

Leesbegrip in de mazetaak:

Aap/ Zie /-Top jij de inconsistentie?

Naam: Karlijn Dievelaar

Studentnummer: s13828378

Masterspecialisatie: Orthopedagogiek, Universiteit Leiden

Begeleider: Prof. Dr. C.A. Espin

Tweede lezer: n.t.b.

Datum: 19-07-2021

Voorwoord

Hierbij mijn masterscriptie over leesbegrip in de mazetaak. Als (instroom)student orthopedagogiek startte ik dit project zonder enige kennis over leesvaardigheidsontwikkeling of begrijpend lezen. Deze uitdaging ben ik aangegaan en gaandeweg heb ik de materie eigen gemaakt, door veel te lezen, vragen te stellen en kritisch alle informatie te analyseren. Bij de totstandkoming van de masterscriptie heb ik hulp gehad van een aantal mensen die ik wil bedanken.

First my supervisor, Christine Espin. With endless patience and enthusiasm, you guided me through the project. Answering all my clueless questions, your honest but constructive feedback and the many times you helped me getting back on track.

Daarnaast wil ik mijn familie en vrienden bedanken voor alle steun en motiverende woorden, een aantal mensen in het bijzonder. Mijn vader en Nicole, voor jullie flexibiliteit en hulp om dit stuk tot een leesbaar geheel te maken. Mijn moeder en Ton, voor het luisteren en opvangen na een lange studiedag. Kristel, mijn studiemaatje die me dag in dag uit motiveerde. En Kirsten voor je kritische wetenschappelijke blik en fijne feedback. Tot slot Yorick, voor je geduld waarmee je al mijn stukken nalas, met me mededacht en mijn twijfels aanhoorde. Maar bovenal voor het eindeloze vertrouwen in mij, zelfs als ik dit zelf af en toe kwijt was.

Ik ben trots op het resultaat en dit deel ik graag met jullie. Dus voor alle docenten, familie, vrienden en andere geïnteresseerden: Veel leesplezier!

Abstract

In recent years the level of reading comprehension among Dutch students has been decreasing (Gubbels, van Langen, Maassen & Meelissen, 2019). Monitoring students' reading progress using the CBM maze task might contribute to a solution. However, there are criticisms of the CBM maze task. Muijselaar and colleagues (2016) claim that the maze task relies on code-related skills, and raising questions about whether the task is an adequate tool for evaluating reading comprehension. This study investigates whether these concerns are legitimate, with the research question: Does the maze task reflect reading comprehension? This question is examined via an experimental design based on the inconsistency paradigm. By using eye tracking data, reading times are compared between consistent and inconsistent texts to determine whether readers notice inconsistencies while completing the maze task. Results show that reading times of a consistent and inconsistent text do not differ when inconsistencies are presented outside a maze selection. However, reading times do increase when inconsistencies are presented in a maze selection, suggesting that completing the maze task reflects not only code-related, but also reading comprehension skills. Follow-up research is indicated to gain more insights in which conditions or circumstances reading comprehension is reflected best in the maze task.

Keywords: mazetaak, leesbegrip, Curriculum-Based Measurement (CBM)
inconsistentie paradigma, eye tracker, leestijd, mazeseselectie

Inleiding

Het leesvaardigheidsniveau van de Nederlandse jongeren gaat de laatste jaren steeds harder achteruit (Gubbels, van Langen, Maassen & Meelissen, 2019). In grootschalig onderzoek (PISA-2012) scoorde Nederlandse vijftienjarige nog bovengemiddeld met een zevende plaats van de 65 deelnemende landen (Kordes, Bolsinova, Limpens, & Stolwijk, 2013). Sindsdien heeft een daling ingezet en in het PISA rapport 2018 scoorde Nederland vergelijkbaar met het gemiddelde van OESO-landen en significant lager dan van de EU-landen. De achteruitgang is het sterkst voor de begripsvaardigheden ‘begrijpen’ en ‘evalueren en reflecteren’, waar Nederland beneden gemiddeld presteert. In dit rapport kwam ook naar voren dat bijna 24% van de jongeren over een dusdanig laag niveau beschikt, dat ze risico lopen op laaggeletterdheid (Gubbels et al., 2019). Dit percentage is in zes jaar tijd aanzienlijk gestegen, in 2012 was dit nog 13,8% (Kordes et al., 2013).

Een goede leesvaardigheid is van groot belang voor de ontwikkeling van kinderen. Het is een essentiële vaardigheid in het academisch functioneren, dat door het hele curriculum verweven zit (Moin, Magiera & Zigmond, 2009; Allington, 2007). Bij de start van het basisonderwijs ligt de focus op het leren lezen, wat gedurende de tijd verschuift naar lezen om te leren (Chall, 1996; Ally & Deshler, 1979). Leerlingen moeten veel (vakinhoudelijke) informatie opnemen door middel van geschreven taal (Moin et al., 2009). De leerlingen die moeite hebben met leesvaardigheid hebben daardoor doorgaans ook meer moeite met andere schoolvakken. Het begrijpen van de lesstof is voor hen extra complex, omdat zij minder goed in staat zijn relevante informatie uit de tekst te verkrijgen (Scruggs & Mastrioueru, 2007; Yovanoff, Duesbery, Alonzo & Tindel, 2005).

De gevolgen van onderontwikkelde leesvaardigheid kunnen ook doorwerken op lange termijn. We leven in een toenemende informatiegestuurde samenleving en beheersing van een goede leesvaardigheid is hierin cruciaal (Gubbels et al., 2019; Rapp, van den Broek, McMaster, Kendeou & Espin, 2007). Niet alleen op professioneel vlak, maar ook op sociaal,

LEESBEGRIP IN DE MAZETAAK

economisch of persoonlijk vlak wordt beroep gedaan op taalvaardigheid. Wie een achterstand heeft, loopt het risico minder goed deel te kunnen nemen aan onze maatschappij (Gubbels et al, 2019; Conoyer et al., 2007).

Voor de aanpak van leesproblemen is gedegen leesonderwijs belangrijk. Op beleidsniveau wordt dit erkend en ministers van Engelshoven en Slob (2019) roepen op tot een leesoffensief ter bevordering van de leesontwikkeling van kinderen in het primair en voortgezet onderwijs. Ook in wetenschappelijke onderzoeken is veel aandacht voor leesontwikkeling op jonge leeftijd en wordt bijvoorbeeld de relevantie van het monitoren van de leerlingenvoortgang benadrukt, zodat vroegtijdige interventie ingezet kan worden (e.g., Espin, Wallace, Lembke, Campbell & Long, 2010; Fuchs & Fuchs, 1992; Deno, 1985).

Curriculum Based Measurement (CBM)

Een veelgebruikte methode voor het monitoren van leerlingen in Nederland is het Cito Volgstelsysteem. Met deze methode maakt een leerling in het primair of voortgezet onderwijs jaarlijks één of enkele toetsen om leesvaardigheid te meten en dit te vergelijken met leeftijdsgenoten (CITO, 2021). Met de dalende cijfers van het leesvaardigheidsniveau van Nederlandse jongeren, is intensievere monitoring gewenst. Een methode die hierop een goede aanvulling kan zijn, is Curriculum Based Measurement (CBM) (Deno, 1985). Deze vorm van formatieve evaluatie wordt veel ingezet in de Verenigde Staten (e.g., FastBridge Learning, 2015). Op gestandaardiseerde wijze wordt het effect van instructies of interventies gemeten in het onderwijs. De taken zijn praktisch en efficiënt opgebouwd. De instructies zijn simpel, de afnameduur is kort (ca. één tot drie minuten) en het nakijken kan snel gebeuren. Dit maakt het mogelijk om de taken regelmatig, bijvoorbeeld wekelijks af te nemen (Deno, 1985, 2003). CBM werkt aan doelen op korte en lange termijn. Zo kunnen interventies snel worden bijgesteld als de korte termijn doelen niet behaald worden en op lange termijn kan de

LEESBEGRIP IN DE MAZETAAK

ontwikkeling nauwkeurig worden bijgehouden door de vele meetmomenten (Hosp, Hosp & Howell, 2007; Deno 2003).

Mazetaak

Een veelgebruikte CBM taak om leesvaardigheid te meten is de mazetaak. Tijdens deze taak lezen leerlingen teksten waarbij in elke zin het zevende woord is weggelaten. Dit woord is vervangen door een keuzemoment dat bestaat uit drie opties, de zogenaamde mazeselectie. Eén van de woorden is passend in de zin, de andere twee woorden zijn incorrect en zijn bedoeld als afleiding. De afleiders mogen niet te veel op het correcte antwoord lijken en uit maximaal twee letters meer of minder bestaan. Het is de taak van de leerling om het goede woord te selecteren (Espin & Foegen, 1996; Fuchs & Fuchs, 1992; Wayman, Wallace, Wiley, Tichá & Espin, 2007). Het is een gestandaardiseerde methode waarbij leerlingen enkele minuten in stilte de tekst lezen en zelf de antwoorden omcirkelen. Dit maakt de test praktisch goed uitvoerbaar en geschikt voor regelmatige afname (Chung, Espin & Steverson, 2018; Deno, Mirkin & Chiang, 1982; Conoyer et al., 2017).

In diverse studies zijn de psychometrische eigenschappen van de mazetaak onderzocht (e.g. Tzivinikou, Tsolis, Kagkara & Theodosiou, 2020; Shin & McMaster, 2019). Validiteitsonderzoek is veelal uitgevoerd met andere (lees)testen als criteriumvariabele. In een recente literatuurreview concludeerden Tzivinikou en collega's (2020) dat de maze over het algemeen hoge correlaties heeft met de andere leestesten, wat wijst op een goede validiteit. Bovendien liet de taak groei van leesontwikkeling over tijd zien (Tzivinikou et al., 2020). De eerdere literatuurreview van Wayman en collega's (2007) kwam tot dezelfde conclusie. Hier werden redelijk tot goede validiteitscoëfficiënten gerapporteerd, variërend tussen de .65 en .86 (Wayman et al., 2007). Ook tonen diverse onderzoeken een goede betrouwbaarheid aan, wat meestal getoetst is aan de hand van de *alternate-form* analyse (e.g. Espin, et al., 2010; Shin Deno & Espin, 2000).

LEESBEGRIP IN DE MAZETAAK

Eén van deze onderzoeken is uitgevoerd door Fuchs en Fuchs (1992). Dit is een longitudinaal onderzoek uit waarin basisschoolleerlingen achttien weken lang twee digitale teksten per week lazen. Ze vonden een sterke correlatie van .77 met de Scholastic Aptitude Test (SAT) en de maze bleek sensitief voor groei gedurende de tijd. Bovendien rapporteerden leraren en leerlingen een hoge tevredenheid van de mazetaak.

Een ander onderzoek komt van Jenkins en Jewell (1993), die de mazetaak vergeleken met de Metropolitan Achievement Test (MAT) en Gates Mac-Ginitie Reading Tests. Ze vonden correlaties tussen de .65 en .75 en deze bleven in de onder- en bovenbouw van de basisschool ongeveer gelijk.

Shinn en collega's (2000) onderzochten 43 basisschoolleerlingen (onderbouw) en vonden een goede predictieve validiteit met de California Achievement Test (CAT) als criteriumvariabele. Ook was er sprake van een significante groei in scores (1.07 goede antwoorden) per maand. De groei was zowel op groepsniveau als individueel niveau aanwezig. Ook werd een goede betrouwbaarheid gerapporteerd, gemiddeld .83 met één maand interval tussen de afnames van de verschillende teksten (Shinn et al., 2000).

Espin en collega's (2010) onderzochten ook onderbouw basisschoolleerlingen, waarbij ze onderscheid maakten in lengte van afname (één, twee of drie minuten) en nakijkmethode (aantal correcte antwoorden of aantal correct minus incorrecte antwoorden). Als criteriumvariabele gebruikten zij de Minnesota Basis Standards Test (MBST). De *Alternate-form* betrouwbaarheidscoëfficiënten waren hoog (.79-.96). De predictieve validiteit varieerde van .75 tot .81. De nakijkmethode bleek weinig invloed te hebben op de resultaten. Qua lengte komt de drie minuten afname als beste naar voren. De mazetaak blijkt bovendien sensitief voor groei (Espin et al., 2010).

De studie van Espin en collega's (2010) had wel wat limitaties. Dit was voor Tichá , Espin en Wayman (2009) de reden om het onderzoek te reproduceren, waarbij de beperkingen

LEESBEGRIP IN DE MAZETAAK

aangepakt konden worden¹. Zij voerden dit uit onder middelbare scholieren, met dezelfde passages als de studie van Espin en collega's (2010). De resultaten waren vergelijkbaar waarbij ze een *alternate-form* betrouwbaarheid van .79 tot .91 en predictieve validiteit tussen de .80 en .88. vonden. Ook was er een significante groei in de mazetaak met 1.29 goede antwoorden per week (Tichá et al., 2009).

Onderzoek naar Nederlandse middelbare scholieren (Chung et al., 2017) toonde ook groei aan, al was deze een stuk kleiner (0.16 goede antwoorden per week). Wel hingen zowel de groei als het niveau van de maze samen met de groei en het niveau van de CITO-VVO. De *alternate-form* betrouwbaarheid scoort in dit onderzoek tussen de .67 en .83.

Muijselaar, Kendeou, de Jong en van den Broek (2016) keken met name naar de vaardigheden waar de mazetaak een beroep op doet. Zij onderzochten specifiek de benodigde inzet van decodeervaardigheden en taalbegrip tijdens het maken. Ze beargumenteerden vanuit eerder onderzoek dat de uitslag van de maze een weerspiegeling is van decodeervaardigheden, maar trokken in twijfel of het ook een goede voorspeller is van taalbegrip. Ze vergeleken dit met de voorspellende waardes van de Gates-MacGinitie reading comprehension test en legden relaties met decodeer- en taalbegripsvaardigheden. De maze en Gates-McGinitie test waren sterk aan elkaar gerelateerd (.69-.79). Als naar de afzonderlijke vaardigheden wordt gekeken, bleken de correlaties met taalbegrip (.49-.65) lager dan met decoderen (.50-.88). Hieruit volgde de conclusie: de mazetaak doet meer beroep op codeervaardigheden dan op taalbegripsvaardigheden. Vervolgens riepen zij de vraag op of maze wel een geschikt middel is om begrijpend lezen bij leerlingen te evalueren.

Begrijpend lezen

Begrijpend lezen is een complexe vaardigheid waar verschillende cognitieve processen aan ten grondslag liggen (e.g. Kendeou, van den Broek, Helder & Karlsson, 2014; Hannon,

¹ In het onderzoek van Tichá et al. (2009) wordt verwezen naar een ongepubliceerde versie van Espin et al. (2010), die is ingezonden voor publicatie in 2009. Dit biedt mogelijk een verklaring voor de jaartallen.

LEESBEGRIP IN DE MAZETAAK

2012). Om een tekst te kunnen begrijpen, moet de lezer allereerst over decodeervaardigheden beschikken, zoals het herkennen en onderscheiden van letters en klanken (Perfetti, Yang & Schmalhofer, 2008). Daarnaast is taalbegrip nodig (Verhoeven & Perfetti, 2008; Gough & Tunmer, 1986). Hier gaat het om meer dan alleen het begrijpen van de zinnen of woorden, want het doel is het begrijpen van de gehele tekst en daar vervolgens een betekenis aan te geven (Kintsch, 1998; Dijk & Kintsch, 1983). In de betekenisgeving staat interactie tussen de tekst, de context en de lezer centraal (McNamara & Magliano, 2009). Bij begrijpend lezen gaat het dus niet alleen over de informatie in de tekst zelf, maar ook de kennis, achtergrond en eerdere ervaringen van de lezer (Duke & Pearson, 2002).

Als er sprake is van succesvol tekstbegrip maakt de lezer een samenhangende mentale representatie van de opbouw van een tekst. Dit wordt een situatiemodel genoemd (Kintsch & van Dijk, 1978). Informatie uit de gehele tekst en achtergrondkennis van de lezer komen samen en vormen een coherent geheel, dat steeds opnieuw bijgewerkt wordt tijdens het lezen (Albrecht & O'Brien, 1993). Om een samenhangende representatie op te bouwen, moeten onder andere verbanden worden gelegd en coherentie gedurende de tekst gemonitord worden (Helder, van den Broek, van Leijenhorst & Beker, 2013; Duke & Pearson, 2002). Bij coherentie-monitoring let de lezer op onderbrekingen, onjuistheden en inconsistenties in een tekst. Dit is cruciaal bij tekstbegrip, omdat de lezer zijn of haar leesgedrag dan kan aanpassen. Door bijvoorbeeld achtergrondkennis toe te passen, terug te kijken of stukken opnieuw te lezen, kan de lezer coherentie herstellen (Duke & Pearson, 2002). Worden onderbrekingen in coherentie niet herkend, dan zal dit de samenhang van de mentale representatie - en daarmee het leesbegrip - verstoren. Dit maakt de mate van coherentie-monitoring een fundamenteel onderdeel van het leesbegrip (Helder, van Leijenhorst & van den Broek, 2016).

LEESBEGRIP IN DE MAZETAAK

Inconsistentieparadigma

Een goede manier om coherentie monitoring te meten is met het inconsistentieparadigma (Albrecht & O'Brien, 1993). Dit paradigma is gebaseerd op de aanname dat tijdens het begrijpend lezen inconsistenties in de tekst wordt opgemerkt. Albrecht en O'Brien (1993) lieten participanten narratieve teksten op een computer lezen, die ze zin voor zin aangeboden kregen. Een deel van de teksten bevatte een inhoudelijke inconsistentie, waarbij informatie in het begin van de tekst en eind van de tekst elkaar tegenspraken. Een voorbeeld is dat Mary als vegetariër geïntroduceerd wordt, maar later een cheeseburger bestelt. Ook was er een consistente variant van hetzelfde verhaal, waarbij het verhaal wel helemaal klopte. De discrepantie kwam pas aan het eind van de tekst naar voren, wat als doel had het langetermijngeheugen te activeren. Voorafgaand aan het lezen, kregen de participanten instructie om op hun eigen tempo te lezen en te letten op begrip. Op het eind van elke tekst werd ze een vraag gesteld over de inhoud van de tekst, maar niet specifiek gevraagd naar de inconsistentie (Albrecht & O'Brien, 1993). In het paradigma wordt de inconsistentiedetectie vastgesteld aan de hand van leestijd van de targetzin. In de targetzin wordt de tegenstrijdige informatie gepresenteerd en kan de inconsistentie herkend worden. De leestijd van de targetzin in een inconsistent verhaal kan worden vergeleken met de leestijd van de targetzin in een consistent verhaal. Hieruit volgt een inconsistentie-effect: vaardige (volwassen) lezers doen langer over het lezen van de zin in een inconsistent verhaal dan een consistent verhaal. Zij merken de coherentieonderbreking op en hebben meer tijd nodig om dit te verwerken (O'Brien, Cook & Gueraud, 2010; Gerrig & O'Brien, 2005).

Uit onderzoek blijkt dat het opmerken van inconsistenties samenhangt met de mate van leesbegrip (Oakhill, Hartt & Samols, 2005). Bij lezers met minder goed tekstbegrip wordt de inconsistentie opgemerkt als de informatie in opvolgende zinnen gepresenteerd wordt, maar gaat dit minder goed als er meerdere zinnen tussendoor moeten worden gelezen. Bij

LEESBEGRIP IN DE MAZETAAK

lezers met goede begripsvaardigheden treedt in beide gevallen het inconsistentie-effect op. Inconsistenties worden dus beter opgemerkt door lezers met een goed tekstbegrip (Long & Chung, 2001; van der Schoot, Rijntjes & van Lieshout, 2012).

Helder en collega's (2016) onderzochten coherentie-monitoring doormiddel van het inconsistentieparadigma. Zij onderzochten bij lezers met sterke en zwakke begripsvaardigheden of zij inconsistenties detecteerden en dit verwerkten in de mentale representatie van de tekst. Ze gebruikten twee groepen participanten: een groep van acht- en negenjarigen en een groep van tien- en elfjarigen (Helder et al., 2016). Hiervoor is gekozen omdat hiertussen de omslag van 'leren lezen' naar 'lezen om te leren' plaatsvindt (Chall, 1996; Helder et al., 2016). De jongere groep houdt zich voornamelijk bezig met basisvaardigheden zoals decoderen en woordenschat, waarbij minder aandacht is voor tekstbegrip. Bij de oudere groep ligt de nadruk op de ontwikkeling van begrijpend lezen, waarbij wordt gericht op de informatie uit de tekst (Best, Floyd & McNamara, 2008).

De participanten lazen 32 verhalen, waarvan de helft inconsistent was. Ze kregen steeds één zin tegelijkertijd op een beeldscherm te zien en moesten op de spatiebalk klikken om naar de volgende zin te gaan. Zo werd de leestijd per zin geregistreerd voor het meten van mogelijk inconsistentie-effect. De lezers beantwoordden eind van het verhaal steeds de vraag: 'Klopt het verhaal?' met ja of nee. Zo combineerden Helder en collega's (2016) online en offline methodes om respectievelijk detectie en verwerking van inconsistenties te meten.

De resultaten bevestigden het concept van het inconsistentieparadigma, waarbij de leestijd voor de targetzin hoger lag bij inconsistente verhalen dan bij consistente verhalen. Zij concludeerden daarbij dat de online inconsistentiedetectie plaatsvindt bij beide leeftijdsgroepen en bij zowel zwakke als sterke (begrijpend) lezers. Offline gegevens toonden dat oudere kinderen en lezers met sterkere begripsvaardigheden beter in staat waren de consistentie van een tekst te beoordelen (Helder et al., 2016).

LEESBEGRIP IN DE MAZETAAK

Huidige studie

Zoals geïllustreerd in voorgaande studies, wordt het inconsistentieparadigma veelal gebruikt om begripsprocessen bij kinderen of volwassenen te onderzoeken (e.g. Helder et al., 2016; Albrecht & O'Brien, 1993). In deze studie wordt het ingezet met een ander doel. In plaats van begripsprocessen te onderzoeken, gaat het om inzicht in de eigenschappen van de mazetaak. Bekeken wordt of een inconsistentie-effect naar voren komt tijdens het maken van de taak. Dit suggereert dat lezers, naast het begrijpen van de losse woorden en zinnen, ook een samenhangende mentale representatie opbouwen van de tekst (Helder et al., 2013; Duke & Pearson, 2002). Indien dit het geval is, betekent dit dat niet alleen decodeervaardigheden worden ingezet bij het maken van de mazetaak, maar er ook sprake is van begrijpend lezen.

Aanleiding om onderzoek te doen naar de weerspiegeling van leesbegrip in de mazetaak komt vanuit het artikel van Muijselaar en collega's (2016), waarin zij kritisch zijn op de werking van de taak. De mazetaak wordt vaak gepresenteerd als een formatieve taak voor leesbegrip (Muijselaar et al., 2016) en het is belangrijk om te achterhalen of de scores daadwerkelijk een goede weerspiegeling zijn van het niveau van leesbegrip. Dit is allereerst relevant om te weten omdat in het onderwijs bij voorkeur valide interventies en meetinstrumenten worden ingezet. Zeker gezien de dalende trend in niveau van leesvaardigheid en begripsvaardigheid in het bijzonder (Gubbels et al., 2019), is het belangrijk om onderzoek te doen naar geschikte testen of methoden voor leesonderwijs. De maze zou een goede bijdrage kunnen leveren, maar daarvoor is het wel belangrijk te weten of de taak geschikt is voor de evaluatie van leesbegrip.

Een ander doel van deze studie, is een bijdrage leveren aan de discussie binnen het academische veld. Naast de eerder besproken onderzoeken naar de validiteit en betrouwbaarheid van de mazetaak, zijn ook enkele onderzoeken uitgevoerd naar specifiek de relatie tussen begrijpend lezen en de mazetaak. In al deze onderzoeken is een vergelijking

LEESBEGRIIP IN DE MAZETAAK

gemaakt tussen de maze en *statetests* die gericht zijn op leesbegrip. Hieruit blijkt een matige relatie te bestaan, met concurrentvaliditeitscoëfficiënten tussen de .52 en .64 (Wiley & Deno, 2005; Tolar et al., 2012; Shin et al., 2019). Hoewel een van de studies een meta-analyse is, (Shin et al., 2019) is het aantal onderzoeken gering. Bovendien zit er beperkte variatie in de methoden van de onderzoeken.

De huidige studie vult de bestaande onderzoeken aan door het gebruik van een experimentele opzet, in plaats van validiteit te meten doormiddel van correlaties met andere testen. De participanten lazen een aantal teksten, waarvoor twee versies bestaan: een inhoudelijk consistente en een inconsistente tekstversie. Een inconsistentie kwam tot uiting in de targetzin (zin 9). Bij een aantal teksten (N=12) stond dit buiten het keuzemoment van de mazeselectie, bij laatste teksten (N=2) staat dit wel in de mazeselectie. Hierbij gaat het niet alleen om detectie van de inconsistentie, maar moeten lezers ook actief anticiperen door te klikken op de tegenstrijdige informatie. Daarnaast wordt ook zin 8 meegenomen in het experiment als controle. Deze wordt aangeduid als non-targetzin, omdat deze altijd kloppend is met het verhaal ongeacht de tekstversie. Om te onderzoeken of de participanten begrijpend lezen tijdens de taak, worden de leestijden in een consistent verhaal vergeleken met de leestijden in een inconsistent verhaal. Wanneer sprake is van tekstbegrip, zal de lezer meer leestijd nodig hebben bij het lezen van een targetzin in een inconsistent verhaal. Bij de non-targetzin zal hier geen verschil in moeten zijn.

Onderzoeksvraag en hypothese

In dit onderzoek staat de vraag centraal: ‘Wordt bij het maken van een CBM mazetaak leesbegrip weerspiegeld?’ Voor het beantwoorden van deze vraag wordt gekeken kijken naar het verschil in leestijden tussen de consistente en inconsistente teksten. Daarnaast wordt onderzocht of er onderscheid zit in de plek waar de inconsistentie geplaatst is: in de

LEESBEGRIP IN DE MAZETAAK

mazeselectie of ergens anders in de targetzin? Hiervoor zijn de volgende hypothesen opgesteld:

- H1) De leestijd van een targetzin is hoger bij een inconsistente tekst dan bij een consistente tekst, terwijl de leestijd van een non-targetzin niet verschilt tussen de tekstversies.
- H2) Het verschil in leestijd van een targetzin tussen een inconsistente tekst en consistente tekst is groter wanneer de inconsistentie in de mazeselectie gepresenteerd wordt dan wanneer het op een andere plek in de targetzin staat.

De eerste hypothese gaat met name om detectie van de inconsistentie, aan de hand van onderbewuste processen. Leesbegrip zal zich uiten in vertraging tijdens het lezen van een inconsistentie in de targetzin. Bij een non-targetzin wordt geen vertraging verwacht, omdat deze als controlemechanisme wordt meegenomen en de tekstversie geen invloed zou moeten hebben op de leestijd van deze zin. De tweede hypothese is opgebouwd vanuit het idee dat er een groter beroep wordt gedaan op leesbegrip als de inconsistentie in de mazeselectie aan het licht komt. Dit zal naar verwachting meer vertraging oproepen in leestijd dan wanneer de inconsistentie op een andere plek in de zin staat.

Methode

Steekproef

Participanten waren 58 leerlingen (38 jongens, 20 meisjes) uit groep 7 (n=24) en 8 (n=34) van vijf verschillende basisscholen in de Randstad. Voor de werving zijn deze scholen benaderd voor een samenwerking. Na akkoord voor de desbetreffende samenwerking is een brief naar de ouders gestuurd waarmee zij toestemming konden geven dat hun kind participeerde in het onderzoek. Om mee te doen aan het onderzoek, moesten de leerlingen over een gemiddelde tot goede leesvaardigheid beschikken. Het inclusie criterium hiervoor

LEESBEGRIP IN DE MAZETAAK

was een drieminutentoets (DMT) of CITO-classificatie van 1, 2 of 3, wat duidt op de hoogste leesniveaus.

Opzet

Design. Het onderzoek had een experimentele opzet, met een *two factor within-subject* design. Hoewel er geen sprake was van meerdere meetmomenten, zijn er wel meerdere vergelijkbare metingen per participant gedaan. De twee factoren die zijn meegenomen in deze studie zijn tekstversie (consistent/inconsistent) en type zin (targetzin/non-targetzin). Een beschrijving van de factoren komt aan de orde onder de kopjes onafhankelijke variabele en afhankelijke variabelen.

Het onderzoek bestond voor elke participant uit het lezen van veertien narratieve teksten, elk bestaande uit tien zinnen. De originele verhalen komen uit de studie van Helder en collega's (2016) en zijn geselecteerd omdat deze de beste resultaten lieten zien op het gebied van inconsistentiedetectering. De teksten zijn voor het huidige onderzoek bewerkt tot een mazetaak. In elke zin stonden één of twee mazeselecties. Op deze plekken is steeds één woord wegegelaten en daarvoor in de plaats stonden drie woorden waaruit de juiste optie aangeklikt moest worden. Slechts één keuzemogelijkheid was correct, de andere twee opties waren duidelijk incorrect. Alle antwoordmogelijkheden bestonden uit hetzelfde aantal letters en waren inhoudelijk niet aan elkaar gerelateerd.

Onafhankelijke variabele. Tekstversie was de onafhankelijke variabele van het onderzoek. Van elke tekst bestonden twee versies, een consistente en een inconsistente variant. De consistente versie was een causaal kloppend verhaal, met logische opbouw. De inconsistente variant was nagenoeg identiek, maar hier is één woord veranderd waardoor tegenstrijdige informatie werd gegeven. Dit verschil zat steeds in de tweede zin van de tekst. De lezer kon de inconsistentie echter pas in de negende zin opmerken, aangezien daar terug werd gerefereerd naar de eerdere informatie. Ter illustratie: in de consistente tekst stond in zin

LEESBEGRIP IN DE MAZETAAK

twee dat Lara een *broertje* krijgt en vervolgens speelt ze in zin negen, zoals verwacht, ook met haar *broertje*. In een inconsistente tekst stond in de tweede zin dat Lara een *zusje* krijgt, maar in zin negen speelt ze met haar nieuwe *broertje*. De informatie is niet direct achter elkaar geplaatst, zodat er meer beroep wordt gedaan op de cognitieve processen binnen leesvaardigheid (bijv. geheugen).

De tegenstrijdige informatie stond bij de meeste inconsistente teksten niet in de mazeselectie. Bij twaalf van de veertien teksten (hierna: reguliere teksten) stond inconsistentie niet in de mazeselectie van targetzin. Het beïnvloedde dus niet direct het keuzemoment in die zin. In de laatste twee teksten die de participanten maakten (hierna: Loes & Lara teksten), was de inconsistentie wel verwerkt in de mazeselectie. Het juiste antwoord voor de taak maakte dan wel een grammaticaal goedlopende zin, maar kwam niet overeen met de eerdere inhoud van de tekst. De participanten moesten dus een discrepantie aanklikken (bijv. het woord *broertje*) om het goede antwoord te geven. Deze teksten zijn zo opgebouwd, omdat hiermee leesbegrip beter tot uiting komt. Er wordt namelijk meer gevraagd van begripsvaardigheden van de lezer, bijvoorbeeld coherentieherstel.

Afhankelijke variabelen. De afhankelijke variabelen waren de leestijden per zin, met daarin onderscheid in tekstversie. Hieruit volgden de variabelen: leestijd targetzin consistente tekst, leestijd targetzin inconsistente tekst, leestijd targetzin consistente tekst en leestijd non-targetzin inconsistente tekst. De targetzin (zin 9) stond centraal, omdat hier een mogelijke inconsistentie geconstateerd kon worden. Daarnaast is de leestijd gemeten van de zin ervoor (zin 8, non-targetzin). Deze diende als controlevariabele, omdat er geen onderscheid zou moeten zijn bij het lezen van deze zin in de verschillende tekstversies. De leestijden werden gemeten met behulp van een *eye tracker* die de oogbewegingen van de participant gedurende de gehele teksten vastlegde. De participanten wisten dat hun oogbewegingen gevolgd werden. Ze kregen de instructie de mazetaak te maken en daarbij werd ze gevraagd naar het scherm te

LEESBEGRIP IN DE MAZETAAK

blijven kijken en niet te veel te bewegen op hun stoel. Met behulp van software is opgenomen hoe lang de participanten erover deden om de gehele zinnen te lezen. Deze data is gecodeerd als leestijd in seconden. Hieruit volgde voor elke participant de leestijd van veertien targetzinnen en veertien non-targetzinnen.

Procedure

Voorafgaand aan de testafname zijn de participanten willekeurig ingedeeld in een van de vier condities. Zo is de steekproef verdeeld in gelijke groepen die elk bestonden uit veertien of vijftien participanten. De condities verschilden van elkaar in volgorde van de teksten en in welke versie van een tekst gelezen werd. In alle condities werden alle veertien teksten gelezen die geselecteerd waren voor dit onderzoek, maar steeds een van de twee versies (consistent of inconsistent). Dit systeem is op twaalf teksten toegepast, waarbij uiteindelijk in elke conditie de helft van de teksten inconsistent was.

Twee bijzondere teksten in de studie waren de Loes en Lara teksten. Deze hadden een andere opbouw van de mazetaak, waarbij de inconsistentie in de mazeselectie gepresenteerd werd. In de overige teksten stond het op een andere plek van de targetzin. Loes en Lara teksten werden in elke conditie als laatste gelezen. Verder is het systeem van deze set teksten vergelijkbaar met de overige teksten: Elke participant las altijd één van teksten in consistente versie en één in inconsistente versie. De welke van de twee eerst gelezen werd verschilde per conditie.²

De testafname vond plaats op de school van de desbetreffende participant. De testen zijn afgenomen door getrainde masterstudenten van de Universiteit Leiden waarvan er twee bij elke testafname aanwezig waren. De testafname vond individueel plaats. Eén van de onderzoekers haalde het kind uit de les op en begeleidde deze naar de ruimte van de testafname. Tijdens de korte wandeling werd alvast kennis gemaakt om de spanning voor het

² Voor een volledig overzicht van tekstversies en afnamevolgorde per conditie, zie bijlage 1.

LEESBEGRIP IN DE MAZETAAK

kind te verlagen. In de testruimte volgde uitleg over de gehele testafname en per onderdeel werd extra specifieke informatie gegeven. Omdat deze studie deel uitmaakt van een groter onderzoek, moesten de participanten bij de start van de afname twee reguliere CBM mazetaken op papier lezen. Vervolgens werd de *eye tracker* op de computer ingesteld en gekalibreerd, zodat de oogbewegingen van de participant geregistreerd konden worden tijdens het lezen. Voordat de officiële dataverzameling startte, maakten de scholieren een oefentaak op de computer waarvan de antwoorden door de participant zelf bekeken werden ter controle. Daarna werden achtereenvolgend de mazetaken gemaakt waarbij de participant met de muis het juiste antwoord op de laptop selecteerde. Halverwege de reeks is een korte pauze gehouden. Na elke tekst werden enkele vragen gesteld om aanvullende data te verkrijgen. De eerste vraag was na elke tekst exact gelijk: ‘Klopte het verhaal?’. Indien dit negatief werd beantwoord, is doorgevraagd om te controleren of de inconsistentie daadwerkelijk gedetecteerd was en er niet gegokt of verkeerd geïnterpreteerd was. Vervolgens werd een specifieke ja/nee-vraag gesteld over de inhoud van de tekst, bijvoorbeeld ‘Trakteert Lara op beschuit met muisjes?’.

Na het verzamelen van de data is alles door de studenten uit de projectgroep gescoord en gecodeerd. Vervolgens is dit ingevoerd in analyseprogramma SPSS. De data van een participant is altijd door twee studenten afzonderlijk gescoord en gecodeerd ter bevordering van de interbeoordelaarsbetrouwbaarheid.

Analyses

Voor de analyse is gebruikgemaakt van de *repeated measures* ANOVA. Voordat dit mogelijk was, moesten de data bewerkt worden. De tekstversie (consistent/inconsistent) was een *within* variabele, wat voor deze studie betekende dat elke participant de helft van de consistente en de helft van de inconsistente teksten las. De analyses werden om deze reden uitgevoerd met gemiddelde scores voor afzonderlijk de consistente en inconsistente teksten van elke participant. Door het middelen van losse scores zijn nieuwe variabelen gecodeerd,

LEESBEGRIP IN DE MAZETAAK

waarbij onderscheid is gemaakt tussen tekstversies en type zinnen. Daaruit volgden de afhankelijke variabelen van dit onderzoek.

Na een *repeated measures* ANOVA om alle variabelen onderling te vergelijken, is een afhankelijke t-test uitgevoerd. Hiermee is meer gedetailleerd gekeken naar de interactie-effecten.

Resultaten

In dit onderzoek stond de vraag centraal of leesbegrip weerspiegeld wordt in de CBM mazetaak. Dit is gedaan door de leestijd van de targetzin te vergelijken tussen een consistente tekst en een inconsistente tekst. Hierbij is onderscheid gemaakt tussen reguliere teksten de Loes & Lara teksten.

Reguliere teksten

Het gaat hier om weerspiegeling van leesbegrip wanneer de inconsistentie niet in de mazeselectie gepresenteerd wordt. De verwachting was dat de leestijd van de targetzin significant hoger is bij een inconsistente tekst dan bij een consistente tekst. Bij de non-targetzin werd geen significant verschil verwacht tussen de tekstversies.

Data-inspectie. Voordat de inhoudelijke analyse plaatsvond, is eerst data-inspectie uitgevoerd. Onderzoek naar de 58 respondenten leverde de volgende gegevens op (Tabel 1).

LEESBEGRIP IN DE MAZETAAK

Tabel 1*Data-inspectie reguliere teksten (inclusief uitbijter)*

| | N | M (in sec) | SD | Range (in sec) | Scheefheid | Gepiektheid |
|----------------------|----|---------------|-------|-------------------|------------|-------------|
| Targetzin | | | | | | |
| Consistente tekst | 58 | 6.622 | 2.429 | 0.76-13.00 | -.150 | .076 |
| Inconsistente tekst | 58 | 6.934 | 2.502 | 2.19-15.74 | 2.312 | 2.932 |
| Non-targetzin | | | | | | |
| Consistente tekst | 58 | 6.968 | 2.543 | 1.72-16.08 | 2.844 | 3.312 |
| Inconsistente tekst | 58 | 6.894 | 2.272 | 2.26-12.42 | .787 | -.338 |

Er waren geen missende data binnen de variabelen. Na analyse naar uitbijters bleken twee respondenten extreme waarden te bevatten. Deze wijken beiden meer dan drie standaarddeviaties af van het gemiddelde, wat wijst op een extreme uitbijter (Moore & McCabe, 2006). Een van deze respondenten wijkt op meerdere teksten af van de bijbehorende gemiddelde leestijd. Uit observaties en afnamegegevens bleek dat de participant bepaalde verhalen nogmaals las na het invullen van het laatste keuzemoment. Daarnaast kon de *eye tracker* niet goed afgesteld worden, dus mogelijk zijn de resultaten hierdoor onvoldoende secuur. De andere respondent blijkt op één specifieke tekst een extreme waarde te hebben. Deze is gehercodeerd als missende waarde, om beïnvloeding van één afwijkende leestijd te voorkomen. Verder zijn alle variabelen bij benadering normaal verdeeld. Ter controle is de data-inspectie zowel met als zonder uitbijter uitgevoerd. De volgende gegevens volgen wanneer de participant met meerdere extreme waardes geëxcludeerd worden uit het onderzoek (Tabel 2).

LEESBEGRIP IN DE MAZETAAK

Tabel 2*Data-inspectie reguliere teksten (exclusief uitbijters)*

| | N | M (in sec) | SD | Range (in sec) | Scheefheid | Gepiektheid |
|----------------------|----|---------------|-------|-------------------|------------|-------------|
| Targetzin | | | | | | |
| Consistente tekst | 57 | 6.587 | 2.436 | 0.76-13.00 | -.044 | .095 |
| Inconsistente tekst | 57 | 6.886 | 2.278 | 2.19-12.66 | .285 | .069 |
| Non-targetzin | | | | | | |
| Consistente tekst | 57 | 6.809 | 2.253 | 1.72-12.73 | .965 | .332 |
| Inconsistente tekst | 57 | 6.903 | 2.291 | 2.26-12.42 | .737 | -.413 |

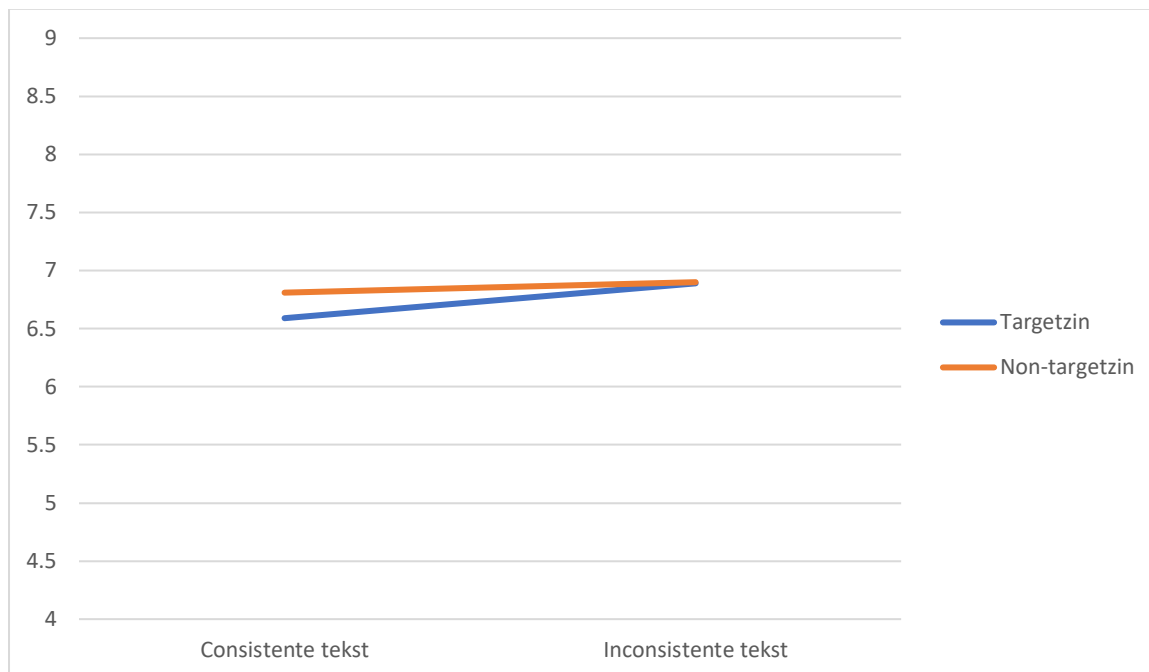
Na exclusie van de uitbijters liggen zowel Z-score van de scheefheid en gepiektheid bij alle variabelen tussen de -3 en 3 (Moore & McCabe, 2006). Omdat de data zo beter binnen de normaalverdeling passen, is gekozen om de analyses uit te voeren zonder deze uitbijters.

Relatie leestijd en tekstversie. Wordt leesbegrip weerspiegeld in de CBM mazetaak wanneer de inconsistentie niet in de mazeselectie wordt gepresenteerd? Een *repeated measures* ANOVA (Tabel 1) laat zien dat er bij de reguliere teksten geen significant verband is tussen leestijd en tekstversie ($F(56,1) = .969$, $p = .185$) of zin ($F(56,1) = .988$, $p = .420$). Ook is er geen significant interactie-effect ($F(56,1) = .991$, $p = .485$) tussen tekst versie en zin.

LEESBEGRIP IN DE MAZETAAK

Figuur 1

Leestijd (in sec) voor de target- en non-targetzin in consistente en inconsistente teksten, reguliere teksten

**Loes & Lara teksten**

Wordt leesbegrip weerspiegeld in de CBM mazetaak wanneer de inconsistentie in een mazeselectie wordt gepresenteerd? Dit is onderzocht aan de hand van de Loes & Lara teksten. De verwachting was dat bij de targetzin het verband tussen leestijd en tekstversie significant verschilt en dit verband sterker is dan bij de reguliere teksten. Bij de non-targetzin werd geen significant verband verwacht.

Data-inspectie. Om te beginnen is ook voor deze variabelen gekeken naar de opbouw van de data (Tabel 3).

LEESBEGRIP IN DE MAZETAAK

Tabel 3*Data inspectie Loes & Lara teksten (inclusief uitbijters)*

| | N | M (sec) | SD | Range (sec) | Scheefheid | Gepiektheid |
|----------------------|----|---------|-------|-------------|------------|-------------|
| Targetzin | | | | | | |
| Consistente tekst | 58 | 6.310 | 3.250 | 0.00-14.05 | .688 | -.269 |
| Inconsistente tekst | 58 | 9.476 | 5.931 | 0.02-29.26 | 3.943 | 2.958 |
| Non-targetzin | | | | | | |
| Consistente tekst | 58 | 7.940 | 3.591 | 0.00-15.65 | .369 | -.006 |
| Inconsistente tekst | 58 | 7.890 | 3.550 | 0.00-18.78 | 2.032 | 2.555 |

Ook uit deze data-inspectie komen geen missende waarden naar voren, door het samenvoegen van gegevens van beide teksten. Wel zitten er extreme waarden in de data van vier verschillende respondenten. Eén van deze respondenten is ook bij de vorige analyse volledig geëxcludeerd, dus deze zal wederom uitgesloten worden. Daarnaast weken drie respondenten af met meer dan drie standaarddeviaties op één specifieke variabele. Deze extreme waarden zijn gehercodeerd naar missende waarden. Na aanpassing van de data is opnieuw een data-inspectie uitgevoerd waar de volgende gegevens uit voortvloeien (Tabel 4).

Tabel 4*Data inspectie tekst Loes & Lara (exclusief uitbijters)*

| | N | M (sec) | SD | Range (sec) | Scheefheid | Gepiektheid |
|----------------------|----|---------|-------|-------------|------------|-------------|
| Targetzin | | | | | | |
| Consistente tekst | 57 | 6.363 | 3.253 | 0.00-14.05 | .579 | -.242 |
| Inconsistente tekst | 55 | 8.862 | 4.934 | 0.02-21.68 | 2.205 | .364 |
| Non-targetzin | | | | | | |
| Consistente tekst | 57 | 7.909 | 3.615 | 0.00-15.65 | .443 | -.043 |
| Inconsistente tekst | 56 | 7.532 | 3.043 | 0.00-16.06 | .141 | 1.486 |

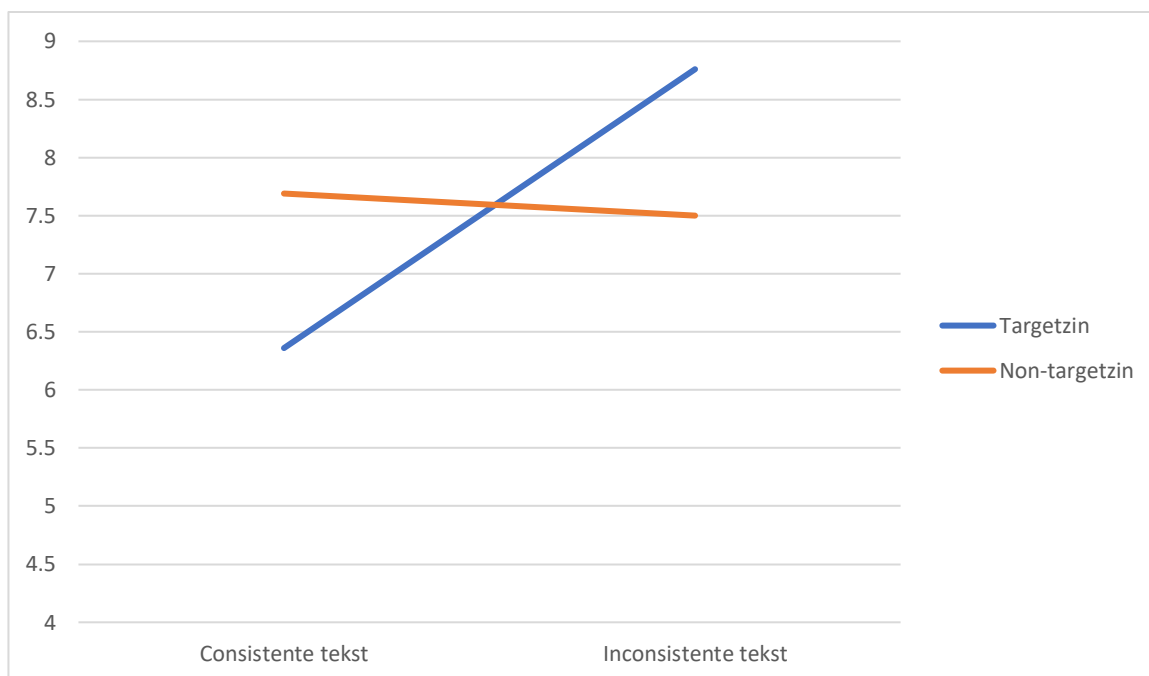
LEESBEGRIP IN DE MAZETAAK

Met het uitsluiten van de uitbijters zijn de data beter verdeeld en is de kans op beïnvloeding van meetfouten verkleind. Gekozen is daarom om de analyses verder uit te voeren exclusief de uitbijters.

Relatie leestijd en tekstversie. Om het verband tussen leestijd en tekstversie voor de target- en non-targetzin afzonderlijk te onderzoeken is wederom een *repeated measures* ANOVA uitgevoerd (Figuur 2). Hieruit blijkt dat er voor de Loes en Lara teksten een significant verband is tussen leestijd en tekstversie ($F(53,1) = .876$, $p = .008$). Tussen leestijd en zin bestaat geen significant verband ($F(53,1) = 1.000$, $p = .903$). Wel bestaat er een significant interactie-effect wanneer zowel tekstversie als zin worden meegenomen als variabelen ($F(53,1) = .857$, $p = .004$).

Figuur 2

Leestijd (in sec) voor de target- en non-targetzin in consistente en inconsistente teksten, Loes en Lara teksten.



Om preciezer te onderzoeken hoe het interactie-effect is opgebouwd, is aanvullend een afhankelijke t-test gedaan. Hieruit blijkt dat de leestijd van de targetzin significant hoger is bij

LEESBEGRIIP IN DE MAZETAAK

een inconsistente tekst dan bij een consistente tekst ($t(54) = -3.713$, $p = .000$). Bij de non-targetzin wordt geen significant verschil gevonden in leestijd ($t(55) = .521$, $p = .604$).

Conclusie

In deze studie is onderzocht of leesbegrip wordt weerspiegeld bij het maken van de mazetaak. Dit is geoperationaliseerd aan de hand van het inconsistentieparadigma. Hierbij zijn de leestijden van de targetzin in consistente en inconsistente teksten met elkaar vergeleken. Allereerst de resultaten van de reguliere teksten, waarbij de inconsistentie buiten de mazeselectie staat. Deze laten zien dat er geen significant verschil is in leestijd van de targetzin tussen de twee tekstversies. De participanten vertragen nauwelijks bij een inconsistentie en er treedt hier geen inconsistentie-effect op. Dit suggereert dat tegenstrijdige informatie niet gedetecteerd wordt of in ieder geval dat er geen duidelijke reactie volgt bij het lezen ervan. Hieruit volgt geen bewijs van leesbegrip tijdens het lezen van de reguliere teksten (Albrecht & O'Brien, 1993).

In de teksten waarin de inconsistentie in de mazeselectie staat (Loes en Lara teksten), zijn wel significante verschillen in de leestijd bij consistente en inconsistente teksten. Voor het lezen van de targetzin in een inconsistente tekst hebben de participanten gemiddeld meer tijd nodig dan in een consistente tekst. De aanwezigheid van een inconsistentie-effect impliceert dat de lezer tegenstrijdige informatie heeft opgemerkt. Dit wijst op leesbegrip, omdat de lezer informatie uit het begin van de tekst opgeslagen heeft en er terugkoppeling plaatsvindt wanneer hiernaar verwezen wordt (Albrecht & O'Brien, 1993).

Bij de non-targetzinnen in zowel de reguliere als de Loes en Lara teksten komen geen significante verschillen naar voren. De participanten doen dus bij alle teksten gemiddeld net zo lang over het lezen van deze 'gewone' zin, ongeacht tekstversie. Het zegt niet direct iets over de mate van leesbegrip, maar wel iets over validiteit van de resultaten in een targetzin.

LEESBEGRIIP IN DE MAZETAAK

Het ondersteunt dat de leestijdverschillen in de targetzin voortkomen uit de manipulatie en niet op toeval of andere factoren berusten.

De hypothese dat de leestijd van een targetzin langer is bij een inconsistente tekst dan bij een consistente tekst wordt uitsluitend bij de Loes en Lara teksten bevestigd en niet bij de reguliere teksten. Hiermee kan de hypothese niet worden aangenomen. De tweede hypothese kan wel worden aangenomen. Het verschil in leestijd tussen een consistente en inconsistente targetzin is groter wanneer de inconsistentie in de mazeselectie gepresenteerd wordt, dan wanneer de inconsistentie op een andere plek in de targetzin staat. Ook de hypothese dat er geen verschil zit in leestijd tussen de non-targetzin de tekstversies wordt bevestigd.

Een mogelijke verklaring voor het verschil in resultaten bij de reguliere en Loes en Lara teksten zou kunnen liggen in de aandachtsturing tijdens het lezen. Het kan zijn dat de lezer de aandacht vooral richt op de mazeselectie, omdat daar meer relevante informatie staat voor het maken van de taak. De stukken ertussen worden dan vluchtiger of oppervlakkiger gelezen (Anderson, 1982). In dat geval zou de lezer de coherentie minder goed monitoren, laat staan coherentie herstellen (Helder et al., 2013). Bij het invullen van de mazeselectie wordt een handeling gevraagd, wat voor een onderbreking in het leesproces zorgt. Mogelijk richt de lezer de aandacht meer op deze delen tekst en bevordert dit inconsistentiedetectie. Bovendien worden lezers beoordeeld op het aangeklikte antwoord in de selectie, wat zou kunnen motiveren om meer tijd te nemen voor coherentieherstel (Ryan & Deci, 2000).

Kan nu worden geconcludeerd dat leesbegrip wordt weerspiegeld bij het maken van de mazetaak? Op basis van het huidige onderzoek kan geen eenduidig antwoord worden gegeven. Op sommige momenten is er sprake van leesbegrip, in ieder geval in het keuzemoment van de mazetaak. In de overige stukken is geen bewijs gevonden van leesbegrip. Op basis daarvan kan met voorzichtigheid worden geconcludeerd dat tijdens het maken van de mazetaak wel enig begripsvaardigheden worden ingezet, maar er (nog)

LEESBEGRIIP IN DE MAZETAAK

onvoldoende onderbouwing is dat de taak een sterke weerspiegeling biedt van het niveau van leesbegrip. Dit lijkt samen te hangen met hoe een tekst is opgebouwd en hoe de mazeselecties worden ingevuld.

Eerder onderzoek op het gebied van lees- of taalbegrip in de mazetaak is voornamelijk gericht op concurrentvaliditeit, waardoor concrete resultaten niet exact vergelijkbaar zijn. Wel komen de conclusies in grote lijnen overeen: de taak hangt samen met leesbegrip, maar de sterkte van de samenhang is matig (Wiley & Deno, 2005; Tolar et al., 2012; Shin et al., 2019). Ondanks het verschil in onderzoeksopzet, is de conclusie van het huidige onderzoek wel vergelijkbaar. Het sluit bovendien aan bij de twijfel van Muijselaar en collega's (2016) over de inzet van begripsvaardigheden bij het maken van de mazetaak.

Wel moet in acht worden genomen dat het huidige onderzoek een aantal beperkingen heeft. Allereerst zijn slechts twee teksten gebruikt waarbij de inconsistentie in de mazeselectie tot uiting komt (Loes en Lara). Daar tegenover staan twaalf teksten waarbij de inconsistentie buiten de mazeselectie gepresenteerd wordt. Bovendien zijn de Loes en Lara teksten altijd op het eind van de test afgenomen. Enerzijds zou dit presentaties negatief kunnen beïnvloeden door bijvoorbeeld vermoeidheid of verminderde concentratie. Anderzijds zou een leereffect een verbetering in prestaties kunnen opleveren. Bekend is dat mazetaken beter worden gemaakt naarmate leerlingen dit herhaaldelijk doen (e.g. Tzivinikou et al., 2019). Bij reguliere teksten zijn de mogelijke effecten van concentratie en herhaling opgevangen door de teksten in willekeurige volgorde af te nemen. Dit is bij de Loes en Lara teksten niet het geval.

Een andere beperking is dat niet alle data zijn meegenomen in de analyses voor het beantwoorden van de hoofdvraag. Aanvullende data kunnen meer inzicht bieden in de leeswijze van de participanten. De antwoorden op de begrips- en consistentievragen na iedere tekst kunnen een beter beeld geven van de bewuste detectie en opslag van de inconsistenties.

LEESBEGRIP IN DE MAZETAAK

Daarnaast kunnen aanvullende data vanuit de meer inzicht geven in de onbewuste processen, zoals aandachtverdeling binnen de individuele zinnen of teruglezen van eerdere zinnen.

Verder kan nog een kritische noot geplaatst worden bij de steekproef. De participanten komen, uit praktische overwegingen, allen uit de randstand. Het betreft kinderen op basisscholen die bereikbaar zijn met het openbaar vervoer. Daarbij bestaat steekproef alleen uit scholieren met een (boven)gemiddeld leesniveau. De resultaten zijn daardoor beperkt generaliseerbaar voor basisschoolscholieren. Bij een benedengemiddeld leesniveau moet rekening worden gehouden met mogelijk een andere relatie tussen de mazetaak en leesbegrip.

Hoewel deze studie beperkingen heeft en veel extra vragen oproept, heeft het een sterke basis. Door het gebruik van het inconsistentieparadigma wordt leesbegrip op procesniveau gemeten. Bij leesbegripstesten moet altijd rekening worden gehouden dat ook andere (lees)vaardigheden ingezet worden bij het maken ervan. In de vergelijking zullen die overige aspecten, bijvoorbeeld coderen, ook een rol spelen. Het inconsistentieparadigma is specifiek gericht op begripsvaardigheden (Albrecht & O'Brien, 1993).

Naast het gericht meten van leesbegrip, is het gebruik van *eye tracking* een sterk aspect van de onderzoeksopzet. Het toepassen van *eye tracking* om inconsistentie-effecten op te sporen is niet eerder ingezet in combinatie met de mazetaak. Door deze technologie wordt een kijk gegeven in het leesproces en aandachtsturing van de participanten. Hiermee kunnen onderbewuste processen beter in kaart worden gebracht, zonder hierin te interveniëren (Veneri, Federighi, Rosini, Federico & Rufa, 2010).

Een ander voordeel het gebruik van korte teksten, waardoor in relatief korte tijd veel data verzamelt kan worden. Één participant kan meerdere teksten lezen tijdens een afnamemoment, wat het experimenteren met verschillende tekstversies en manieren van tekstopbouw makkelijker maakt. Naast de mogelijkheid om veel diversiteit in

LEESBEGRIP IN DE MAZETAAK

wetenschappelijke inzichten te verkrijgen, bevordert het ook de betrouwbaarheid van de resultaten.

Deze sterke basis biedt een goed uitgangspunten om op voort te bouwen in vervolgonderzoek. Uiteraard is het belangrijk om daarbij rekening te houden met de besproken beperkingen. Een advies is dat er meer aandacht is voor de opbouw van de mazetaak en wat de invloed daarvan is op de mate waarin leesbegrip tot uiting komt. Meer variatie in teksten is gewenst en er kan bijvoorbeeld aandacht komen voor de plek van de mazeselectie en keuze voor afleiders. Het vermoeden is dat wanneer de afleiders inhoudelijk meer gerelateerd zijn aan de inhoud van de tekst, leesbegrip beter tot uitdrukking komt (Conoyer et al., 2017).

De combinatie van offline data en meer gedetailleerde online data kan een goede aanvulling zijn. Door al deze informatie mee te nemen in de analyse worden bewuste en onbewuste processen gecombineerd, wat zal helpen om beter te begrijpen welke vaardigheden tot uiting komen in de mazetaak.

Tot slot zou in het licht van generaliseerbaarheid een bredere populatie meegenomen kunnen worden. Het vermogen tot coherentie monitoring ontwikkelt zich gedurende de basisschoolperiode (Markman, 1979). Jongere kinderen blijken minder goed in staat inconsistenties te herkennen, evenals leerlingen met zwakkere begripsvaardigheden (Helder et al., 2016). Voor inzet van de mazetaak in het onderwijs is het relevant om te weten welke vaardigheden ingezet worden per leeftijd of leesniveaus.

Samengevat is vervolgonderzoek geïndiceerd om nader te onderzoeken welke vaardigheden ingezet worden bij het maken van de mazetaak. De boodschap voor de praktijk is dan ook om niet uitsluitend de mazetaak in te zetten voor het monitoren van het niveau van begrijpend lezen. Voor nu is er te weinig zekerheid of, en in welke hoedanigheid, het een

LEESBEGRIP IN DE MAZETAAK

geschikt instrument is voor dit doeleinde. Wetenschappelijk onderzoek kan hier een goede bijdrage aan leveren en zo voor meer begrip zorgen in de toekomst.

Literatuur

- Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of experimental psychology: Learning, memory, and cognition*, *19*(5), 1061-1070. <https://doi.org/10.1037//0278-7393.19.5.1061>
- Alley, G. R., & Deshler, D. D. (1979). *Teaching the learning disabled adolescent: Strategies and methods*. Denver, CO: Love
- Allington, R. L. (2007). Intervention all day long: New hope for struggling readers. *Voices from the Middle*, *14*(4), 7-14.
- Anderson, R. C. (1982). Allocation of attention during reading. In A. Flammer, W. Kintsch (Red.), *Advances in Psychology* (Vol. 8, pp. 292-305). Amsterdam: North-Holland Publishing Company.
- Best, R. M., Floyd, R. G., & McNamara, D. S. (2008). Differential competencies contributing to children's comprehension of narrative and expository texts. *Reading Psychology*, *29*, 137–164. <http://dx.doi.org/10.1080/02702710801963951>
- Cito, (2021, 4 mei). (z.d.) Toetskalenders 2020-2021. Geraadpleegd via <https://www.cito.nl/onderwijs/primair-onderwijs/toetskalenders>.
- Chall, J. (1996). *Stages of reading development (2nd ed.)*. Fort Worth, Tex.: Harcourt Brace.
- Chung, S., Espin, C. A., & Stevenson, C. E. (2018). CBM maze-scores as indicators of reading level and growth for seventh-grade students. *Reading and writing*, *31*(3), 627-648. <https://doi.org/10.1007/s11145-017-9803-8>
- Conoyer, S.J., Lembke, E.S., Hosp, J.L., Espin, C.A., Hosp, M.K., & Poch, A.L., (2017). Getting More From Your Maze: Examining Differences in Distractors, *Reading & Writing Quarterly*, *33*(2), 141-154. <https://doi.org/10.1080/10573569.2016.1142913>
- Deno, S. L. (2003). Developments in Curriculum- Based Measurement. *The Journal of Special Education*, *37*(3), 184-192. <https://doi.org/10.1177/00224669030370030801>

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional children*, 52(3), 219-232. <https://doi.org/10.1177/001440298505200303>

Deno, S. L., Mirkin, P. K., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children*, 49(1), 36-45. <https://doi.org/10.1177/001440298204900105>

Duke, N. K., & Pearson, P. D. (2002). Effective practices for developing reading comprehension. In A. E. Farstrup, & S. J. Samuels (Eds.), *What research has to say about reading instruction* (pp. 205-242). Newark, DE: International Reading Association.

Espin, C. A., & Foegen, A. (1996). Validity of general outcome measures for predicting secondary students' performance on content-area tasks. *Exceptional Children*, 62(6), 497-514. <https://doi.org/10.1177/001440299606200602>

Espin, C. A., Wallace, T., Lembke, E., Campbell, H., & Long, J. (2010). Creating a progress-monitoring system in reading for middle-school students: Tracking progress toward meeting high-stakes standards. *Learning Disabilities Research and Practice*, 25(2), 60-75. <https://doi.org/10.1111/j.1540-5826.2010.00304.x>.

FastBridge Learning. (2015). *Formative assessment system for teachers (FAST): Technical manual*. Minneapolis, MN: Author.

Fuchs, L. S., & Fuchs, D. (1992). Identifying a measure for monitoring student reading progress. *School Psychology Review*, 21(1), 45-58. <https://doi.org/10.1080/02796015.1992.12085594>

Gerrig, R. J., & O'Brien, E. J. (2005). The scope of memory-based processing. *Discourse Processes*, 39(2), 225-242. <https://doi.org/10.1080/0163853x.2005.9651681>.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7(1), 6-10. <https://doi.org/10.1177/074193258600700104>

- Gubbels, J., Van Langen, A. M. L., Maassen, N. A. M., Meelissen, M. R. M. (2019). Resultaten PISA-2018 in vogelvlucht. Enschede: Universiteit Twente. <https://doi.org/10.3990/1.9789036549226>
- Hannon, B. (2012). Understanding the relative contributions of lower-level word processes, higher-level processes, and working memory to reading comprehension performance in proficient adult readers. *Reading Research Quarterly*, 47(2), 125–152. <https://doi.org/10.1002/RRQ.013>.
- Helder, A., Broek, P., van den, Leijenhorst, L. van, & Beker, K. (2013). Sources of comprehension problems during reading. In B. Miller, L. Cutting & P. McCardle (Eds.), *Unraveling reading comprehension: Behavioral, neurobiological, and genetic components* (pp. 43-53). Baltimore, md: Paul H. Brookes Publishing.
- Helder, A., Van Leijenhorst, L., & van den Broek, P. (2016). Coherence monitoring by good and poor comprehenders in elementary school: Comparing offline and online measures. *Learning and Individual Differences*, 48, 17-23. <https://doi.org/10.1016/j.lindif.2016.02.008>
- Hosp, M.K., Hosp, J.L., & Howell, K.W. (2007). *The ABC's of CBM. A Practical Guide to Curriculum- Based Measurement*. New York: The Guilford Press.
- Jenkins, J. R., & Jewell, M. (1993). Examining the validity of two measures for formative teaching: Reading aloud and maze. *Exceptional Children*, 59(5), 421–432. <https://doi.org/10.1177/001440299305900505>
- Kendeou, P., van den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning Disabilities Research & Practice*, 29(1), 10-16. <https://doi.org/10.1111/ldrp.12025>
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

LEESBEGRIP IN DE MAZETAAK

- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(1), 363–394. <https://doi.org/10.1037/0033-295X.85.5.363>
- Kordes, J., Bolsinova, M., Limpens, G., & Stolwijk, R. (2013). Resultaten PISA-2012 in vogelvlucht. *Praktische kennis en vaardigheden van 15-jarigen. Nederlandse uitkomsten van het Programme for International Student Assessment (PISA) op het gebied van wiskunde, natuurwetenschappen en leesvaardigheid in het jaar 2012*. Verkregen via www.pisa-nederland.nl/resultaten/.
- Long, D. L., & Chong, J. L. (2001). Comprehension skill and global coherence: A paradoxical picture of poor comprehenders' abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(6), 1424–1429. <https://doi.org/10.1037/0278-7393.27.6.1424>.
- Markman, E.M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. *Child Development*, 50(3), 643-655. <https://doi.org/10.2307/1128929>
- McNamara, D.S., & Magliano, J. P. (2009). Towards a comprehensive model of comprehension. In B. Rose (Ed.), *The psychology of learning and motivation* (pp.297-384). New York, NY: Academic Press.
- Moin, L. J., Magiera, K., & Zigmond, N. (2009). Instructional activities and group work in the US inclusive high school co-taught science class. *International Journal of Science and Mathematics Education*, 7, 677–697. <https://doi.org/10.1007/s10763-008-9133-z>
- Moore, D.S., & McCabe, G.P. (2006). *Introduction to the practice of statistics. (5th Edition)*. New York: W.H. Freeman and Company.
- Muijselaar, M. M., Kendeou, P., de Jong, P. F., & van den Broek, P. W. (2017). What does the CBM-Maze test measure? *Scientific Studies of Reading*, 21(2), 120-132. <https://doi.org/10.1080/10888438.2016.1263994>

LEESBEGRIIP IN DE MAZETAAK

- Oakhill, J., Hartt, J., & Samols, D. (2005). Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and Writing, 18*, 657–686. <http://doi.org/10.1007/s11145-005-3355-z>.
- O'Brien, E. J., Cook, A. E., & Gueraud, S. (2010). Accessibility of outdated information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(4), 979–991. <http://doi.org/10.1037/a0019763>.
- Perfetti, C., Yang, C. L., & Schmalhofer, F. (2008). Comprehension skill and word-to-text integration processes. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 22*(3), 303-318. <https://doi.org/10.1002/acp.1419>
- Rapp, D. N., Broek, van den, P. V. D., McMaster, K. L., Kendeou, P., & Espin, C. A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific studies of reading, 11*(4), 289-312. <https://doi.org/10.1080/10888430701530417>
- Ryan, R.M., & Deci, E.L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78. <https://doi.org/10.1037/110003-066X.55.1.68>
- Scruggs, T. E., & Mastropieri, M. A. (2007). Science learning in special education: The case for constructed versus instructed learning. *Exceptionality, 15*(2), 57–74. <https://doi.org/10.1080/09362830701294144>
- Shin, J., & McMaster, K. (2019). Relations between CBM (oral reading and maze) and reading comprehension on state achievement tests: A meta-analysis. *Journal of school psychology, 73*(2), 131-149. <https://doi.org/10.1016/j.jsp.2019.03.005>
- Shin, J., Deno, S. L., & Espin, C. (2000). Technical adequacy of the maze task for curriculum based measurement of reading growth. *The Journal of Special Education, 34*(3), 164–172. <https://doi.org/10.1177/002246690003400305>

LEESBEGRIP IN DE MAZETAAK

- Tichá, R., Espin, C., & Wayman, M. (2009). Reading Progress Monitoring for Secondary-School Students: Reliability, Validity, and Sensitivity to Growth of Reading-Aloud and Maze-Selection Measures. *Learning Disabilities Research & Practice, 24*(3), 132-142. <https://doi.org/10.1111/j.1540-5826.2009.00287.x>
- Tolar, T. D., Barth, A. E., Francis, D. J., Fletcher, J. M., Stuebing, K. K., & Vaughn, S. (2012). Psychometric properties of maze tasks in middle school students. *Assessment for Effective Intervention, 37*(3), 131–146. <https://doi.org/10.1177/1534508411413913>
- Tzivinikou, S. , Tsolis, A. , Kagkara, D. and Theodosiou, S. (2020). Curriculum Based Measurement Maze: A Review. *Psychology, 11*(10), 1592-1611. <https://doi.org/10.4236/psych.2020.1110101>
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press. <http://dx.doi.org/10.1177/1461445606059565>
- van Engelshoven, I., & Slob, A. (2019, 3 december). Kamerbrief over leesoffensief. Geraadpleegd via: <https://www.rijksoverheid.nl/documenten/kamerstukken/2019/12/03/kamerbrief-over-leesoffensief>.
- van der Schoot, M., Reijntjes, A., & van Lieshout, E. C. (2012). How do children deal with inconsistencies in text? An eye fixation and self-paced reading study in good and poor reading comprehenders. *Reading and Writing, 25*(7), 1665–1690. <http://dx.doi.org/10.1007/s11145-011-9337-4>.
- Veneri, G., Federighi, P., Rosini, F., Federico, A., & Rufa, A. (2010). Influences of data filtering on human–computer interaction by gaze-contingent display and eye-tracking applications. *Computers in Human Behavior, 26*(6), 1555-1563. <https://doi.org/10.1016/j.chb.2010.05.030>
- Verhoeven, L., & Perfetti, C. (2008). Advances in text comprehension: Model, process and development. *Applied Cognitive Psychology, 22*(3), 293–301. <https://doi.org/10.1002/acp.1417>

- Wayman, 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. <https://doi.org/10.1177/00224669070410020401>
- 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. <https://doi.org/10.1177/07419325050260040301>
- Yovanoff, P., Duesbery, L., Alonzo, J., & Tindal, G. (2005). Gradelevel invariance of a theoretical causal structure predicting reading comprehension with vocabulary and oral reading fluency. *Educational Measurement: Issues and Practice, 24*(3), 4-12. <https://doi.org/10.1111/j.1745-3992.2005.00014.x>

Bijlage 1**Overzicht van tekstversies en afnamevolgorde per conditie**

| Conditie A* | | Conditie B* | | Conditie C* | | Conditie D* | |
|-------------|----------|-------------|----------|-------------|----------|-------------|----------|
| Lin | C | Lin | I | Nina | C | Nina | I |
| Esther | I | Esther | C | Marit | I | Marit | C |
| Hans | I | Hans | C | Tom | I | Tom | C |
| Simon | C | Simon | I | Esther | C | Esther | I |
| Kas | I | Kas | C | Simon | I | Simon | C |
| Linda | C | Linda | I | Hans | C | Hans | I |
| Marit | C | Marit | I | Lin | C | Lin | I |
| Pauze | | | | | | | |
| Sara | I | Sara | C | Mark | I | Mark | C |
| Mark | C | Mark | I | Kas | C | Kas | I |
| Tom | I | Tom | C | Linda | I | Linda | C |
| Sanne | C | Sanne | I | Sanne | C | Sanne | I |
| Nina | I | Nina | C | Sara | I | Sara | C |
| <i>Lara</i> | <i>I</i> | <i>Lara</i> | <i>C</i> | <i>Loes</i> | <i>I</i> | <i>Loes</i> | <i>C</i> |
| <i>Loes</i> | <i>C</i> | <i>Loes</i> | <i>I</i> | <i>Lara</i> | <i>C</i> | <i>Lara</i> | <i>I</i> |

* Naam van de tekst gevolgd door tekstversie. C = consistente versie, I = inconsistente versie.