorieën – zoals het opleidingsniveau van de moeder – verloopt de analyse op basis van een referentiecategorie. Bijvoorbeeld: een leerling in de referentiecategorie heeft een moeder met een leidinggevende functie. De regressieanalyse onthult dan of de score van een leerling die een moeder heeft met een elementair beroep het beduidend beter of minder goed doet dan de leerlingen in de referentiegroep.

Resultaten: kwaliteit

Figuur 1 biedt een totaaloverzicht van de Engelse taalvaardigheid van de getoetste doelgroepen. In Vlaanderen haalt de gemiddelde leerling vlot het luister- en leesniveau B2 van het ERK. Dat is het hoogste niveau dat ESLC kon toetsen. In Nederland is het gemiddelde luisterniveau ook B2, maar de leerlingen doen het er iets minder goed voor lezen, waarvoor ze gemiddeld het niveau B1 halen. Het gemiddelde schrijfniveau van de leerlingen is in beide landen ook B1. In absolute



Figuur 1. Verdeling in procent van de leerlingen in Vlaanderen en Nederland volgens het behaalde ERK-niveau voor luisteren, lezen en schrijven

termen is dat goed: zowel in Vlaanderen als in Nederland is de gemiddelde leerling een onafhankelijke taalgebruiker. Bovendien zijn er amper leerlingen die het niveau A1 nog niet halen en dus nog geen basisgebruiker zijn.

Ook relatief gezien zijn de resultaten bevredigend. In een vergelijking van de gemiddelde scores van de 15 Europese landen en taalgemeenschappen die het Engels toetsten, plaatst Vlaanderen zich voor elke vaardigheid in de top-drie en Nederland in de top-vijf. Puur op basis van deze gegevens lijkt de relatieve kwaliteit van het onderwijs Engels in Vlaanderen en Nederland dus zeker aanvaardbaar te zijn. Andere landen die het bijzonder goed deden zijn Estland, Malta en Zweden. Zij vullen de top-vijf aan.

Resultaten: mate van ongelijkheid

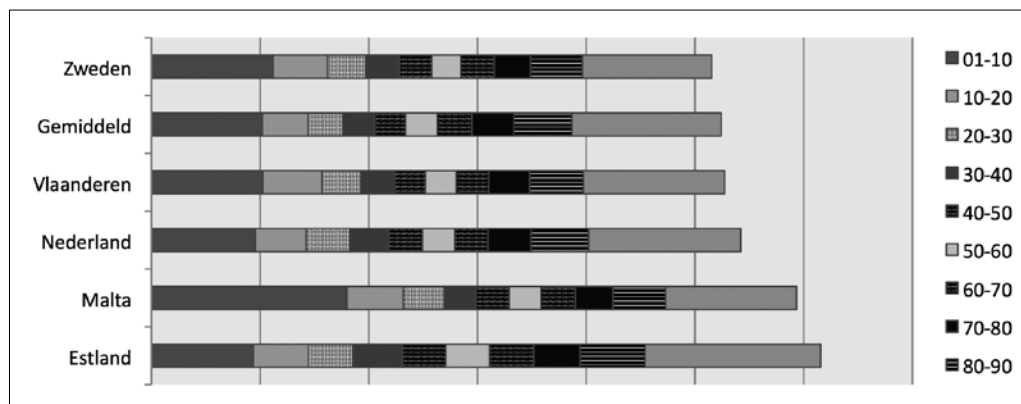
De figuren 2, 3 en 4 (zie pag. 8) geven de afstand weer tussen de 1% zwakste en de 1% sterkste leerlingen in Nederland en Vlaanderen. Om deze gegevens beter in een kader te plaatsen, laten de figuren ook de equivalente scorekloof zien in de drie andere landen in de top-vijf: Estland, Malta en Zweden. De lijst wordt aangevuld met

een indicator voor de situatie in Europa als geheel. De balk naast 'gemiddeld' geeft het gemiddelde weer van de scorekloof van alle landen en taalgemeenschappen die deelnamen aan ESLC.

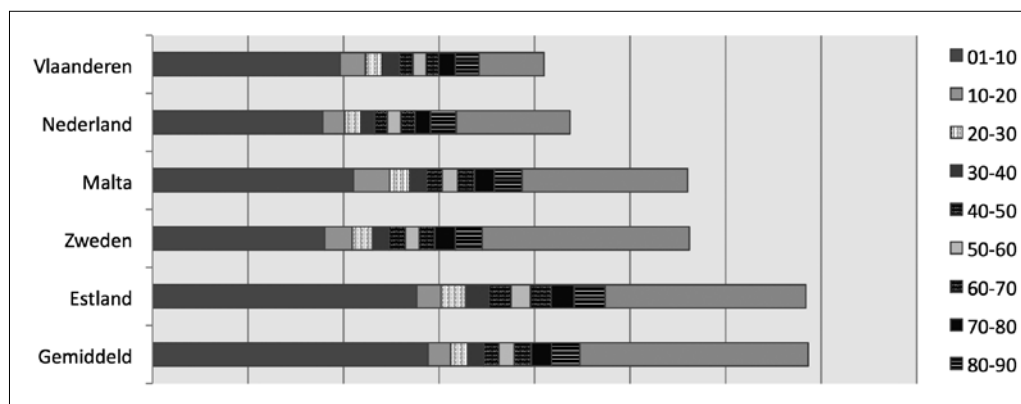
De figuren 2 en 3 hebben betrekking op de receptieve vaardigheden. Hieruit blijkt dat Vlaanderen en Nederland erin slagen om de ongelijkheid in de scores voor Engelse leesvaardigheid relatief beperkt te houden. Het verschil tussen de sterkste en de zwakste leerlingen is kleiner dan gemiddeld over de andere deelnemende landen heen. In Nederland is de ongelijkheid zelfs nog wat beperkter dan in Vlaanderen. Dat komt vooral doordat de scorespreiding in de Vlaamse staartgroep groter is dan in de Nederlandse. Voor luistervaardigheid stemt de scorekloof in Vlaanderen ongeveer overeen met het gemiddelde. Die in Nederland is net wat groter dan gemiddeld. Beide landen komen voor de receptieve vaardigheden telkens goed uit de vergelijking met de andere landen in de top-vijf. In Malta en Estland hinken de 10% zwakste leerlingen verder achterop ten opzichte van hun 10% sterkste landgenoten dan in de lage landen. De Zweedse scores voor luisteren vertonen ook meer ongelijkheid dan de Vlaamse en de Nederlandse,



Figuur 2. Afstanden tussen de percentielscores voor Engelse leesvaardigheid



Figuur 3. Afstanden tussen de percentielscores voor Engelse luistervaardigheid



Figuur 4. Afstanden tussen de percentielscores voor Engelse schrijfvaardigheid

maar de leesscores lopen minder uiteen.

Het beeld voor Engelse schrijfvaardigheid in figuur 4 is beduidend anders dan dat voor de receptieve vaardigheden. Het toont aan dat de scores van de gemiddeld presterende jongeren in zowat alle landen erg dicht bij elkaar liggen, maar dat de leerlingen in de staartgroep en die in de kopgroep onderling grote verschillen vertonen. De extremen zijn dus meer uitgesproken. Uit de gemiddelde scorekloof bij alle deelnemers aan ESLC blijkt dat vooral de minder goed presterende landen, die niet individueel opgenomen zijn in figuur 4, dit patroon vertonen. In de best presterende landen is de totale scorekloof altijd minder groot dan gemiddeld. Vlaanderen loopt in deze groep voorop als land met de kleinste verschillen. Het wordt gevolgd door Nederland.

Resultaten: correlaten van ongelijkheid

Tabel 1 (zie pag. 10) presenteert verschillende achtergrondkenmerken van leerlingen die kunnen samenhangen met hun prestaties voor het Engels. Het aantal boeken thuis fungeert als indicator van cultureel kapitaal. Bij gebrek aan informatie over het gezinsinkomen wordt de socio-economische situatie van de leerling geduïd door indicatoren zoals het diploma van de ouders of het hebben van een eigen slaapkamer.

Zowel in Vlaanderen als in Nederland blijkt er een significante samenhang te bestaan tussen de Engelse taalvaardigheid en verschillende achtergrondkenmerken.¹ In Vlaanderen meten we bijvoorbeeld een betere Engelse luistervaardigheid bij leerlingen die in België geboren zijn en deel uitmaken van een gezin met een groter cultureel kapitaal en ouders die een masterdiploma en een leidinggevend beroep hebben. Leerlingen in Nederland kunnen minder goed overweg met

Engelse teksten als ze thuis geen Nederlands spreken of niet in Nederland geboren zijn, of als hun moeder het lager onderwijs niet heeft afgemaakt of een dienstverlenend of elementair beroep heeft.²

Samen verklaren de variabelen in tabel 1 maar liefst 23% van de variantie voor Engelse leesvaardigheid in Vlaanderen (zie hiervoor tabel 2 op pag. 11). Voor luistervaardigheid en schrijfvaardigheid is dat respectievelijk 20% en 28%. In Nederland hangen de verschillen wat vaker samen met andere factoren dan de sociale achtergrond. De achtergrondkenmerken in tabel 1 verklaren daar 16% (leesvaardigheid), 15% (luistervaardigheid) of 8% (schrijfvaardigheid) van de variantie.

Conclusie

Het Vlaamse en het Nederlandse onderwijsstelsel blijken er beide in te slagen om de gemiddelde leerling naar een hoog niveau van Engelse taalvaardigheid te brengen. Echter, als de gemiddelde vaardigheidsscore hoog ligt maar de staartgroep erg zwak presteert, kunnen we nog steeds niet spreken van een succesvol onderwijssysteem. In dat opzicht is het goed nieuws dat het Vlaanderen en Nederland lukt om de scorekloof tussen de sterkste en de zwakste leerlingen beperkter te houden dan in andere landen die hoge gemiddelde scores behalen.

Voor Nederland is deze relatief grote gelijkheid geen uitzonderlijke vaststelling. Als onderwijssysteem liet het eerder al een behoorlijke mate van gelijkheid zien voor andere schoolvakken. In het PISA-onderzoek van 2009 was de scorespreiding voor wetenschappelijke geletterdheid ongeveer gelijk aan het gemiddelde, en voor wiskundige geletterdheid was deze beduidend kleiner dan gemiddeld. Voor Vlaanderen was de situatie in PISA 2009 echter anders, met voor beide domeinen een veel grotere scorespreiding

	LEZEN		LUISTEREN		SCHRIJVEN	
	VI	NI	VI	NI	VI	NI
Geslacht: vrouw	-.17**	ns	ns	-.26**	ns	ns
Beroepssituatie moeder: voltijds	ns	ns	ns	ns	ns	.38*
Beroepssituatie vader	ns	ns	ns	ns	ns	ns
Diploma moeder (r: lager sec.)						
Master of equivalent	ns	.47*	.53*	ns	ns	ns
Lager onderwijs niet afgemaakt	ns	-.75**	ns	ns	ns	ns
Diploma vader (r: lager sec.)						
Bachelor of equivalent	ns	ns	ns	ns	.62**	ns
Master of equivalent	ns	ns	.42*	ns	.82**	ns
Geen eigen slaapkamer hebben	ns	ns	ns	-.46*	ns	ns
Aantal boeken thuis (r: 0-10)						
26-100	.41**	ns	.37***	ns	.49**	ns
101-200	.53***	.32*	.55***	.36*	.65***	ns
201-500	.82***	.41*	.76***	.41**	.96***	ns
Beroep moeder (r: leidinggevend)						
Technicus of vakspecialist	ns	ns	ns	ns	-.36*	ns
Dienstverlener, verkoper	ns	-.39**	ns	ns	ns	ns
Geschoold landbouwer	ns	ns	-.86***	-1.20*	ns	ns
Assembleur, machinebediener	ns	ns	ns	ns	ns	-1.23**
Elementair beroep	ns	-.41*	-.37*	ns	ns	ns
Beroep vader (r: leidinggevend)						
Wetenschapper/artiest	ns	.26*	ns	.27**	ns	ns
Geschoold landbouwer	ns	ns	-.62**	ns	-.55***	ns
Ambachtsman	ns	ns	-.43**	ns	-.46*	ns
Elementair beroep	ns	ns	-.48*	ns	ns	ns
Geen: trekt bijstand	ns	ns	ns	ns	-1.08*	ns
Immigratiestatus (r: autochtoon)						
Immigrant eerste generatie	ns	-.52*	-.35*	ns	ns	ns
Immigrant tweede generatie	ns	ns	ns	-.37*	ns	ns
Thuis taal niet Nederlands	-.52**	-.69***	ns	ns	-.84***	ns

*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, ns: $p > 0.05$, r: referentiecategorie

Tabel 1. Regressiecoëfficiënten voor achtergrondkenmerken als predictor van Engelse lees-, luister- en schrijfvaardigheid in Vlaanderen en Nederland

	Vlaanderen			Nederland		
	LEZEN	LUISTEREN	SCHRIJVEN	LEZEN	LUISTEREN	SCHRIJVEN
Schoolniveau	39	23	24	24	19	0
Leerlingniveau	11	15	20	11	13	13
Totaal	23	20	28	16	15	8

Tabel 2. Proportie variantie (in procent en per niveau) in de Engelse lees-, luister- en schrijfvaardigheid van leerlingen die verklaard wordt door alle variabelen uit tabel 1 samen

dan gemiddeld (De Meyer & Warlop, 2010). Voor het Engels blijkt Vlaanderen dit patroon nu te doorbreken.

Mogelijk komt de beperktheid van de scorekloof deels doordat het verwerven van het Engels niet beperkt blijft tot de schoolbanken. Er kan zeker een effect zijn van de vele leerkanalen die zich in Vlaanderen en Nederland voordoen buiten de school. Het eindrapport van ESLC meldt bijvoorbeeld dat Engelstalige televisieprogramma's doorgaans gedubd zijn in ongeveer de helft van de deelnemende landen, maar niet in Vlaanderen en Nederland. De leerlingen komen ook via andere media regelmatig in contact met het Engels. Zo luistert 87% van de jongeren in de landen die deelnemen aan ESLC meerdere keren per week naar Engelstalige muziek. In Vlaanderen en Nederland is dat nog bijna 10% meer. Ruim 70% van de Vlaamse en de Nederlandse leerlingen bezoekt minstens maandelijks Engelstalige websites. Voor 36% van de Vlaamse en 40% van de Nederlandse jongeren ligt die frequentie zelfs op meerdere keren per week, en ook dat is ongeveer 10% meer dan gemiddeld. Het hoeft dan ook niet te verbazen dat in Vlaanderen 37% en in Nederland 43% van de leerlingen meerdere keren per maand een Engelstalig gesprek voert op het internet. In tegenstelling tot een vak als wiskunde, dat nagenoeg uitsluitend afhankelijk is van lessen op school, lijkt het Engels door deze bijkomende leerkanalen

buiten de school minder vatbaar voor grote prestatieverschillen.

Toch wil de beperktheid van de scorekloof niet zeggen dat Vlaanderen en Nederland schoolvoorbeelden zijn op het vlak van gelijkheid in het onderwijs Engels. In de relatief beperkte verschillen die er zijn spelen de achtergrondkenmerken van leerlingen een substantiële rol. In Vlaanderen verklaren ze zelfs ongeveer een vierde van de variantie in de vaardigheidsscores. Dit aandeel ligt zelfs nog enkele percentages hoger dan de impact van achtergrondkenmerken voor wiskunde in het PISA-onderzoek van 2012, waaruit bleek dat Vlaanderen van al de goed presterende landen het meest vatbaar was voor invloeden van de socio-economische status van leerlingen (OECD, 2013). Ook Nederland slaagt er niet volledig in om de thuissituatie van de leerling buiten spel te zetten voor Engels. Daar wegen achtergrondkenmerken net als in het PISA-onderzoek echter wat minder dan in Vlaanderen, met verklaarde varianties van slechts 8% voor schrijven tot 16% voor luisteren.

In Nederland maar vooral in Vlaanderen is een belangrijke opdracht dus om negatieve invloeden van een laag socio-economisch of cultureel gezinskapitaal verder in te dijken. Het is immers zeer onwaarschijnlijk dat alle kansarme leerlingen minder talent hebben voor talen dan leerlingen die geboren werden in gunstige gezinssituaties. Het meer buiten spel zetten van de sociale achtergrond ten

opzichte van de Engelse taalvaardigheid is een zaak waar het beleid aan kan werken, maar het kan ook op schoolniveau of op vakniveau aangepakt worden. Zowel beleidsmatig (in projecten) als vanuit de school of op initiatief van de leerkracht Engels kan er bijvoorbeeld gezorgd worden voor een betere huiswerkbegeleiding voor wie daar thuis niet op kan rekenen. Een ander mogelijk actiepunt is om de studiekeuzes van leerlingen meer te baseren op hun talenten en interesses. In Vlaanderen (en ook in Nederland) komen leerlingen met een lage socio-economische status namelijk vaker terecht in het technisch of beroepsgericht secundair onderwijs (VLOR, 2014), waar het leerstofaanbod voor onder meer Engels beperkter is. Zowel een betere oriëntatie vanuit de school als een grootschaligere aanpak vanuit het beleid kan hier helpen. Verder kan een mentaliteitswijziging nodig zijn. Onderwijseffectiviteitsonderzoek vergelijkt de prestaties van scholen doorgaans na correctie voor de achtergrondkenmerken van hun leerlingen. Die vergelijking wordt 'eerlijker' genoemd. Mogelijk verdringt deze werkwijze echter de pertinente vraag of een onderwijssysteem dat toelaat dat leerlingen met een zwakkere socio-economische achtergrond systematisch lagere vaardigheidsscores behalen nog wel eerlijk is. Het onderwijs moet streven naar gelijke resultaten voor gelijke talenten en inspanningen. Zeker voor een taalvak als het Engels, dat belangrijk is op de arbeidsmarkt, zijn dergelijke gelijke kansen essentieel.

NOTEN

1. Merk op dat de effectgroottes alleen vergeleken kunnen worden als ze betrekking hebben op dezelfde vaardigheid. Aangezien elke vaardigheid op een andere schaal werd gemeten, is het niet mogelijk om uit deze tabel af te leiden of 'geslacht' bijvoorbeeld een groter effect heeft op leesvaardigheid dan op luistervaardigheid.

2. Op basis van deze analyses is het niet mogelijk om te concluderen dat de betrokken achtergrondkenmerken de scorekloof rechtstreeks veroorzaken. We kunnen enkel een samenhang vaststellen tussen variabelen die gemeten zijn. De mechanismen achter deze samenhang moeten verder onderzocht worden voor we er uitspraken over kunnen doen.

LITERATUUR

- Barkaoui, K. (2013). Using multilevel modeling in language assessment research: A conceptual introduction. *Language Assessment Quarterly*, 10(3), 241–273.
- De Meyer, I. & Warlop, N. (2010). *Leesvaardigheid van 15-jarigen in Vlaanderen: De Eerste Resultaten van PISA 2009*. Brussel: Departement Onderwijs & Vorming.
- European Commission (2008). *Speaking for Europe: Languages in the European Union*. Luxembourg: Office for Official Publications of the European Communities.
- European Commission (2012a). *First European Survey on Language Competences: Final report*. Luxembourg: Publications Office of the European Union.
- European Commission (2012b). *First European Survey on Language Competences: Technical report*. Luxembourg: Publications Office of the European Union.
- Kingston, N. M., Kahl, S. R., Sweeney, K., & Bay, L. (2001). Setting performance standards using the body of work method. In G. J. Cizek (Ed.), *Setting Performance Standards: Concepts, Methods and Perspectives* (pp. 219–248). Mahwah, NJ: Lawrence Erlbaum.
- Masters, G.N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149–174.
- Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Setting Performance Standards: Concepts, Methods and Perspectives* (pp. 249–281). Mahwah, NJ: Lawrence Erlbaum.

OECD (2013). *PISA 2012 Results: Excellence Through Equity: Giving Every Student the Chance to Succeed (Volume II)*. Paris: PISA-OECD Publishing.

OECD (2012). *Equity and Quality in Education: Supporting Disadvantaged Students and Schools*. Paris: OECD Publishing.

Perlmann-Balme, M. (2013). European Survey on Language Competences – Comparability of A1 level competences across 5 languages. In M. Milanovic & C. J. Weir (Eds.), *Studies in Language Testing: Exploring Language Frameworks* (pp. 85–102). Cambridge: Cambridge University Press.

Robinson, M. (2013). Innovation in language test development. *Research Notes*, 52, 7–12.

Rubin, D.B. (1987) *Multiple Imputation for Nonresponse in Surveys*. New York: J. Wiley & Sons.

Sammons, P. (2010). Equity and Educational Effectiveness. In P. Peterson, E., Baker & B. McGaw (Eds.) *International Encyclopedia of Education, Volume 5, Leadership and Management – Politics and Governance* (pp 51–57). Oxford: Elsevier.

Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, 75(3), 417–453.

Van der Schoot, F. (2009). Cito variation on the bookmark method, in *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR) – A Manual*. Strasbourg: Council of Europe.

VLOR (2014). *Analytical report: Education for pupils with a migrant background in Flanders*. Geraadpleegd op <http://www.eunec.eu>.

Wolter, K. M. (2007). *Introduction to Variance Estimation*. New York: Springer.

Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, 31(2/3), 114–128.

KATRIJN DENIES coördineerde de Vlaamse uitvoering van de European Survey on Language Competences en was praktijklector in de lerarenopleiding Engels van de KU Leuven. In 2015 rondt ze haar doctoraatsonderzoek in de pedagogische wetenschappen af aan het Centrum voor Onderwijseffectiviteit en Evaluatie van de KU Leuven. In haar onderzoek bestudeert ze de vreemdetaalbeheersing van jongeren en hoe die samenhangt met kenmerken van henzelf en van het onderwijs dat ze krijgen.

E-mail: <katrijn.denies@ppw.kuleuven.be>