

Een leerlingvolgsysteem voor het
voortgezet onderwijs van leerlingen
met Nederlands als tweede taal

De validiteit en betrouwbaarheid van
de maze en woordenschat

Rianne Feijt; s0734454
Datum: 14 augustus 2012

1e begeleider: Prof. Dr. C. A. Espin
2e beoordelaar: S. Chung MSc

Voorwoord

Vijf jaar geleden stapte ik het faculteitsgebouw van de Sociale Wetenschappen binnen. Het enige wat ik toen wist is dat ik een studie wilde gaan doen waarbij het begeleiden van jonge mensen op de voorgrond staat. Hoe dat eruit moest gaan zien daar had ik eigenlijk nog geen idee van.

Nu, vijf jaar later, heb ik een scriptie geschreven die het resultaat is van keuzes die ik de afgelopen vijf jaar heb gemaakt. Deze keuzes heb ik kunnen maken dankzij inspirerende en enthousiasmerende mensen die ik in die vijf jaar ben tegen gekomen. Het voert te ver om alle mensen te noemen maar het tegenkomen van Christine Espin en haar enthousiasme om leerlingen niet buiten te sluiten maar te zorgen dat er gekeken wordt naar hun individuele groei en ontwikkeling is voor mij richtinggevend geweest. En heeft uiteindelijk geleid tot deze scriptie.

Het doen van het onderzoek voor de scriptie vond ik leuk en interessant maar het schrijven vond ik een “a hell of a job”. Gelukkig heb ik het niet alleen hoeven doen. Als eerste waren er de leerlingen die hun best hebben gedaan om de toetsen zo goed mogelijk te maken maar ook de docenten die er hun tijd en energie in hebben gestoken ondanks hun drukke schoolprogramma's. Vervolgens was er gedurende het hele proces de niet aflatende steun van Chris die met een bijna bovenmenselijk geduld wekelijks mijn klaagzang heeft aangehoord en Siuman mij altijd kon overtuigen dat het goed zou komen. Als laatste Mels, mijn vriend, die er voor heeft gezorgd dat ik tijd en energie kon steken in de scriptie. Allemaal super bedankt!

Inhoudsopgave	Pagina
Voorwoord	2
Samenvatting	4
Inleiding en theoretisch kader	5
Curriculum-Based Measurement (CBM)	6
Methode	8
Steekproef en proefpersonen	8
Onderzoeksinstrumenten	9
Onafhankelijke variabelen	9
De maze	9
De woordenschat	9
Criterium variabelen	10
De Trajecttoets	10
Het schoolniveau	12
Ontwikkeling van de taken	13
Procedure	13
Analysetechnieken	14
Resultaten	14
Betrouwbaarheid	15
Validiteit	16
Maze en woordenschat gecombineerd	17
Conclusie	19
Beperkingen van het onderzoek	20
Toekomstig onderzoek	20
Referenties	22

Samenvatting

In het onderwijs aan leerlingen die Nederlands als tweede taal moeten leren ontbreekt een passend leerlingvolgsysteem. Curriculum-based measurement (CBM) is een leerlingvolgsysteem uit Amerika waarbij door middel van korte taken de groei en ontwikkeling van schoolse vaardigheden gemeten kan worden. In Amerika is hiernaar al veel onderzoek is gedaan. In deze studie wordt onderzocht of een Nederlandse maze en woordenschattaak betrouwbaar en valide meetinstrumenten zijn om Nederlandse leesvaardigheid te meten. In totaal hebben 72 leerlingen meegedaan tussen de 13 en 18 jaar en afkomstig uit 22 landen. De resultaten laten zien dat de maze en de woordenschattaak betrouwbare meetinstrumenten maar de betrouwbaarheid is groter bij twee minuten dan bij één minuut. Ook zijn de maze en de woordenschattaak ruim gemiddeld tot goede voorspellers voor zowel het niveau van leesvaardigheid als het schoolniveau. Echter zijn de maze en woordenschattaak samen geen betere voorspeller, dit geldt voor zowel leesvaardigheid als schoolniveau.

Het Centraal Bureau voor de Statistiek rapporteerde dat 5327 tieners tussen de 12 en 18 jaar naar Nederland kwamen in 2010. Het leren van de Nederlandse taal is een verplicht onderdeel van de Inburgering voor Nieuwkomers (Wet Inburgering, 2007) en heeft prioriteit omdat het de sociale mobiliteit en integratie van minderheidsgroepen vergroot (Verspoor & Cremer, 2008).

Deze jeugdige nieuwkomers hebben, net als alle kinderen in Nederland, tot achttien jaar recht op onderwijs. In Nederland moeten kinderen tot 18 jaar verplicht naar school, tot 16 jaar heet dit leerplicht en daarna tot 18 jaar kwalificatieplicht. Het doel van het onderwijs is dat iedereen in Nederland bij het verlaten van het voortgezet onderwijs in het bezit is van een startkwalificatie om de arbeidsmarkt te kunnen betreden (Leerplichtwet, 1969). Een startkwalificatie betekent dat het onderwijs minmaal is afgerond op MBO 2 niveau. Als leerlingen van 16 jaar nog geen startkwalificatie hebben dan wordt de leerplicht verlengd tot 18 jaar. De onderwijs carrière van deze nieuwkomers start in Nederland veelal op een Internationale Schakelklas (ISK) als zij in een grotere gemeente wonen om daarna hun onderwijs carrière te vervolgen in het reguliere voortgezet onderwijs. Indien er geen ISK in de gemeente is dan starten zij direct in het reguliere voortgezet onderwijs met extra begeleiding. De achtergronden van deze jonge nieuwkomers zijn zeer divers zowel wat betreft het land van herkomst als het niveau van taalvaardigheid van de moedertaal. Ondanks deze verschillen is het voor alle kinderen van belang voor zowel de schoolcarrière als de latere kansen op werk en inkomen om zo snel mogelijk de Nederlandse taal te leren. Een taal bestaat uit verschillende vaardigheden zoals luisteren, spreken, schrijven en lezen.

Het kunnen lezen van een taal is niet alleen noodzakelijk om later maatschappelijk te kunnen functioneren (Organisation for Economic Cooperation and Development (OECD), 2001; Lyon, 2001; Vernooij, 2004) maar is tevens een noodzakelijke voorwaarde om met succes andere schoolse vaardigheden te kunnen ontwikkelen (Koda & Zehler, 2008; Lipka & Siegel, 2011) en is van groot belang voor een schoolcarrière (Verspoor & Cremer, 2008). Het succesvol volgen van onderwijs hangt niet alleen af van beheersing van de Nederlandse taal om te kunnen communiceren met anderen, maar leerlingen moeten de Nederlandse taal beheersen om te kunnen gebruiken bij het leren op school (Paradise, J., Genesee, F., & Crago, M. B., 2011).

Voor leerlingen die Nederlands als tweede taal (NT2-ers) moeten leren betekent het dat zij in kortere tijd dan hun leeftijdgenoten (NT1-ers) de taal moeten leren lezen om dezelfde kansen te kunnen krijgen op school om uit te stromen met een startkwalificatie.

Gemiddeld duurt het twee jaar voordat zij kunnen doorstromen naar het regulier voortgezet onderwijs. Voor de school is een belangrijke opdracht om op elk moment in de opleiding te weten welke ontwikkeling een leerling doormaakt voor zowel de interne doorstroming als de doorstroom naar het regulier voortgezet onderwijs.

In het onderwijs aan NT2-ers tussen de 12 en 18 jaar die Nederlands als tweede taal moeten leren is er geen sprake van een jaarklassensysteem. Leerlingen kunnen gedurende het hele schooljaar instromen en doorstromen als zij het gewenste niveau hebben bereikt. Voor het bepalen van de voortgang en het niveau op een bepaald moment kunnen scholen gebruik maken van beschikbare meetinstrumenten zoals gestandaardiseerde tests en rapportcijfers. Volgens een onderzoek van de Inspectie voor Onderwijs in 2004 beschikten de scholen voor nieuwkomers echter niet over een goed leerlingvolgsysteem. De resultaten op gestandaardiseerde testen laten de verschillen tussen de leerlingen zien op een bepaald moment en niet de verschillen bij één leerling gemeten over een tijdsinterval (Shin, Deno, & Espin, 2000). De bestaande gestandaardiseerde toetsen zijn geschikt om een uitstroom niveau te kunnen meten of een vaardigheid ten opzichte van landelijke normgroepen. Veelal kunnen deze toetsen maar één of enkele keren per jaar worden afgenomen op een voorgeschreven moment, zijn er maar enkele varianten van een toets, zijn zij over het algemeen erg tijdrovend en kosten veel geld. Hierdoor zijn ze minder geschikt om het niveau en de groei van leesvaardigheid bij een individuele leerling op elk gewenst moment te kunnen meten. Een geschikt leerlingvolgsysteem hiervoor is het in de USA ontwikkelde Curriculum Based Measurement (CBM).

Curriculum Based Measurement (CBM)

CBM is een methode ontwikkeld in de USA om de vooruitgang van de student en het effect van instructie van de docent te meten gebaseerd op het jaarcurriculum. Deze metingen moeten valide en betrouwbaar zijn, makkelijk toe te passen door docenten, makkelijk te zijn begrijpen en niet teveel kosten (Deno, 1985; Fuchs & Deno, 1991). In eerste instantie is CBM ontwikkeld in de USA voor leerlingen in het speciaal onderwijs als leerlingvolgsysteem om de ontwikkeling van leerlingen te meten en docenten te ondersteunen bij hun beslissingen. In de loop der jaren zijn er meerdere toepassingen gevonden voor de deze methode. Het doel bleef om schoolse vaardigheden op een voor docenten eenvoudige manier gegevens te verzamelen en te verwerken zodat de groei en ontwikkeling van de leerling zichtbaar werd en docenten te helpen om het lesprogramma te verbeteren de onderwijskundige aanpak bijgesteld werd. (Deno, 1985; Fuchs & Deno, 1991; Deno, 2003; Stecker, Fuchs & Fuchs, 2005). Ook bleek dat het goed gebruikt kon worden bij het nemen van beslissingen op individueel, klas en

schoolniveau (Pierce, McMaster, Deno, 2010) voor verschillende niveaus van onderwijs. Het curriculum is het uitgangspunt en op basis hiervan werden taken ontwikkeld.

De taken die ontwikkeld zijn bestaan uit korte toetsen die frequent, wekelijks kunnen worden afgenomen, klassikaal en individueel. De goede antwoorden worden geteld. De resultaten ervan worden op een eenvoudige wijze in een grafiek gezet en zijn voor betrokkenen eenvoudig te interpreteren (Hosp, Hosp & Howell, 2007). Op basis van onderzoeken werden standaardtaken ontwikkeld voor hardop lezen, begrijpend lezen, schrijven en rekenen (Deno, 2003). Voor het meten van de voortgang bij leesvaardigheid zijn de maze en de hardop-lezen taak veel gebruikt.

Onderzoek naar de bruikbaarheid van CBM-taken voor NT2-ers zijn er nauwelijks. Er is één onderzoek bekend van Wiley en Deno (2005). Dit onderzoek laat zien dat zowel de maze als een hard-op-lezen test het niveau kunnen voorspellen op de score van een Amerikaanse standaardtest, the Minnesota Comprehension Assessment in reading. Dit onderzoek was gedaan bij leerlingen van groep 5 en 7. Naar de bruikbaarheid van CBM-taken bij NT2-ers binnen het voortgezet onderwijs is zover bekend tot nu toe geen onderzoek bekend. Wel is uit onderzoek bij NT1-ers binnen het voortgezet onderwijs gebleken dat de hard-op-lees-taak minder gevoelig was voor het meten van groei dan in het basisonderwijs.

Hoewel de hardop-lees-taak in het basisonderwijs een hogere correlaties had met de standaardtoetsen (Wiley & Deno, 2005) bleek dat het in het voortgezet onderwijs minder tot niet gevoelig was voor het meten van groei (Wayman e.a., 2007; Tichà, Espin & Wayman, 2009). De maze bleek gevoelig te zijn voor de groei van leesvaardigheid (McMaster, e.a.,2006) en had hoge correlatie met de standaardtoetsen hoewel de correlaties in het voortgezet onderwijs lager waren dan in het basisonderwijs (Deno,2003).

Om voor het voortgezet onderwijs een tweede bruikbare taak te kunnen gebruiken is gekozen voor het toetsen van de woordenschat. Een reden hiervoor is dat uit onderzoek in het verleden is gebleken dat kennis van de woordenschat samenhangt met leesvaardigheid (Stahl, Jacobson, Davis & Davis, 2006). Er is geen onderzoek bekend naar de betrouwbaarheid en de validiteit van de woordenschattaak als indicator van leesvaardigheid. Uit onderzoek (Espin & Deno, 1994-1995; Busch & Espin, 2003) bleek wel dat scores op een woordenschattaak voor Engels en zaakvakken samenhangen met hardop-lezen en schoolprestaties van leerlingen binnen het voortgezet onderwijs.

Wat tot nu toe bekend is over de bruikbaarheid van CBM-taken voor NT2-ers in het voortgezet onderwijs is dat de maze een betrouwbaar en valide meetinstrument is om de groei van leesvaardigheid te meten bij NT1-ers in het voortgezet onderwijs en voor NT2-ers op de

basisschool in Amerika. Het is niet bekend of dit ook het geval is bij NT2-ers binnen het voortgezet onderwijs in Nederland. De betrouwbaarheid en validiteit van een woordenschat toets als CBM-taak is, zover bekend, zowel in Nederland als in Amerika nog niet eerder onderzocht.

Het doel van dit onderzoek is het onderzoeken van de bruikbaarheid van twee in Nederland nieuw ontwikkelde CBM taken, maze en woordenschat, als meetinstrument voor het meten van het niveau van leesvaardigheid van leerlingen die Nederlands als tweede taal moeten leren.

Centrale onderzoeksvraag:

Wat is de betrouwbaarheid en validiteit van een maze en woordenschat taak als indicator van leesvaardigheid in de Nederlandse taal voor leerlingen in het voortgezet onderwijs die Nederlands als tweede taal moeten leren?

De eerste subvraag is: Wat is de (paralleltest) betrouwbaarheid van de maze en woordenschattaak?

De tweede subvraag is: Zijn de maze en de woordenschattaak valide meetinstrumenten om het niveau van leesvaardigheid te voorspellen?

De derde subvraag is: Zijn de maze en woordenschattaak valide meetinstrumenten om de fase van het onderwijs waarin een leerling is geplaatst te kunnen voorspellen?

De vierde subvraag is: Zijn de maze en woordenschattaak samen een betere voorspeller voor het leesniveau en voor het schoolniveau dan apart?

Methode

Steekproef en proefpersonen

Deze studie was een onderdeel van een groter onderzoek naar de betrouwbaarheid en validiteit van meetinstrumenten ontwikkeld voor het meten van groei van lees- en taalvaardigheid. De studie was gedaan op een school voor voortgezet onderwijs in een grote stad in het westen van Nederland. De school verzorgt, naast het onderwijs voor NT2-ers, ook onderwijs voor de Praktijkschool (Pro), Voorbereidend Middelbaar Beroepsonderwijs (VMBO), Middelbaar Algemeen Voortgezet Onderwijs (MAVO), Hoger Algemeen Voortgezet Onderwijs (HAVO) en Voorbereidend Wetenschappelijk Onderwijs (VWO).

De proefpersonen in het onderzoek waren leerlingen van 13 tot en met 18 jaar die korter dan twee jaar in Nederland zijn en het Nederlands onvoldoende beheersen waardoor zij (nog) niet in staat waren om aan het reguliere voortgezet onderwijs deel te kunnen nemen. Er deden 72 leerlingen uit 9 klassen mee aan het onderzoek waarvan 28 mannen (39 %), 37 vrouwen (51 %) en 7 leerlingen (10 %) waarvan deze gegevens niet beschikbaar waren. De

gemiddelde leeftijd was 15,7. Er waren 53 leerlingen die onderwijs volgden in de 2^e fase en 19 leerlingen die dit deden in de 3^e fase. Het gemiddelde rapportcijfer voor Nederlands was een 7, het laagste cijfer was een 5 en het hoogste een 8.4. Van 65 leerlingen was hun geboorteland bekend en zij kwamen uit 22 verschillende landen. Hiervan kwamen er 14 uit Polen (19 %), 12 uit Turkije (17 %), 7 uit Bulgarije (10 %), en uit de andere landen per land één tot vijf leerlingen.

Onderzoeksinstrumenten

Onafhankelijke variabelen. Er waren voor dit onderzoek twee meetinstrumenten gebruikt. Een Nederlandse maze (maze) en een woordenschattoek (woordenschat).

De maze. De teksten van de maze waren informatief met een lengte van ongeveer 400 woorden waarin om de zeven woorden van het volgende woord een meerkeuze was gemaakt door twee vervangers toe te voegen. De leerlingen lezen gedurende twee minuten de tekst en omcirkelen het antwoord waarvan zij denken dat die het beste past in de tekst. Door de onderzoekers van de universiteit Leiden waren twee teksten geschreven met een moeilijkheidsgraad van AVI7 niveau, wat iets lager is dan het CEF A1 niveau (Lindhout, 2009). De twee teksten hadden als onderwerp: “Bewegen en sporten: Goed voor de geest” en “Koninginnedag”.

De teksten voor de maze waren op de volgende manier tot stand gekomen. De eerste zin was in tact gelaten en daarna was elk zevende woord in de tekst gemarkeerd; bij eigennamen van mensen, dieren, dingen, aardrijkskundige namen, cijfers of jaartallen was het achtste woord gemarkeerd. Bij elk gemarkeerd woord waren twee vervangers gezocht die in lengte maximaal 1 of 2 letters verschillen van het gemarkeerde woord, grammaticaal en qua betekenis niet in de tekst passen en elke vervanger mocht maar één keer in de tekst voorkomen. Per groepje van drie woorden waren de gemarkeerde woorden en de vervangers in willekeurige volgorde geplaatst, zodat het gemarkeerde woord gemiddeld even vaak op de eerste, tweede, derde positie kwam. De woorden van de maze waren vet, onderstreept en gescheiden door een / en tussen haakjes, als volgt: (**woord1 / woord2 / woord3**). De woorden van de maze kunnen verdeeld worden over twee regels. Ook over de typografie waren afspraken gemaakt. Het lettertype was Times New Roman, puntgrootte 13, de regelafstand 1.75, de kantlijnen boven, links en rechts op 2,5 cm., de titel was vetgedrukt, de kantlijn was niet naar rechts uitgelijnd, woorden waren niet afgebroken. Een complete tekst mocht niet meer dan twee bladzijden beslaan. (zie bijlage Ontwikkeling mazetaak).

De woordenschat. De taken van de woordenschat bestonden uit 27 woorden bovenaan de pagina geplaatst in een tabel en daaronder 25 omschrijvingen van woorden met

daarachter een stippellijn waar het juiste woord moest worden opgeschreven. De onderzoekers van de universiteit Leiden ontwikkelden twee taken waarbij gebruik gemaakt was van het Basis Woordenboek Nederlands (De Kleijn, 2001).

De taken voor de woordenschat waren op de volgende wijze tot stand gekomen. Uit het Basis Woordenboek, bestaande uit 2143 woorden, waren door de onderzoekers per taak random een pagina getrokken om vervolgens random een woord te kiezen van die bladzijde. Op deze wijze maakten de onderzoekers per taak een lijst van bestaande uit 27 woorden. Bij willekeurig 25 van deze woorden zochten de onderzoekers de omschrijving. De omschrijving van de getrokken woorden zochten zij vervolgens op in het van Dale Pocketwoordenboek Nederlands als tweede taal, NT2. De betekenissen opgenomen in dit woordenboek zijn zo veel mogelijk geschreven met de ruim 2000 woorden die staan in het Basiswoordenboek Nederlands. Een taak was zo gemaakt dat het op een A-vier formaat papier past. Het aantal woorden was kleiner dan het aantal omschrijvingen om de kans op het goed raden van de overgebleven woorden te verminderen. Aan de linkerkant van de pagina waren de definities gegeven. Leerlingen lezen de definities één voor één door en schrijven het bijpassende woord achter de definitie. De woorden staan bovenaan de pagina weergegeven.

Criterium variabelen. Er zijn twee criteriumvariabelen gebruikt. De eerste was een nationale taaltoets (Trajecttoets NT2) ontwikkeld om het taalniveau Nederlands bij tweedetalers te kunnen meten de schoolniveau. De tweede was de indeling van groepen leerlingen in verschillende fases op basis van de beslissingen van de school.

De Trajecttoets. De trajecttoets NT2 van het Bureau InterCulturele Evaluatie (ICE) was een toets is ontwikkeld om te kunnen meten of mensen een voldoende beheersing hebben van de Nederlandse taal om te kunnen inburgeren en was in 2009 goedgekeurd als exameninstrument door de Inspectie van het Onderwijs. De toets omvatte de onderdelen lezen, schrijven, spreken en een gesprek voeren en was gebruikt om het niveau van taalvaardigheid te meten. Omdat in dit onderzoek alleen was gekeken naar de betrouwbaarheid en validiteit van de maze en woordenschat om leesvaardigheid te meten was alleen het onderdeel lezen van de trajecttoets worden gebruikt.

De score op de toets, een niveau van leesvaardigheid, wordt aangegeven volgens de Common European Framework of Reference for Languages (CEF) van A1 tot B2. A1 staat voor een beginnende lezer, zie figuur 1, voor het verschil in leesniveau op basis van de CEF normering door Het Talencentrum van de Rijksuniversiteit Groningen Overgenomen van website: <http://www.rug.nl/talencentrum/overhetTalencentrum/CEF> (oktober 2011).

- A1. U kunt bekende woorden en eenvoudige zinnen lezen en begrijpen, bijvoorbeeld in briefjes, op posters en catalogi.
- A2. U kunt korte, eenvoudige teksten lezen. U kunt specifieke, voorspelbare informatie halen uit eenvoudige teksten, zoals advertenties, menu's, dienstregelingen. U kunt korte, eenvoudige, persoonlijke brieven begrijpen.
- B1. U kunt teksten begrijpen die hoofdzakelijk bestaan uit zeer frequente, dagelijkse of aan uw werk gerelateerde taal.
- B2. U kunt de beschrijving van gebeurtenissen, gevoelens en wensen in persoonlijke brieven begrijpen. U kunt artikelen en verslagen begrijpen over actuele onderwerpen waarin de schrijver stelling neemt, of een standpunt verwoordt. U kunt modern literair proza begrijpen.

Figuur 1. Een omschrijving van de vaardigheden behorende bij de verschillende niveaus van leesvaardigheid zoals aangegeven door de CEF.

Er waren in totaal zeven toetsniveaus waarbij er varianten zijn voor leerlingen met weinig leerervaring (L) en voor leerlingen met veel leerervaring (S) De leerlingen in deze steekproef maakten de toetsen 2S, 3L, 3S, 4L of 5L. De score op een test was een indicatie voor verschillende niveaus van leesvaardigheid, zie figuur 2.

De leerlingen kregen de toets aangeboden waarvan verwacht wordt dat die zo goed mogelijk aansloot bij hun niveau. Als de scores twijfel opriepen bestond de mogelijkheid om een hogere dan wel lagere toets af te nemen om zo meer zekerheid te krijgen over het niveau. Als leerlingen twee of meer toetsen moesten maken om het niveau te bepalen is de score van de laatst gemaakte toets mee genomen in dit onderzoek. Op basis van de uitslag werd er een toetsscore gegeven die kan oplopen van 'op weg naar A1' tot "A2". Deze toetsen waren in de periode september tot november afgenomen.

Voor dit onderzoek waren niet de ruwe scores van de Trajecttoets gebruikt maar het CEF leesniveau verdeeld over twee groepen. Een groep leerlingen met een leesniveau "Op weg naar A1" en "A1" en een groep leerlingen met leesniveau "Op weg naar A2" en "A2".

<i>Leesniveau CEF</i>	<i>Trajecttoets lezen</i>	<i>Score</i>
Op weg naar A1	3L	7-11
	2S	9-13
A1	3L	12-16
	2S	14-20
Op weg naar A2	3L	17-18
	4L	10-13
	5L	7-9
	2S	21-2
A2	4L	14-18
	5L	10-15
	3S	17-20

Figuur 2. Leesniveau volgens de CEF-normering en verschillende Trajecttoetsen met de scores die een indicatie zijn voor het leesniveau.

Het schoolniveau. Het schoolniveau was bepaald aan de hand van verschillende meetinstrumenten, de indruk van de docent en gegevens van vorige schoolervaringen, indien aanwezig. Leerlingen waren rechtstreeks uit het buitenland gekomen. Er was begonnen met een intake-toets onder andere bestaande uit een toets voor EMT, NDT, een gesprekje, woordjes laten lezen en de observatie van de Remedial Teacher van de school. Leerlingen waren gestart in fase 1 waarbij onderscheid was gemaakt tussen leerlingen die al een beetje Nederlands kunnen lezen en praten en leerlingen die analfabeet zijn. Na 6 tot 8 weken was er opnieuw getoetst en dan onder andere ook de Trajecttoets. Hierna waren zij doorgestroomd naar fase 2. Als zij in fase 2 voldoende voortgang hebben doorgemaakt waren zij doorgestroomd naar fase 3. Na fase 3 verlaten zij de school en gaan naar het reguliere voortgezet onderwijs. De school hanteert geen jaarklassensysteem maar leerlingen stromen

door naar de volgende fase op het moment dat er daarvoor voldoende criteria voor zijn.

Leerlingen kunnen niet langer dan twee jaar op deze school blijven.

Ontwikkeling van de taken

Voor de maze was in een pilot met ongeveer zeven leerlingen vanuit het voortgezet onderwijs bepaald welke teksten het meest aansluiten bij de doelgroep. Op basis van de uitkomsten van deze pilot was de keuze gemaakt om de teksten: “Sport en bewegen” en “Koninginnedag” te gaan gebruiken. Er was voor gekozen om zowel de score na één minuut als na twee minuten te meten. Uit eerder onderzoek in de USA door Espin, Wallace, Lembke, Campbell, en Long (2010) was gebleken dat de betrouwbaarheid van de maze bij 2, 3 en 4 minuten tussen de .80 en .88 lag. Voor gebruik in het onderwijs is een zo kort mogelijke tijd om een taak af te nemen het meest efficiënt. Ook zou er gescoord worden op goede antwoorden en op foute antwoorden. Bij drie foute antwoorden achter elkaar was het scoren gestopt en was gerekend tot het voorgaande goede antwoord.

Ook de woordenschattaken waren meegenomen in een kleine pilot. Omdat aan deze pilot Nederlandse leerlingen deelnamen vanuit het reguliere voortgezette onderwijs traden er, zoals verwacht, lichte plafond effecten op. Omdat de taken specifiek waren ontwikkeld voor leerlingen met een lagere dan gemiddelde leesvaardigheid zoals NT2 leerlingen was er vervolgens een pilot gedaan in maart 2011 op een school in een middelgrote stad in het westen van Nederland bij negentien NT2 leerlingen. De woordenschattaken waren gescoord op één, twee en drie minuten. Bij drie minuten waren er verscheidene leerlingen die aan het eind van de toets kwamen, het plafondeffect. Daarom was er voor gekozen om alleen de scores op één en twee minuten te gebruiken. Bij de woordenschattaak waren zowel de goede en de foute antwoorden geteld en ook was er onderscheid gemaakt tussen goed en fout gespelde woorden. Voor het onderzoek waren de scores voor de goede antwoorden gebruikt zonder onderscheid tussen goed en fout gespeld. Verschil tussen de scores met goed en fout gespelde antwoorden bleek niet significant van elkaar te verschillen.

Procedure

Voor alle meetinstrumenten waren docenteninstructies geschreven die exact beschrijven welke instructie de docent moet geven. Voorafgaande aan de daadwerkelijke afname was er een trainingsmiddag waarbij de mentoren uitleg hebben gekregen over CBM en de verschillende taken en konden oefenen met het afnemen van de taken. In verband met het verzoek van de school om in verschillende klassen verschillende taken af te nemen waren er setjes met daarin de verschillende taken per klas gemaakt.

De eerste meting was in maart en april 2011 gedaan. Er was door medewerkers van de Universiteit Leiden geobserveerd aan de hand van een van tevoren gemaakt observatie schema. Hierbij was voornamelijk gekeken naar de juistheid van de instructie en de tijdsbewaking. Als uit deze observaties was gebleken dat toetsen niet volgens de procedure zijn afgenomen waren de scores van deze klassen niet meegenomen in het onderzoek. Als de eerste meting goed was verlopen kan de mentor de tweede meting doen zonder observanten. In de periode mei en juni 2011 was de tweede meting gedaan. Voor dit onderzoek waren de gegevens van de tweede meting gebruikt omdat bij de tweede meting er meer leerlingen waren die de toetsen met zowel de onafhankelijke variabele, maze en woordenschat, hebben gemaakt en waarvan ook de criteriumvariabelen, leesniveau en schoolniveau, beschikbaar waren.

De toetsen waren nagekeken door studenten van de universiteit Leiden. De studenten hadden hiervoor een training gekregen. Tijdens de training was er gecontroleerd op betrouwbaarheid. Ook waren de toetsen nogmaals, als ze waren nagekeken, steekproefsgewijs door andere studenten gecontroleerd.

Analysetechnieken

Om de eerste onderzoeksvraag naar de betrouwbaarheid en de validiteit van de meetinstrumenten maze en woordenschattaak te kunnen beantwoorden was een Pearson's correlatie gedaan. Bij de woordenschattaak was als eerste gekeken of er verschil was tussen de scores met of zonder naar de spelling middels een t-toets. Voor de tweede en de derde onderzoeksvraag naar de validiteit van de maze en de woordenschat in relatie tot leesniveau en schoolniveau was een t-toets gedaan. De vierde onderzoeksvraag en vijfde onderzoeksvraag waren beantwoord met een stapsgewijze logistische regressie.

Resultaten

In het onderzoek wordt gekeken naar de betrouwbaarheid en de validiteit van de meetinstrumenten maze en woordenschat die gebruikt kunnen worden voor het voorspellen van leesvaardigheid. In tabel 1 staan de gemiddelde scores na 1 en 2 minuten van zowel de maze A en B en de woordenschattoets A en B. Bij de maze is de gemiddelde score na 2 minuten ongeveer twee maal zo groot als na 1 minuut. Bij de woordenschat is het verschil tussen 1 en 2 minuten ongeveer anderhalf keer zo groot.

Tabel 1

De gemiddelde scores, de standaardafwijking en de hoogste en laagste score op de A en B variant van de meetinstrumenten maze en de woordenschat na 1 en 2 minuten.

Variabele ^a	1 minuut				2 minuten			
	M	SD	Laagste	Hoogste	M	SD	Laagste	Hoogste
Maze A	11.56	3.98	1	20	21.93	7.47	1	39
Maze B	10.90	3.95	2	21	22.47	8.48	2	39
Woordenschat A	5.29	2.02	1	10	8.44	3.13	3	15
Woordenschat B	5.57	1.88	2	9	7.68	3.07	2	15

^a $n = 72$

Betrouwbaarheid

De eerste onderzoeksvraag was: “Wat is de (paralleltest) betrouwbaarheid van de twee meetinstrumenten?”. Deze vraag wordt onderzocht door de Pearson’s correlatie te berekenen tussen de A en B variant van zowel de maze als de woordenschat. De betrouwbaarheid is berekend voor de beide meetinstrumenten na 1 en na 2 minuten.

Tabel 2

De correlatie tussen de maze- en woordenschattest A en B bij de score na 1 en 2 minuten.

Variabele	1 minuut	2 minuten
	Correlatie	Correlatie
Maze	.62*	.74*
Woordenschat	.41*	.68*

^a $n = 72$. * $p < .01$, tweezijdig.

In tabel 2 staan de betrouwbaarheidscoëfficiënten voor A en B variant van de twee meetinstrumenten. De correlatie tussen de A en B variant van de maze is hoger dan de

woordenschat bij zowel 1 minuut ($r = .62$ respectievelijk $.41$) als bij 2 minuten ($r = .74$ respectievelijk $.68$). Bij beiden zijn de correlaties hoger bij 2 minuten dan bij 1 minuut.

Validiteit

De tweede onderzoeksvraag was: “Zijn de maze- en de woordenschattoets valide meetinstrumenten om het niveau van leesvaardigheid te kunnen meten?”. De validiteit is onderzocht door het verschil in gemiddelde scores van leerlingen op leesniveau 0 op de Trajecttoets NT2 te vergelijken met die van leerlingen op leesniveau 1 op de Trajecttoets NT2.

Voor deze analyse zijn de scores van 2 minuten op de maze en de woordenschat gebruikt omdat zij een hogere betrouwbaarheid hadden dan de scores van 1 minuut. Ook zijn de scores van de A en B variant bij elkaar opgeteld om een stabiel beeld te geven.

In tabel 3 staan de resultaten van de analyse voor de maze en de woordenschat. Leerlingen van niveau 1 scoren gemiddeld ongeveer 9 punten hoger op de maze, ongeveer 49, dan leerlingen van niveau 0, ongeveer 40. Dit verschil was significant. Op de woordenschat scoren leerlingen van niveau 1 gemiddeld ongeveer 4 punten, ongeveer 18, hoger dan niveau 0, ongeveer 14. Ook dit verschil was significant. Voor beide testen is er sprake van een ruim gemiddeld effect, respectievelijk Cohen’s d $.65$ en $.66$.

Tabel 3

Resultaten van de t-toets naar het verschil in de gemiddelden op de maze- en woordenschattoets bij de trajecttoets.

Variabele	Trajecttoets		t	p	Cohen’s d
	Leesnivea 0 ^a	Leesnivea 1 ^b			
	M (SD)	M (SD)			
Maze	39.50 (14.8)	48.79 (13.6)	- 2.77	.007	0.65
Woordenschat	14.24 (5.73)	17.82 (5.15)	- 2.79	.007	0.66

^a $n = 34$. ^b $n = 38$

De derde onderzoeksvraag was: “Zijn de maze en woordenschat valide meetinstrumenten om de fase van het onderwijs te kunnen bepalen waarin een leerling is geplaatst?”. De validiteit is

onderzocht door het verschil in gemiddelde scores van leerlingen in fase 2 te vergelijken met die van leerlingen in fase 3.

In tabel 4 staan de resultaten. Leerlingen uit fase 3 scoren gemiddeld ongeveer 16 punten hoger, ongeveer 56, op maze dan leerlingen uit fase 2, ongeveer 40. Dit verschil was significant. Op de woordenschat scoren leerlingen uit fase 3 gemiddeld ongeveer 4 punten hoger, ongeveer 19 dan uit fase 2, ongeveer 15. Dit verschil was ook significant. Bij de maze is er sprake van een groot effect en bij de woordenschat van een ruim gemiddeld effect, respectievelijk Cohen's d 1.28 en .78.

Tabel 4

Resultaten van de t-toets naar het verschil in de gemiddelden op de maze en woordenschat bij het schoolniveau.

Schoolniveau					
Variabele	Fase 2 ^a	Fase 3 ^b	t	p	Cohen's d
	M (SD)	M (SD)			
Maze	40.28 (14.27)	55.89 (9.71)	- 4.41	.001	1.28
Woordenschat	15.08 (5.78)	19.05 (4.31)	- 2.73	.008	0.78

^a $n = 53$. ^b $n = 19$

Maze en woordenschat gecombineerd

De vierde onderzoeksvraag was: “Zijn de maze en woordenschat samen een betere voorspeller voor het leesniveau dan apart?” en de vijfde onderzoeksvraag was: “Zijn de maze en woordenschat samen een betere voorspeller voor het schoolniveau dan apart?”. Beiden zijn onderzocht met behulp van een stapsgewijze logistische regressie. De logistische regressie laat zien dat de twee variabelen samen geen betere voorspelling geven dan apart. Dit geldt zowel voor het leesniveau als het schoolniveau. Voor beiden geldt, dat de maze als eerste is toegevoegd in de equation en onderzocht voor een significante bijdrage aan de variantie (Leesniveau, Wald statistiek 1.84 en $p = .18$; Schoolniveau, Wald statistiek 8.84 en $p = .003$). Het toevoegen van de woordenschat geeft ook geen significant proportie van de variantie (Leesniveau, Wald statistiek .57 en $p = .81$.; Schoolniveau, Wald statistiek 1.94 en $p = .16$).

Tabel 5

Aannemelijkheid van de overeenkomst van de score op de maze en de woordenschat met de score op leesniveau.

	X^2	ρ	R^2	Wald	ρ
Stap 1	7.42	.006*	.075		
Maze				6.44	.011**
Stap 2	2.01	.156	.094		
Maze en				1.84	.176
woorden-				1.94	.164
schat					

^a = Mc Fadden pseudo R-square. * $p < .01$; ** $p < .05$.

Tabel 6

Aannemelijkheid van de overeenkomst van de score op de maze, de woordenschat met het schoolniveau.

	X^2	ρ	R^{2a}	Wald	ρ
Stap 1	19.01	.0001*	.229		
Maze				11.93	.001*
Stap 2	.057	.812	.229		
Maze				8.84	.003*
Woorden-				.57	.812
schat					

^a = Mc Fadden pseudo R-square. * $p < .01$; ** $p < .05$. Tabel 5

Conclusie

We onderzochten de betrouwbaarheid en validiteit van de maze en de woordenschattoets om het leesniveau en schoolniveau te voorspellen van NT2-ers tussen de 12 en 18 jaar. We onderzochten ook de vraag of de maze en de woordenschat samen een betere indicatie geven dan apart. De resultaten laten zien dat beide varianten van de maze en de woordenschat betrouwbaar zijn en dat de maze en de woordenschat beide het leesniveau en het schoolniveau kunnen voorspellen maar samen geen betere indicatie geven zijn dan apart. De betrouwbaarheid van de maze en de woordenschat is zowel onderzocht bij één als bij twee minuten. Het gemiddelde aantal scores op de maze is na twee minuten ongeveer twee keer zo hoog als na één minuut. Het lijkt erop dat leerlingen gemiddeld in hetzelfde tempo doorlezen. Dit kan betekenen dat één minuut ook een betrouwbare meting is. Echter de betrouwbaarheid is hoger bij twee minuten dan bij één minuut.

Bij de woordenschattoets laat het resultaat iets anders zien. De verwachting is dat als de tijd twee maal zo lang is de score ook tweemaal zo hoog is, net als bij de maze, maar in het geval van de woordenschattoets is dit niet het geval. Ook bij de woordenschattoets is de betrouwbaarheid hoger bij twee minuten. De reden waarom de maze en de woordenschattoets apart een valide indicatie geven voor het leesniveau maar samen gemeten beide geen indicerende waarde meer laten zien wordt is te verklaren doordat de hoge correlatie tussen de maze en de woordenschattoets erop duiden dat ze beiden voor een groot gedeelte een zelfde construct meten wat ook wordt gemeten door de criteriumvariabele, het leesniveau. Mogelijke verklaring is dat het fonologisch bewustzijn zowel een betrouwbare voorspeller is voor de vaardigheid van het lezen van woorden maar ook een belangrijke factor bij begrijpend lezen (Lipka & Siegel, 2011).

De criteriumvariabele lijkt een ander construct te meten maar ook hier zijn de maze en de woordenschattoets geen samen geen betere voorspeller voor het schoolniveau.

Onderzoek in Amerika naar het gebruik van de maze als CBM-taak bij NT2-ers (Wiley & Deno, 2005) in het basisonderwijs heeft aangetoond dat de maze een betrouwbaar en valide meetinstrument is voor leesvaardigheid. Dit onderzoek laat zien dat de maze ook voor NT2-ers binnen het voortgezet onderwijs in Nederland een betrouwbaar en valide meetinstrument is om leesvaardigheid te meten.

Beide instrumenten kunnen in het voortgezet onderwijs bij NT2-ers gebruikt worden om het niveau van leesvaardigheid te meten. Beiden kunnen ook gebruikt worden om beslissingen die moeten worden genomen voor instroom, doorstroom en uitstroom van deze leerlingen te ondersteunen.

De taken zijn goed bruikbaar in het onderwijs omdat zij op elk willekeurig moment kunnen worden afgenomen, ze zowel individueel als klassikaal kunnen worden afgenomen en het weinig tijd kost om ze af te nemen.

Als vervolgonderzoek kan aantonen dat de maze en woordenschat als CBM-taken gevoelig zijn voor groei van leesvaardigheid dan kunnen deze taken in het voortgezet onderwijs aan NT2-ers gebruikt worden als onderdeel van een leerlingvolgsysteem.

Beperkingen van het onderzoek

De steekproef was klein en de proefpersonen volgden onderwijs op één school. De gegevens waren gescoord in één meetmoment waardoor alleen de paralleltest betrouwbaarheid is gemeten en niet de test-hertest betrouwbaarheid. Bij de validiteit van de maze en de woordenschat als indicator van leesniveau zijn kanttekeningen bij te plaatsen. Ten eerste is de Trajecttoets geen gestandaardiseerde toets volgens de COTAN richtlijnen. Ten tweede was het nodig om de gegevens vanuit de Trajecttoets om te zetten in een dichotome variabele. Hierdoor was de enige voorspelling die gedaan kan worden, op basis van de scores gemeten met de maze of de woordenschat, of de leerlingen een hoog of laag leesniveau hadden op de Trajecttoets.

Er was in het onderzoek geen rekening gehouden met de moedertaal als variabele. Dit kan van invloed zijn geweest op de resultaten. Door een andere testcultuur uit het land van herkomst of onbekendheid met testen kan ervoor zorgen dat deze leerlingen onvoldoende eraan gewend zijn om test zo snel mogelijk te maken (Okawa, 2008).

Toekomstig onderzoek

In dit onderzoek is de betrouwbaarheid en validiteit van de maze en de woordenschat onderzocht, een eerste stap in de ontwikkeling van een leerlingvolgsysteem voor leerlingen die Nederlands als tweede taal moeten leren.

Een volgende stap is het onderzoeken van de gevoeligheid voor groei zodat de ontwikkeling van leesvaardigheid gevolgd kan worden. Als eerste zal een grotere steekproef nodig zijn met leerlingen die Nederlands als tweede taal moeten leren van verschillende scholen verspreid over Nederland. Ten tweede zijn meerdere meetmomenten door de tijd heen nodig zal om de betrouwbaarheid over tijd te kunnen meten. Ten derde zal het leesniveau gemeten moeten worden met een gestandaardiseerde (COTAN, 2009) toets waarbij de variabele op interval niveau gemeten kan worden. Als laatste kan de invloed van de moedertaal op het niveau van leesvaardigheid onderzocht worden. Om in het algemeen de mogelijkheden voor toekomstig onderzoek te verbeteren, en de bruikbaarheid voor docenten te vergroten, is het aan te raden de taken te digitaliseren (Stecker, Fuchs, & Fuchs, 2005).

De bedoeling is uiteindelijk om een leerlingvolgsysteem ontwikkelen die de groei en ontwikkeling meten van Nederlandse leesvaardigheid van leerlingen die in korte tijd Nederlands moeten leren. Op grond van de gemeten ontwikkeling kunnen docenten hun lesprogramma op individueel, klas of school aanpassen (Pierce, McMaster, & Deno, 2010). In de toekomst kunnen de meetinstrumenten wellicht ook ingezet worden om onderscheid te kunnen maken tussen leerlingen die lager scoren dan verwacht mag worden omdat ze de Nederlandse taal nog onvoldoende beheersen en leerlingen bij wie er sprake is van onderliggende tekorten op het gebied van leesvaardigheid (Deno & Wiley, 2005).

Referenties

- Boers, T. (2008). Hoe schat een NT2-cursist zijn eigen taalniveau in? En wat is de inschatting van de docent? *Levende Talen Tijdschrift*, 9(1), 27-32.
- Brown-Chidsey, R., Johnson, Jr., P., & Fernstrom R. (2005). Comparison of Grade-Level Controlled and Literature-Bases Maze CBM Reading Passages. *School Psychology Review*, 34(3), 387-394.
- Busch, T. W., & Espin, C. A. (2003). Using Curriculum-Based Measurement to Prevent Failure and Assess Learning in the Central Content Areas. *Assessment for Effective Intervention*, 28(3&4), 49-58. doi: 10.1177/073724770302800306
- Centraal Bureau voor de Statistiek (2011). *Immi- en emigratie naar geboorteland, leeftijd (31 december) en geslacht*. Gevonden op <http://cbs.nl>.
- De Klein, P. (2001). *Woordenboek Nederlands*. Groningen: Noordhoff Uitgevers B.V. ISBN: 9789001470715
- Deno, S. L. (1985). Curriculum-Based Measurement: The Emerging Alternative. *Exceptional Children*, 52(3), 219-232.
- Deno, S. L. (2003). Curriculum-Based Measures: Development and Perspectives. *Assessment for Effective Intervention*, 28(3&4), 3-12. doi: 10.1177/073724770302800302
- Deno, S. L. (2003). Development in Curriculum-Based Measurement. *Journal of Special Education*, 37(3), 184-192. doi: 10.1177/00224669030370030801
- Deno, S. L., & Wiley, H. I. (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.
- Espin, C. A., & Deno, S. L. (1994-1995). Curriculum-Based Measures for Secondary Students: Utility and Task Specificity of Text-Based Reading and Vocabulary Measures for Predicting Performance on Content-Area Tasks. *Diagnostique*, 20(1-4), 121-142.
- Espin, C., Wallace, T., Lembke, E., Campbell, H., & Long J. D. (2010). Creating a Progress-Monitoring System in Reading for Middle-School Students: Tracking Progress Toward Meeting High-Stakes Standards. *Learning Disabilities Research & Practice*, 25(2), 60-75.
- Foegen, A., Espin, C. A., Allinder, R. M., & Markell, M. A. (2001). Translating Research Into Practice: Preservice Teachers' Beliefs About Curriculum-Based Measurement. *The Journal of Special Education*, 34(4), 226-236.
- Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic Distinctions Between Instructionally Relevant Measurement Models. *Exceptional Children*, 57(5), 488-500.
- Huijgen, M., Verburg, M. (1996). *Van Dale Basiswoordenboek van de Nederlandse taal*. Utrecht: Van Dale. ISBN: 9789066488687.
- Hosp, M. K., Hosp, J. L. & Hotel, K. W. (2007). *The ABC's of CBM. A Practical Guide to Curriculum-Based Measurement*. New York: Guilford Press.
- Inspectie van Onderwijs (2004). *Het onderwijs eerste opvang anderstaligen*. Utrecht: Inspectie van het Onderwijs. Gevonden op <http://www.onderwijsinspectie.nl/actueel/publicaties/Het+onderwijs+eerste+opvang+anderstaligen.html>

- Inspectie van het Onderwijs (2009). *Bureau ICE; Onderzoek kwaliteit exameninstrumenten van examenleveranciers 2009. Rapportage kwaliteit exameninstrumenten bij leverancier Bureau ICE*. Gevonden op <https://zoek.officielebekendmakingen.nl/blg-109703.pdf>
- Inspectie van het Onderwijs (2009). *Onderzoek kwaliteit exameninstrumenten van Examenleveranciers 2009. Rapportage kwaliteit exameninstrumenten bij leverancier*. Utrecht: Inspectie van Onderwijs.
- Koda, K., & Zehler, A. M. (2010). Introduction; Conceptualizing reading universals, cross-linguistic variations, and second language literacy development. In K. Koda & A. M. Zehler (Eds.), *Learning to Read across Languages: Cross-Linguistics Relationships in First- and Second-Language Literacy Development*. (pp. 1-9). New York and London: Routledge.
- Lindhout, P. (2009). *Taalproblemen bij BRZO- en ARIE bedrijven, een onderschat gevaar? Een verkennend onderzoek naar het raakvlak tussen taalproblemen en zware ongevallen. Bijlage 10 bij thesis; Resultaten deelonderzoek 4 – Methode voor het meten van leesbaarheid van documenten (Doctoraal scriptie, TU-Delft/MoSHE)*.
- Lipka, O., & Siegel, S. (2011). The development of reading comprehension skills in children learning English as a second language. *Read Writ*. Online publicatie. doi: 10.1007/s11145-011-9309-8/Springer Science+Business Media B.V. 2011.
- McMaster, K. L., Wayman, M. M. & Cao, M. (2006). Monitoring the Reading of Secondary-Level English Learners: Technical Features of Oral Reading and maze Tasks. *Assessment for Effective Intervention*, 31(17), 17-31. doi: 10.1177/073724770603100402
- Nova College; Leercentrum voor Nieuwe Nederlanders. Jaarplan 2009 – 2010 van <http://www.johandewittcollege.nl>
- Nova College; Leercentrum voor Nieuwe Nederlanders. Leerplan Nederlands als Tweede Taal 2007-2011 van <http://www.johandewittcollege.nl> .
- OECD, Organization for Economic Co-operation and Development. (1997). *Literacy Skills for the Knowledge Society*. Parijs: OECD.
- Okawa, J. B. (2008). Considerations for the Cross-cultural Evaluation of Refugees and Asylum Seekers. In L.A. Suzuki & J.G. Ponterotto (Eds.), *Handbook of Multicultural Assessment* (pp. 165 - 195). San Francisco CA: Jossey-Bass.
- Paradis, J., Genesee, F., & Crago M. B. (2011). Schooling in a Second Language. In S. F. Warren & M. E. Fey (Eds.), *Dual Language Development & Disorders; A Handbook on Bilingualism & Second Language Learning* (pp. 164-196). Baltimore: Brookes.
- Pierce, R. L., McMaster, K. L., & Deno, S. L. (2010). The Effects of Using Different Procedures to Score Maze Measures. *Learning Disabilities Research & Practice*, 25(3), 151-160.
- Shin, J., Deno, S. L., & Espin, C. A. (2000). Technical Adequacy of the Maze Task for Curriculum-Based Measurement of reading Growth. *Journal of Special Education*, 34(3), 164-172. doi: 10.1177/002246690003400305
- 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)* (pp. 284-302). New York, London: The Guilford Press.
- Stecker, P. M., Fuchs, L. S., & Fuchs, D. (2005). Using Curriculum- Measurement to Improve Student Achievement: Review of Research. *Psychology in the Schools, 42*(8), 795-819. doi: 10. 1002/pits.20113
- Tichá, R., Espin, C. A., & Wayman, M. M. (2009). Reading Progress Monitoring for Secondary-School Students, Validity, and Sensitivity to Growth of Reading-Aloud and Maze-Selection Measures. *Learning Disabilities Research and Practice, 24*(3), 132-142.
- Vernooy, D. (2008). Als leerlingen in het voorgezet onderwijs niet goed lezen.
- Verspoor, M., & Cremer, M. (2008). Research on foreign-language teaching and learning in the Netherlands (2002-2006). *Language Teaching, 41*(2), 183-211. doi: 10.1017/S0261444807004879
- Wayman, M. M., Wallace, T., Wiley, H. I., Tichá, R., & Espin, C. A. (2007). Literature Synthesis on Curriculum-Based Measurement in Reading. *The Journal of Special Education, 41*(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. doi: 10. 1177/07419325050260040301