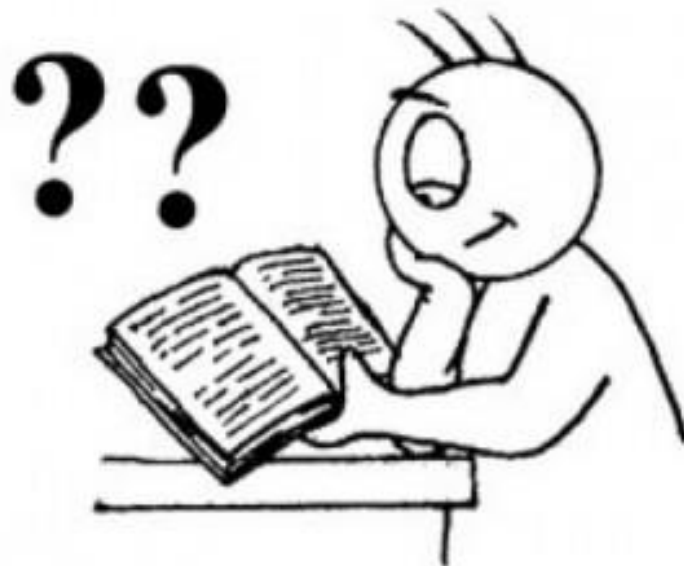

De ontwikkeling van een leerlingvolgsysteem voor leerlingen met Nederlands als tweede taal in het voortgezet onderwijs

De betrouwbaarheid en validiteit van de taalvaardigheidstoets



Masterscriptie

Master Education and Child Studies; Clinical Child and Adolescent Studies

Universiteit Leiden

Sylvana Gordijn; 0941646

Datum: 29-01-2013

Scriptiebegeleider: Prof. Dr. C.A. Espin

Tweede lezer: S. Chung MSc

Abstract

In deze studie is er onderzoek gedaan naar de betrouwbaarheid en validiteit van een taalvaardigheidstoets als indicator voor algemene taalvaardigheid. Participanten waren 135 (73 meisjes en 58 jongens) leerlingen uit het tweede leerjaar van het voortgezet onderwijs van één NT2 onderwijslocatie uit de Randstad. Het CBM meetinstrument was de taalvaardigheidstoets en de criteriumvariabelen in dit onderzoek waren de rapportcijfers Nederlands van de leerlingen. De resultaten gaven aan dat de taalvaardigheidstoets zowel betrouwbaar als valide is als indicator voor de algemene taalvaardigheid. De betrouwbaarheid varieerde van .709 tot .843, wat een hoog positief verband is. De correlaties tussen het rapportcijfer Nederlands en de taalvaardigheidstaken varieerden tussen de .386 en de .518, wat een matige positieve samenhang is. Het sterkste verband is gevonden aan het einde van het schooljaar. Hieruit kan geconcludeerd worden, dat de taalvaardigheidstoets een redelijke indicator is voor de algemene taalvaardigheid van de leerlingen met Nederlands als tweede taal.

Inleiding

Uit internationaal onderzoek van de Inspectie van het Onderwijs uit 2009 blijkt dat de kwaliteit van het onderwijs in Nederland voldoende is in vergelijking met het buitenland. Uit nationale nieuwsberichten komen echter andere signalen: “Twee derde pabostudenten zakt voor taaltoets” (2007); “Nijmeegse studenten kunnen niet spellen” (2011); “Taalniveau hbo’s bedroevend slecht” (2012). In de Leerplichtwet (1969) staat als doel dat iedereen in Nederland bij het verlaten van het voortgezet onderwijs in het bezit is van een startkwalificatie. Deze is nodig om de arbeidsmarkt te kunnen betreden en houdt in dat minimaal het MBO niveau 2 is afgerond. Om de startkwalificatie te kunnen behalen moet een leerling de Nederlandse taal beheersen.

De Rijksoverheid werkt daarom hard om de kwaliteit van het voortgezet onderwijs te verbeteren en dan vooral op het gebied van taal en rekenen (Rijksoverheid, 2012). De oorzaak van de slechte taalbeheersing wordt al teruggekoppeld naar de peuterleidsters. In maart 2012 heeft Minister van Onderwijs, Marja van Bijsterveldt, afspraken gemaakt met het kabinet. Er komt geld vrij om het taalniveau van het personeel van de voor- en vroegschoolse educatie te verhogen, waardoor peuters met een minder grote taalachterstand aan het basisonderwijs kunnen beginnen (“Miljoenen voor aanpak taalachterstand peuters“, 2012). Daarnaast wil de Rijksoverheid dat scholen de prestaties van leerlingen duidelijk gaan bijhouden met een leerlingvolgsysteem. Op deze manier kan er gekeken worden op welk niveau een leerling zit en of er misschien ingegrepen moet worden (Rijksoverheid, 2012).

Het is belangrijk dat elke leerling ondersteund wordt bij het aanleren van de Nederlandse taal. Echter is het voor de één moeilijker dan de ander. Zo zijn er veel leerlingen in Nederland die de taal als tweede taal moeten aanleren om zo mee te kunnen komen in de maatschappij.

Nederlands als tweede taal

In 2011 zijn er 16.370 mensen tussen de 10 en 20 jaar naar Nederland geïmmigreerd (Centraal bureau voor de Statistiek, 2012). Op het moment dat zij in Nederland wonen, geldt ook dat zij moeten voldoen aan een startkwalificatie voordat zij de arbeidsmarkt op kunnen. Voor de kinderen betekent dit vaak dat zij veelal zullen beginnen in een Internationale Schakelklas (ISK). Deze leerlingen zullen in een korte tijd de Nederlandse taal eigen moeten maken om zo mee te kunnen komen in de lessen. Het is belangrijk dat leerkrachten zicht hebben op de

ontwikkeling van de leerlingen, waar de leerlingen tegenaan lopen of waar ze extra ondersteuning in nodig hebben. Dit wordt in het reguliere onderwijs vaak gedaan doormiddel van een leerlingvolgsysteem. Uit een onderzoek van de Inspectie van het Onderwijs (2004) blijkt echter dat deze NT2 scholen niet in het bezit zijn van een goed leerlingvolgsysteem.

Het Cito volgsysteem

In zowel het basis- als voortgezet onderwijs in Nederland wordt er gebruik gemaakt van het Cito Volgsysteem. Met dit volgsysteem worden de prestaties gemeten van leerlingen op kernvaardigheden gedurende meerdere jaren. Vervolgens kunnen deze vergeleken worden met de landelijke norm (Sanders, 2001). Aan de hand van de volgsysteemrapportages kunnen de prestaties van individuele leerlingen vergeleken worden met leerlingen van hetzelfde of zelfs ander onderwijstype. Nadelen van dit volgsysteem zijn, dat er maar één versie van de toets is per niveau (vmbo/havo/vwo). Daarnaast is de afnametijd erg intensief en is er maar één keer per jaar een afname voor de tweede en derde klassen. Voor de eerste klassen is dit twee keer per jaar, aan het begin en aan het einde van het schooljaar (Sanders, 2001). Hierdoor is er geen mogelijkheid om op tijd in te grijpen wanneer leerlingen achter gaan lopen. Dit volgsysteem kan geen inzicht geven in de leervorderingen die een leerling maakt, juist omdat het zo weinig wordt afgenomen.

Het huidige volgsysteem sluit niet aan op de behoefte van de Rijksoverheid om de prestaties van leerlingen duidelijk in kaart te brengen. Op deze manier krijg je inzicht in een momentopname, maar niet in de leervorderingen van een leerling. Een volgsysteem dat wel aansluit op de behoefte van de Rijksoverheid is het curriculum-based measurement.

Curriculum-Based Measurement (CBM)

Deno (1985) heeft een assessmentmethode ontwikkeld waarbij frequent, herhaalde observaties gevraagd worden van schoolvaardigheden. Het is een korte test of taak die regelmatig wordt afgenomen. Aan de hand van deze scores kan snel gezien worden of een leerling vooruitgang boekt of waar een leerling tegenaan loopt. Op deze manier kan op een leerkracht op een flexibele en aansluitende manier aansluiten bij de onderwijsbehoeften van de leerlingen. Zo kunnen leerlingen ook vergeleken worden met medestudenten en met zichzelf gedurende hun gehele schoolcarrière (Deno, 1985; McMaster et al., 2007). Vanaf de jaren '80 wordt de CBM gebruikt om data te verzamelen met een brede range van verschillende doelen (Deno, 2003). Tegenwoordig wordt de CBM ook gebruikt om te kijken of het mogelijk is om scores op toetsen of testen te voorspellen (Deno, 2003). Wright (1992) geeft aan dat de kracht van de CBM juist ligt in het voorspellen van de toetsscores en het

aantonen of een interventie aanslaat of dat deze voor de individuele leerling aangepast moet worden.

In het review van Wayman, Wallace, Wiley, Espin (2007) komt naar voren dat er de afgelopen jaren veel onderzoeken zijn gedaan met CBM en lezen. De meeste onderzoeken zijn uitgevoerd bij leerlingen van de basisschool. Deze onderzoeken waren vooral gericht op de specifieke vooruitgang van lezen, zoals bijvoorbeeld technisch lezen. Recent zijn er meer onderzoeken uitgevoerd die ook hebben gekeken naar de bredere vooruitgang van de algehele leesvaardigheid van leerlingen.

Wayman et al. (2007) concludeert dat er in verschillende onderzoeken sterke correlaties zijn gevonden voor de betrouwbaarheid en validiteit van de CBM. Uit een meta-analyse van Yeo (2010) komen tevens hoge correlaties ($r > .689$) naar voren, wat de indruk geeft dat CBM een goede voorspeller is voor het mondelinge leesniveau van leerlingen op andere leesvaardigheidstoetsen. Wayman et al. (2007) geven aan dat vooral de CBM maze taak de vooruitgang van lezen het best kan voorspellen en de groei kan weergeven van alle leerlingen op de middelbare school voor sociale vakken. Uit een onderzoek van Fuchs en Fuchs (1992) kwam zelfs naar voren dat de CBM maze taak op de computer een hoge beoordeling kreeg van de leerkrachten. Het bedekt volgens hen een groot deel van de leesvaardigheid en het was gemakkelijk af te nemen. Zelfs studenten gaven aan de CBM maze leuk te vinden om te doen. Daarnaast was de CBM maze taak ook gevoelig voor de verschillende groei van studenten, welke goed te zien was op een grafiek wat de vooruitgang is van een student.

Ondanks dat er geen betrouwbaarheid en validiteit is aangetoond wordt er een algemene conclusie over de CBM getrokken in het onderzoek van Fuchs, Deno & Mirkin (1984). Het gebruik van de CBM resulteerde in significante prestatieverbeteringen bij de leerlingen. Tevens kwam naar voren dat zowel leerkracht als leerling beter op de hoogte waren van de onderwijsvorderingen. Hierdoor werden leerkrachten gestimuleerd om meer verschillende interventies te proberen wanneer bleek dat een leerling geen vooruitgang boekte.

Over het algemeen zijn er een redelijk aantal positieve resultaten gevonden in het onderzoek naar de validiteit en betrouwbaarheid van de CBM met betrekking tot leesvaardigheid, maar hoe zit dit met taalvaardigheid?

Curriculum-Based Measurements in Taalvaardigheid (CBM-T)

De online encyclopedie, Encyclo (2012), omschrijft taalvaardigheid als volgt: “Taalvaardigheid is het vermogen om te luisteren, spreken, lezen en schrijven in samenhang en continuïteit”. Uit een onderzoek van Stahl, Jacobson, Davis en Davis (2006) komt naar voren dat er een samenhang is tussen lezen en woordenschat. Dit zou betekenen dat als er meer gelezen wordt, de woordenschat groeit en dit een positieve invloed heeft op de taalvaardigheid. Er zijn weinig onderzoeken gedaan naar woordenschat en CBM.

Curriculum-Based Measurements bij NT2 leerlingen in het middelbaar onderwijs

Tot op heden is er nog niet veel onderzoek gedaan naar CBM bij leerlingen met Nederlands als tweede taal in het middelbaar onderwijs. In de Verenigde Staten is er onderzoek gedaan naar CBM en leerlingen met Engels als tweede taal. Baker & Good (1995) hebben gekeken naar de betrouwbaarheid en validiteit van de CBM bij tweetalige leerlingen en eentalige leerlingen. Ze hebben in totaal 76 leerlingen onderzocht, 50 leerlingen met Spaans als tweede taal en 26 eentalige leerlingen. De leerlingen hebben 13 weken lang, twee maal per week een CBM leestaak gemaakt. Hieruit kwam naar voren dat eentalige leerlingen hoger scoorden op deze taken dan de tweetalige leerlingen. Het verschil bleek echter niet significant.

In 2005 hebben Klein en Jamerson een onderzoek gedaan naar de leesvaardigheid van leerlingen met Engels als tweede taal. Er is data verzameld van 4000 autochtone en Spaanse leerlingen van groep 3 tot 5. Er is gekeken of er een bias was voor vloeiend voorlezen om de scores op een schoolse vaardigheden test te voorspellen. Hieruit kwam naar voren dat de leerlingen die Engels hadden als tweede taal, vaker overschat werden dan leerlingen met Engels als eerste taal (Klein & Jamerson, 2005). Bij beide onderzoeken werd wel benoemd dat de CBM leervaardigheidstaak voor beide groepen leerlingen voldoende werkt voor een algemene meting voor leesvaardigheid.

Wiley en Deno hebben in 2005 een onderzoek gedaan naar leerlingen met Engels als tweede taal. Er is onderzocht of CBM gebruikt kan worden om te voorspellen welke leerlingen het risico lopen om uit te vallen op een schoolse vaardigheden test. Het onderzoek is afgenomen in het basisonderwijs en er zijn gemiddelde tot sterke correlaties gevonden tussen de test en de CBM taak. Er wordt geconcludeerd dat door beide taken, CBM “oral-reading” en de maze-taak, te combineren er meer voorspellende waarde was voor de uitslagen voor lezen op de schoolse vaardigheden test. Dit resultaat was gevonden voor zowel de eentalige als tweetalige leerlingen.

Ramirez en Shapiro (2007) hebben onderzocht of het vloeiend lezen van de eerste taal (Spaans) een voorspeller kan zijn voor het vloeiend lezen in een tweede taal (Engels). Een totaal van 68 tweetalige leerlingen van groep 3 tot en met 8 hebben meegedaan aan het onderzoek. In de herfst, winter en lente zijn er CBM taken afgenomen. De resultaten laten zien dat er redelijke hoge correlaties zijn met de verschillende datamomenten, met uitzondering van groep 7. En er zijn ook significante resultaten gevonden dat het vloeiend lezen van Spaans in het begin van het jaar en het lezen van Engels aan het eind van het jaar kunnen voorspellen. Ramirez en Shapiro (2007) benoemen dat CBM een goed instrument is om de relatie tussen de eerste taal en tweede taal te evalueren.

In 2011 is er een onderzoek geweest van Buijs binnen het NT2 middelbaar onderwijs in Nederland. Deze studie had 105 leerlingen uit de tweede fase van een schakelklas. Er is gekeken naar de validiteit en betrouwbaarheid van de CBM lees- en woordenschattaak als indicator voor de algemene Nederlandse taalvaardigheid bij deze leerlingen. Hieruit kwam naar voren dat de CBM lees- en woordenschattaken redelijke indicators zijn voor het taalvaardigheidsniveau van de leerlingen. Echter, er werd wel ter discussie gesteld of de huidige CBM taken voldoende valide zijn om de vaardigheden te meten.

Geconcludeerd kan worden dat er in de Verenigde Staten al enkele onderzoeken zijn gedaan naar CBM en het leren van een tweede taal. Er zijn redelijk wat positieve resultaten gevonden dat CBM kan helpen met het aanleren van een tweede taal. In Nederland is er bijna nog geen onderzoek naar gedaan.

Dit onderzoek

Nederland heeft een onderwijsvolgsysteem, maar dit systeem staat niet toe om in een korte periode te kunnen oordelen op welk niveau een leerling zit. Curriculum-based measurement is een manier om de resultaten van leerlingen op een gemakkelijke en overzichtelijke manier weer te geven. Het is een eenvoudig volgsysteem dat op een korte en effectieve manier een beeld kan creëren over het niveau van een leerling. Zo zou het gemakkelijk moeten zijn om de ‘zwakkere’ leerlingen sneller te identificeren en een aangepaste instructie te geven. De taalvaardigheidstaken duren kort, 2 minuten, en het programma kijkt de taken na. Dit zou een efficiënte en gemakkelijke manier zijn om de taalvaardigheid van een leerling bij te houden.

In dit onderzoek wordt er gekeken of met een meetinstrument voor taalvaardigheid, de taalvaardigheidstoets, iets gezegd kan worden over de taalvaardigheid van een leerling. Het doel van dit onderzoek, tevens de hoofdvraag, is: “Kan de taalvaardigheidstoets gebruikt worden om het taalvaardigheidsniveau van een leerling op het voortgezet onderwijs met

Nederlands als tweede taal in Nederland te bepalen?”. Deze vraag wordt beantwoord door middel van de volgende onderzoeksvragen:

Onderzoeksvraag 1: Is de taalvaardigheidstoets betrouwbaar?

Deelvraag 1.1: Is er een samenhang tussen de scores van aangrenzende taken? De verwachting is dat er een samenhang gevonden wordt tussen deze meetpunten.

Onderzoeksvraag 2: Is de taalvaardigheidstoets valide?

Deelvraag 2.1: Is er een samenhang tussen de scores, aan het begin van het schooljaar, op de taalvaardigheidstaken en andere taalvaardigheidsprestaties (cijfers NL)? (concurrente validiteit). De verwachting is dat er een samenhang wordt gevonden voor de cijfers Nederlands.

Deelvraag 2.2: Is er een samenhang tussen de scores, aan het begin van het schooljaar, op de taalvaardigheidstaak en de scores van de taalvaardigheidstaak aan het einde van het schooljaar? (predictieve validiteit). De verwachting is dat de taalvaardigheidstaken een stabiele score produceren, zodat met scores aan het begin van het jaar iets gezegd kan worden over de mogelijke verwachting van groei in een schooljaar en dus de scores aan het eind van een schooljaar.

Deelvraag 2.3: Is er een samenhang tussen de scores aan het einde van het schooljaar op de taalvaardigheidstaken en de scores op andere taalvaardigheidstoetsen (cijfers NL)? (concurrente validiteit) De verwachting is dat er een samenhang wordt gevonden voor de cijfers Nederlands.

Methode

Steekproef

Dit onderzoek is een deel van een groter onderzoek over het CBM volgsysteem. De participanten zijn afkomstig uit het voortgezet onderwijs van een Nederlandse Internationale Schakelklas uit de Randstad. Alleen leerlingen uit leerjaar twee van dezelfde locatie hebben meegedaan in dit onderzoek. In totaal waren dit 135 leerlingen. De steekproef bevat 73 meisjes (54.1%), 58 jongens (43.0 %). De leeftijd van de leerlingen is gemiddeld 15,19 (SD: 1.774). De meeste leerlingen (38) komen uit Oost-Europa en Rusland (28.1 %), 38 leerlingen komen uit niet-westerse en Midditerraanse landen (28.1 %), 8 leerlingen komen uit Nederland (5,9 %), 7 leerlingen uit deels Nederlands sprekende landen (5.2 %) en 2 uit West-Europese landen (1.5 %) en 35 leerlingen komen uit de resterende landen (25.9 %).

De leerlingen hebben bezwaarbriefjes meegekregen naar huis. Ouders konden hierin aangeven of zij bezwaar hadden dat hun kind(eren) meededen aan dit onderzoek. Geen van de ouders uit dit onderzoek hadden bezwaar.

Meetinstrumenten

Hier worden de meetinstrumenten die voor dit onderzoek zijn gebruikt uitgebreid beschreven. De gebruikte meetinstrumenten bestaan uit de volgende instrumenten, de taalvaardigheidstoets, dit is een taal- en woordenschattoets. En de criterium variabele, Schoolcijfer Nederlands, rapport 1 en rapport 3.

Taalvaardigheidstoets: De taalvaardigheidstoets functioneert in de Elektronische LeerOmgeving (ELO). Dit is een webbased toetsstelsel waarmee de taalvaardigheid en woordenschat van een leerling gemeten kan worden. Wanneer het programma geopend is, wordt er begonnen met twee voorbeeldvragen en de instructie van de leerkracht. Vervolgens krijgt een leerling twee minuten de tijd om de meerkeuzetaken te maken. Er zijn omschrijvingen van verschillende definities te lezen en hierbij zoekt de leerling de correcte definitie, die onder de omschrijving worden gegeven met vier andere definities (negatieve alternatieven). Wanneer de twee minuten voorbij zijn stopt het programma en kun je niet meer verder. Het programma berekent zelf de score en slaat deze op. Het idee was gebaseerd op een onderzoek van Espin, Shin en Busch (2005), hierin werd geconcludeerd dat deze taalvaardigheidstaak een valide indicator is voor sociale studies op het middelbaar onderwijs. Hieronder vallen de vakken zoals aardrijkskunde, geschiedenis en maatschappijleer.

Deze taalvaardigheidstaak is iets anders dan die van Espin et. Al. (2005). Voor deze taalvaardigheidstaak zijn er twee versies ontwikkeld, één voor het VMBO en één voor het NT2 onderwijs. In dit onderzoek is er alleen gebruik gemaakt van de NT2 versie. De “Nederlands als tweede taal” (NT2) versie, heeft zijn definities uit het “Van Dale pocketwoordenboek Nederlands als tweede taal (NT2).” (Stumpel & Verburg; 2009). De gebruikte afleiders komen uit het basis Woordenboek Nederlands (De Kleijn, 2001). De eerste druk van dit woordenboek is gemaakt in samenwerking met het ministerie van Onderwijs en Wetenschap. Het bevat 2143 basiswoorden uit de Nederlandse taal. De woorden opgenomen in dit woordenboek komen overeen met de inhoud van het examen ‘Het Certificaat Nederlands als vreemde taal’ basiskennis (het middenniveau).

Voor het ontwikkelen van de verschillende taken is er een lijst opgezet met alle goede antwoorden. Dit is twintig keer gedaan, zodat er twintig verschillende taken ontwikkeld konden worden. Op deze manier was er voor twintig weken lang elke keer een andere lijst. Deze lijsten zijn ontstaan aan de hand van de papieren versie en het gebruik van een online randomizer. In deze randomizer zijn alle definities en afleiders ingevoerd, zodat deze in een random volgorde georganiseerd werden. Daarna is gekeken of er geen dubbele woorden in zaten. Wederom is de randomizer gebruikt om de vier foute alternatieven te krijgen per taak. Wederom is hier een controle uitgevoerd op dubbele alternatieven, zodat niet hetzelfde woord twee keer bij één taak zou staan of het woord in de eerstvolgende taak ook weer voor kwam. Dit is gedaan om de gokkans gelijk te houden. De dubbele woorden werden vervangen door woorden die nog niet in de taken voorkwamen. Wanneer alle definities, antwoorden en negatieve alternatieven waren gecontroleerd kon de taak ingevoerd worden in het computerprogramma.

Criterion Toets: Cijfers Nederlands. De schoolcijfers bepalen het niveau en de vooruitgang van een individuele leerling op een bepaald tijdstip. In dit onderzoek is er gebruik gemaakt van de cijfers zoals die zijn aangeleverd door de school. Over de manier waarop het cijfer tot stand is gekomen bestaan geen eenduidige criteria. Een schoolcijfer is een waardering van een schoolprestatie uitgedrukt in een getal. In Nederland is wettelijk vastgelegd, dat op scholen werk van leerlingen wordt beoordeeld door middel van de getallen 1 tot en met 10 (Eindexamenbesluit VO, 2012). Binnen deze locatie wordt het schoolcijfer mede bepaald door de inzet van de leerlingen. Hoe meer een leerling zich inzet hoe hoger het cijfer uitvalt.

(Meet)Procedure

Aan het begin van het schooljaar hebben leerkrachten een training gevolgd over het afnemen van onder andere de taalvaardigheidstoets. Nadat de leerkrachten de training hadden afgerond, zijn de klassen begonnen met het maken van de taken. Over een periode van een schooljaar hebben de leerlingen wekelijks een opdracht gemaakt achter de computer die hun taalvaardigheid meet. De scores op de taken werden automatisch berekend en opgeslagen. De eerste taken, 1 & 2, zijn geobserveerd om te zien of de leerkracht en leerlingen begrepen wat er van hun verwacht werd. Tenslotte werd aan het einde, taak 18 & 19, ook geobserveerd. Uit de observaties bleek dat de leerkrachten duidelijk de instructies volgde van hun training.

Resultaten

Data inspectie

In tabel 1 worden de beschrijvende gegevens van de verschillende taken weergegeven. In het onderzoek hebben de leerlingen 18 taken gemaakt. Maar bij een aantal taken waren er te weinig respondenten en zijn deze taken uit het onderzoek gehaald. Het gaat om taak 13 ($N=29$), taak 14 ($N=7$), taak 15 ($N=0$) en taak 16 ($N=0$).

Tabel 1

Beschrijvende gegevens van de taalvaardigheid taken ($N=135$)

<i>Taak</i>	<i>N</i>	<i>Missings</i>	<i>M</i>	<i>SD</i>	<i>Skewness</i>	<i>Min</i>	<i>Max</i>
1	108	27	9.69	5.137	.504	0	24
2	114	21	11.08	4.297	.055	1	22
3	108	27	11.69	4.769	.233	1	22
4	111	24	10.41	4.220	.345	2	23
5	98	37	10.60	4.998	.368	0	25
6	108	27	11.41	4.799	.233	2	23
7	111	24	11.76	4.491	.271	3	23
8	101	34	12.47	4.812	.240	1	24
9	104	35	9.75	4.831	.237	2	22
10	94	41	11.67	4.395	.319	2	22
11	105	30	12.65	3.954	.054	3	22
12	83	52	10.23	4.615	.488	3	24
18	92	43	13.13	5.183	.069	2	25
19	91	44	13.34	4.662	-.037	4	23

Note: Taak 1 & 2 zijn afgenomen op dezelfde dag, net als taak 18 & 19. Tevens zijn taak 1 en 18 dezelfde taak, net als taak 2 en 19.

In tabel 1 is te zien dat de scores een range hebben van 9.6 tot 13.4, ze variëren van week tot week. De scheefheid heeft een range van -0.037 tot 0.504. Over het geheel is er sprake van een zwakke scheefheid. De hoge scheefheid komt voor bij taak 1 en taak 12. De gepiekttheid was bij alle taken redelijk stabiel.

In tabel 2 is er een overzicht gegeven van de numerieke variabelen van dit onderzoek “Cijfers Nederlands rapport 1” en “Cijfers Nederlands rapport 3”. De cijfers zijn afkomstig uit

twee momenten in het schooljaar, namelijk periode december 2011 en mei 2012. De gemiddelde score blijft redelijk gelijk bij beide rapporten. De standaarddeviatie is wel wat groter aan het einde van het jaar. De scheefheid is minimaal, beide liggen rondom de 0.

Tabel 2

Beschrijvende gegevens verdeling numerieke variabelen

		<i>Cijfers NL</i>	
		<i>Rapport</i>	<i>Rapport</i>
		<i>1</i>	<i>3</i>
N	Valid	119	124
	Missing	16	11
	Mean	6.931	6.831
	Std. Deviation	.980	1.044
	Skewness	-.153	.019
	Minimum	4.80	4.20
	Maximum	9.10	9.30

Betrouwbaarheid van de taalvaardigheidstoets

Door middel van het berekenen van de correlaties wordt er geprobeerd antwoord te geven op de vraag “*Is de taalvaardigheidstoets betrouwbaar?*” De verwachting is dat er een samenhang gevonden wordt tussen deze meetpunten. Er is gekeken of er sprake is van een verband tussen de opvolgende taken zoals bij taak 1 & 2; 2 & 3; 3 & 4; 4 & 5;.....18 & 19? Door het berekenen van de correlaties van de opeenvolgende taken is er gekeken of de resultaten significant zijn en hoe hoog te gevonden correlaties zijn.

Taak 1 & 2 zijn op dezelfde dag afgenomen evenals taak 18 & 19, bij de andere taken zat er gemiddeld een week tussen de afnames. Taak 1 en 18 bevatten dezelfde tekst, dit geldt ook voor taak 2 en 19. In tabel 3 zijn de correlaties te zien van de aansluitende taken. De correlaties hebben een range van .71 tot .84. Alle correlaties zijn significant. De laagste correlatie die gevonden is tussen taak 11 en 12 ($r = .71, p < .001$). Er kan benoemd worden dat er een sterk positief verband is tussen de aansluitende teksten.

Tabel 3

Bivariate correlatie aansluitende taalvaardigheidstaken

	N	<i>r</i>
Taak 1&2	107	.77
Taak 2&3	100	.82
Taak 3&4	98	.80
Taak 4&5	90	.80
Taak 5&6	90	.80
Taak 6&7	101	.81
Taak 7&8	93	.84
Taak 8&9	88	.75
Taak 9&10	83	.84
Taak 10&11	83	.77
Taak 11&12	73	.71
Taak 18&19	90	.84

Note: Alle resultaten zijn significant bij $p < .001$

Taak 1 & 2 zijn afgenomen op dezelfde dag, net als taak 18 & 19. Tevens zijn taak 1 en 18 dezelfde taak, net als taak 2 en 19.

Validiteit van de taalvaardigheidstoets

Door middel van concurrente validiteit en predictieve validiteit wordt er geprobeerd antwoord te geven op de vraag “*Is de taalvaardigheidstoets valide?*” De verwachting is dat er een samenhang wordt gevonden tussen de scores op de taalvaardigheidstaken en de cijfers Nederlands.

Om deze vraag te beantwoorden wordt er gekeken naar de relatie tussen “Baseline begin” & “Rapport cijfers NL 1” (concurrente validiteit); “Baseline begin” & “Rapport cijfers NL 3” (predictieve validiteit) en “Baseline eind” & “Rapport cijfers NL 3” (concurrente validiteit). “Baseline begin” bestaat uit de scores van taak 1 en 2 opgeteld en weer gedeeld door 2. “Baseline Eind” Bestaat uit de scores van taak 18 en 19 bij elkaar opgeteld en tevens gedeeld door 2. In tabel 4 staan de gegevens weergegeven van bovengenoemde berekeningen.

Tabel 4

Correlatieanalyse onafhankelijke variabelen

	Pretest Cijfers (Rapport 1 cijfers NL)	Posttest Cijfers (Rapport 3 cijfers NL)
	<i>r</i>	<i>r</i>
Pretest taalvaardigheidstaken (Baseline begin)	.39 N =104	.39 N = 104
Posttest taalvaardigheidstaken (Baseline eind)		.52 N = 89

Note: ¹ Alle resultaten zijn significant bij $p < .01$

In de resultaten is er over het algemeen een matig positieve lineaire samenhang te zien met de onafhankelijke variabelen. De sterkste samenhang is te zien aan het einde van het schooljaar, bij het verband tussen het laatste rapportcijfer en de eindmeting van de taken. Er is sprake van een zwakke predictieve validiteit, de scores op de eerste taken kunnen de eindrapport cijfers zwak voorspellen. En er is sprake van een concurrente validiteit, er is een matige samenhang tussen de scores op de eerste taken en de cijfers Nederlands op het eerste rapport. Echter, er is een redelijke samenhang tussen de eindmeting van de taken en de Nederlands cijfers.

Als extra is er vervolgens gekeken naar de scatterplots om te zien of uitbijters invloed uitoefenden op de lage correlaties, maar dit bleek niet het geval te zijn. Er is duidelijk sprake van een geringe samenhang.

Daarna is er gekeken of er een betere correlatie te vinden is tussen het gemiddelde van taak 2 en 3 en de cijfers Nederlands aan het begin van het schooljaar. Dit omdat taak 1 als enige behoorlijk scheef verdeeld is zoals te zien in tabel 1. Het resultaat is $r = .46$, dit is een hogere samenhang dan bij de vergelijking met de baseline en de cijfers Nederlands aan het begin van het jaar.

Discussie en Conclusies

Het doel van dit onderzoek was om te onderzoeken of de taalvaardigheidstoets een betrouwbare en valide voorspeller is van de taalvaardigheid bij leerlingen met Nederlands als tweede taal in het voortgezet onderwijs in Nederland. De eerste deelvraag is: “Is de taalvaardigheidstoets betrouwbaar?”. Alle gevonden correlaties zijn significant ($p < .001$). De laagste correlatie die gevonden is tussen taak 11 en 12 ($r = .71$). Een mogelijke verklaring hiervoor is de kleinere N in vergelijking met de andere taken. De resultaten laten zien dat de taalvaardigheidstoets een betrouwbaar meetinstrument is. Er worden hoge correlaties gevonden bij de opeenvolgende taken. Er is duidelijk een sterk positief lineair verband gevonden. Dit betekent dat de score op de ene taak de score op de volgende taak kan voorspellen.

De tweede deelvraag is: “Is de taalvaardigheidstoets valide?”. Er is gekeken naar verschillende vormen van validiteit. Er is gekeken naar de concurrente validiteit in het begin en aan het eind. Daarnaast is er ook gekeken naar de predictieve validiteit van de taalvaardigheidstoets. De resultaten laten een matige positieve samenhang zien, van .386 tot .519. De sterkste samenhang is te zien bij de concurrente validiteit aan het einde van het jaar. Hier is een verband gevonden tussen de eindrapportcijfers en de scores op de taalvaardigheidstoets. De resultaten over de predictieve validiteit zijn matig, er is geen sprake van een sterk verband. Concluderend kan gezegd worden dat er een matige sprake is van validiteit bij de taalvaardigheidstoets. Alle correlaties zijn niet allemaal even sterk, maar wel significant. Aan het einde van het schooljaar is de relatie het sterkst.

Er kunnen wel enkele vraagtekens gezet worden bij de hogere concurrente validiteit aan het einde van het schooljaar. De rapportcijfers Nederlands, waarmee de taalvaardigheid taken vergeleken zijn, zijn gebaseerd op de behaalde cijfers op verschillende toetsen en de inzet van de leerling. Een leerling die zich meer heeft ingezet voor een vak zal op deze onderwijslocatie een hoger cijfer krijgen dan een leerling die minder inzet heeft getoond.

Ook kan er getwijfeld worden aan de lage concurrente validiteit aan het begin van het schooljaar. In tabel 1 toont taak 1 een grote scheefheid, welke afwijkt van de scheefheid bij de andere taken. Omdat taak 1 als enige afwijkt met zijn scheefheid bestaat het vermoeden dat de leerlingen nog niet wisten hoe de taak ging. Mogelijk hierdoor zijn de scores meer verdeeld bij taak 1 dan op de andere taken ($SD = 5.137$). Dit zou dan ook de lage samenhang kunnen verklaren bij de concurrente validiteit. Daarna is er gekeken of er een betere correlatie te vinden is tussen het gemiddelde van taak 2 en 3 en de cijfers Nederlands aan het begin van het

schooljaar. Dit omdat taak 1 als enige behoorlijk scheef verdeeld is zoals te zien in tabel 1. Het resultaat is $r = .46$, dit is een hogere samenhang dan bij de vergelijking met de baseline en de cijfers Nederlands aan het begin van het jaar. Dit komt vermoedelijk omdat taak 1 nog te nieuw was en de leerlingen nog moesten leren hoe de taak gemaakt moest worden. Bij taak 2 en verder hadden de leerlingen de taak onder de knie en scoorden hierdoor beter.

In dit onderzoek is er een redelijk bewijs dat de taalvaardigheidstoets een valide en betrouwbaar instrument is om het taalvaardigheidsniveau te bepalen van leerlingen met Nederlands als tweede taal. De gevonden resultaten komen ongeveer overeen met het eerdere onderzoek van Buijs (2011), dit onderzoek had echter meer sterkere data om een duidelijke conclusie te kunnen trekken. In het onderzoek van Espin, Shin en Busch (2005) is aangetoond dat taalvaardigheidstaken een goede voorspeller zijn voor sociale studies en nu ook voor het algemene taalvaardigheidsniveau. De taken kunnen gemakkelijk individueel of klassikaal en in korte tijd afgenomen worden in het onderwijs. De taalvaardigheidstoets kan gebruikt worden voor beslissingen over instroom, interventies, doorstroom en uitstroom van leerlingen. Ook in dit onderzoek is sprake van meer duidelijke data, waardoor deze conclusies getrokken konden worden.

Beperkingen van dit onderzoek zijn het aantal participanten, de steekproef is niet heel klein, maar groter had een beter beeld gegeven. In vergelijking met het onderzoek van Baker & Good (1995), Ramirez en Shapiro (2007) en Buijs (2011) is er in dit onderzoek sprake van een grotere steekproef. Echter, alle leerlingen komen uit één onderwijslocatie en uit één leerjaar. Er kunnen tevens vraagtekens gezet worden bij de werkelijke validiteit van de taalvaardigheidstoets. Bij dit onderzoek ontbrak een goed gestandaardiseerde criteriumvariabele. Omdat de gebruikte variabele mede gebaseerd is op een inschatting van een leerkracht kan er een grote foutmarge zijn, waardoor niet met zekerheid te zeggen is dat de taalvaardigheidstoets werkelijk valide is.

Als in een vervolgonderzoek de betrouwbaarheid en validiteit nogmaals bewezen kan worden met een betere criteriumvariabelen, en als vervolgens aangetoond kan worden dat de taken gevoelig zijn voor groei, dan is de taalvaardigheidstaak een leerlingvolgsysteem dat gemakkelijk is in gebruik en dat weinig tijd kost. Tevens wordt voorgesteld om een extra oefentaak toe te voegen. In dit onderzoek hadden de leerlingen twee oefenvragen, maar bij de eerste echte vraag was er nog een grote verdeling van scores. Dit kan wijzen op het feit dat na twee oefentaken de leerlingen de taak nog niet voldoende begrepen. Door een oefentaak extra toe te voegen kan deze spreiding misschien verholpen worden.

Dit onderzoek heeft aangetoond dat CBM mogelijk een goed aansluitend meetinstrument kan zijn binnen het voortgezet onderwijs voor leerlingen met Nederlands als tweede taal. Wanneer het instrument werkelijk betrouwbaar en valide is gebleken, kan dit systeem naast een huidig volgsysteem gebruikt worden op scholen. Doormiddel van CBM is er de mogelijkheid om in een korte periode duidelijk inzicht te krijgen en een bepaalde vaardigheid van een leerling. Op deze manier kan er vroeg ingegrepen worden en de leerling meer ondersteuning krijgen.

Referenties

- Baker, S. K., & Good, R. H. (1995) Curriculum-based measurement of English reading with bilingual Hispanic students: A validation study with second-grade students. *School Psychology Review*, 24(4), 561-578.
- Buijs, S (2011) *Ontwikkeling van een leerlingvolgsysteem voor Nederlands als Tweede Taal leerlingen in het middelbaar onderwijs: Betrouwbaarheid en validiteit van de maze- en woordenschattaak*. Universiteit Leiden: Masterscriptie
- Centraal Bureau voor de Statistiek (2012) *Immi- en emigratie naar geboorteland, leeftijd (31 December 2011) en geslacht*. Verkregen op 18 november 2012 van <http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=03742&D1=0-4&D2=0&D3=0,32,99-100&D4=0,1&D5=0&D6=a&HDR=T&STB=G1,G2,G3,G4,G5&VW=T>
- Deno, S.L. (1985) Curriculum-Based measurements: The emerging alternative. *Exceptional Children*, 52, 219-232.
- Deno, S.L. (2003) Developments in Curriculum-Based Measurement. *Journal of special education*, 37(3), 184-192.
- Encyclo (2012) *Begrip: Taalvaardigheid*. Verkregen op 15 mei 2012 van <http://www.encyclo.nl/begrip/Taalvaardigheid>
- Espin, C. A., Shin, J., & Busch, T. W. (2005) Curriculum-based measurement in the content areas: vocabulary matching as an indicator of progress in social studies learning. *Journal of Learning Disabilities*, 38(4), 353-363.
- Eindexamenbesluit VO (2012) *Wet- en regelgeving*. Verkregen op 23 oktober 2012 van http://wetten.overheid.nl/BWBR0004593/HoofdstukV/Artikel47/geldigheidsdatum_23-10-2012
- Fuchs, L. S., Deno, S. L., & Mirkin, P. (1984) The effects of frequent curriculum based measurement and evaluation on pedagogy, student achievement and student awareness of learning. *American Educational Research Journal*, 21, 449-460.
- Fuchs, L. S., & Fuchs, D. (1992) Identifying a measure for monitoring student reading progress. *School Psychology Review*, 21, 45-58.
- Inspectie van Onderwijs (2004) *Het onderwijs eerste opvang anderstaligen*. Verkregen op 18 november 2012 van <http://www.onderwijsinspectie.nl/actueel/publicaties/Het+onderwijs+eerste+opvang+anderstaligen.html>

- Inspectie van het Onderwijs (2012) *Bureau ICE; Toolkit Onderwijs en Arbeidsmarkt*. Verkregen op 19 april 2012 van <http://www.toets.nl/wat-kunt-u-meten-met-de-toa-55.html>.
- Inspectie van onderwijs. (2012) *Internationale vergelijkingen*. Verkregen op 9 mei 2012 van <http://www.onderwijsinspectie.nl/onderwerpen/Kwaliteit+van+het+onderwijs/Internationale+vergelijkingen>
- Klein, J. R., & Jimerson, S. R. (2005) Examining ethnic, gender, language, and socioeconomic bias in oral reading fluency scores among Caucasian and Hispanic students. *School Psychology Quarterly*, 20, 23–50.
- De Klein, P. (2001). *Woordenboek Nederlands*. Groningen: Noordhoff Uitgevers B.V. ISBN: 9789001470715
- Leerplichtwet 1969 (2012) *Wet- en regelgeving*. Verkregen op 5 november 2012 van http://wetten.overheid.nl/BWBR0002628/geldigheidsdatum_05-11-2012
- McMaster, K.L., Du, X. & Petursdottir, A. (2007) *Technical Features of Beginning Writing Measures*. Research Institute on Progress Monitoring, Minnesota.
- Miljoenen voor aanpak taalachterstand peuters (2012, 12 maart) *Nu.nl*. Verkregen op 10 mei 2012 van <http://www.nu.nl/politiek/2761118/miljoenen-aanpak-taalachterstand-peuters.html>
- Nijmeegse studenten kunnen niet spellen (2011, 26 mei) *Sp!ts*. Verkregen op 9 mei 2012 van http://www.spitsnieuws.nl/archives/binnenland/2011/05/studenten_radboud_uni_kunnen_n.html
- Rijksoverheid. (2012) *Kwaliteit voortgezet onderwijs*. Verkregen op 9 mei 2012 van <http://www.rijksoverheid.nl/onderwerpen/voortgezet-onderwijs/kwaliteit-voortgezet-onderwijs>
- Ramírez, de, R.D. & Shapiro, E.S. (2007) Cross-language relationship between Spanish and English oral reading fluency among Spanish-speaking English language learners in bilingual education classrooms. *Psychology in the Schools*, 44(8), 795-806.
- Sanders, P (2001) *Toetsen op school*. Arnhem: Stichting Cito Instituut voor Toetsontwikkeling.
- Stahl, S. A., Jacobson, M. G., Davis, C. E., & Davis, R. L. (2006) Prior Knowledge and Difficult vocabulary in the Comprehension of Unfamiliar Text. In K. A. D. Stahl & M. C. Mc Kenna (Eds.), *Reading Research at Work (Foundations of Effective Practice)*, 284-302. New York, London: The Guilford Press.
- Stumpel, R & Verburg, M. (2009) *Van Dale pocketwoordenboek Nederlands als tweede taal*

- (NT2). Uitgeverij Van Dale Lexicograaf: Nederland
- Taalniveau hbo's bedroevend slecht (2012, 3 april) *De Telegraaf*. Verkregen op 9 mei 2012 van http://www.telegraaf.nl/binnenland/11837100/___Taalniveau_hbo_s_bedroevend_slecht__.html
- Twee derde pabostudenten zakt voor taaltoets (2007, 16 januari) *De Ondernemer*. Verkregen op 9 mei 2012 van <http://www.deondernemer.nl/binnenland/662/Pabostudenten-zakken-voor-taaltoets.html>
- Wayman, M.M., Wallace, T.W., Wiley, H.I. & Espin, C.A. (2007) Literature Synthesis on Curriculum-Based Measurement in Reading. *Journal of special education*. 42 (2), 85-120.
- Wiley, H.I. & Deno, S.L. (2005) Oral reading and maze measures as predictors of success for English learners on a state standards assessment. *Remedial and Special Education*. 26(4), 207-214
- Wright, J. (1992) *Curriculum-based measurement: A manual for teachers*. Verkregen op 15 mei 2012 van <http://www.jimwrightonline.com/pdfdocs/cbaManual.pdf>
- Yeo, S. (2010) Predicting Performance on State Achievement Tests Using Curriculum-Based Measurement in Reading: A Multilevel Meta-Analysis. *Remedial and Special Education*, 31 (6), 412-422.