

## **B2-niveau als toelatingseis voor de universiteit: een goed idee?**

Een onderzoek naar de schrijfvaardigheid van NT2-studenten met B2-niveau

Scriptie MA Linguistics: Nederlands als tweede taal; docent en expert

**Naam:** Loes van den Heuvel

**Studentnummer:** S4381998

**Definitieve versie:** 01-08-2019

**1<sup>e</sup> lezer:** Prof. Dr. R. Schoonen

**2<sup>e</sup> lezer:** Dr. J. Klatter-Folmer

## Inhoudsopgave

Samenvatting .....	3
1. Introductie.....	4
1.1 Het ERK .....	4
1.2 Kritiek op het ERK.....	6
1.3 Toegang tot de universiteit .....	8
1.4 B2-niveau als geschikte toegangseis voor de universiteit .....	9
1.5 Huidig onderzoek .....	10
2. Methode.....	11
2.1 Participanten .....	11
2.2 Procedure.....	13
2.3 Instrumenten .....	13
2.4 Design.....	18
2.5 Data-analyse .....	18
3. Resultaten .....	19
3.1 Opdracht ‘samenvatting’ .....	19
3.2 Opdracht ‘e-mail over onderzoek’ .....	23
3.3 Verband tussen beoordeling taken.....	27
3.4 Tekstvakken.....	27
4. Discussie.....	28
4.1 Samenvatting .....	28
4.2 E-mail .....	30
4.3 Verband tussen beoordeling taken.....	31
4.4 Tekortkomingen .....	31
4.5 Aanbevelingen voor de praktijk .....	32
4.6 Aanbevelingen voor vervolgonderzoek.....	33
5. Conclusie .....	34
Referenties.....	35
Bijlagen .....	38
Bijlage 1: Originele tekst over hoogbegaafdheid .....	38
Bijlage 2: Tekst ‘samenvatting’, niveau B2- .....	42
Bijlage 3: Tekst ‘samenvatting’, niveau B2+ .....	43
Bijlage 4: Tekst ‘samenvatting’, niveau NT1.....	44
Bijlage 5: Tekst ‘e-mail over onderzoek’, niveau B2- .....	45
Bijlage 6: Tekst ‘e-mail over onderzoek’, niveau B2+ .....	46
Bijlage 7: Tekst ‘e-mail over onderzoek’, niveau NT1 .....	47
Bijlage 8: Overzicht vragen in Qualtrics.....	48

## Samenvatting

In deze studie is onderzocht of het veronderstelde B2-niveau van studenten die Nederlands als tweede taal leren (NT2-ers) op het gebied van schrijfvaardigheid voldoende is om in Nederland naar de universiteit te gaan. Dit werd gedaan door twee tekstsoorten, een samenvatting en een e-mail over een onderzoek, voor te leggen aan docenten van de Radboud Universiteit in Nijmegen, die ervaring hebben met lesgeven aan eerstejaarsstudenten. Deze teksten waren geschreven door studenten die al B2-niveau behaald hadden (B2+), door studenten die nog geen B2-niveau behaald hadden (B2-) en door moedertaalsprekers (NT1) van het Nederlands. De docenten moesten deze teksten beoordelen op de zeven aspecten *adequaatheid*, *begrijpelijkheid*, *grammaticale correctheid*, *correcte spelling*, *samenhang*, *logische opbouw* en *gepast woordgebruik*. Ook werd er bij iedere tekst gevraagd of de betreffende tekst als geheel voldeed aan het niveau dat de docent verwachtte van een eerstejaarsstudent aan hun opleiding. Het antwoord op deze vraag werd gezien als een algemeen oordeel over de tekst. Alle vragen moesten beantwoord worden op een 7-puntslikertschaal. De teksten van de moedertaalsprekers werden bij beide teksten op alle aspecten ruim voldoende beoordeeld. Ook werden ze op bijna alle aspecten significant beter dan de teksten van de tweedetaalleerders beoordeeld. Bij de samenvatting was er geen verschil tussen de scores van de B2- en de B2+ groep en werden veel aspecten van de B2+ groep als onvoldoende beoordeeld. Bij de e-mail werden de teksten van de B2+ groep beter beoordeeld dan de B2- groep. De aspecten van de e-mails van de B2+ groep werden ook allemaal voldoende beoordeeld. Verder bleek dat het algemene oordeel van beide teksten sterk samenhang met de beoordelingen van de afzonderlijke aspecten. Ten slotte werd er geen verband gevonden tussen de beoordeling van de twee verschillende taken. Deze uitkomsten laten zien dat er op basis van deze studie geen eenduidig antwoord is te geven op de vraag of B2-niveau een voldoende niveau is om NT2-ers toe te laten tot de universiteit. De uitkomsten van dit onderzoek tonen wel dat er meer dan één schrijfproduct van een student verzameld moet worden om het taalniveau van deze student vast te stellen.

## 1. Introductie

Niet iedereen die in Nederland woont, heeft het Nederlands als moedertaal verworven. Er zijn ook mensen die het Nederlands als tweede taal hebben geleerd. Deze mensen worden ook wel NT2-ers genoemd. Het leren van het Nederlands als tweede taal kan verschillende redenen hebben. Vluchtelingen moeten de taal bijvoorbeeld leren om in te burgeren en expats moeten soms de taal leren om in Nederland te kunnen werken. Sommige van deze NT2-ers willen in Nederland een studie volgen aan de universiteit of het hbo. Voor toegang tot een studie met Nederlands als voertaal moeten studenten het Nederlands op een bepaald niveau beheersen. De huidige studie richt zich op het schrijfvaardigheidsniveau dat verwacht wordt van NT2-ers die een studie aan de universiteit willen volgen.

### 1.1 Het ERK

Het vaststellen van het niveau van leerders van een tweede taal gebeurt tegenwoordig vaak met behulp van het Europees Referentiekader (ERK). Het ERK werd voor het eerst uitgebracht in 2001 (Council of Europe, 2001). In 2008 werd de eerste versie vertaald in het Nederlands. De doelstellingen van het ERK omvatten:

- De samenwerking tussen onderwijsinstellingen in verschillende landen te bevorderen en mogelijk te maken;
- Een solide basis te vormen voor wederzijdse erkenning van taalkwalificaties;
- Leerders, docenten, materiaalontwikkelaars, examencommissies en onderwijsbestuurders te helpen hun inspanningen te situeren en te coördineren. (Nederlandse Taalunie, 2008, p. 10)

Met het ERK werd er een gemeenschappelijke basis gelegd voor de uitwerking van lesprogramma's, leerboeken en dergelijke in heel Europa. Middels het ERK zouden ook belemmeringen weggenomen kunnen worden in de communicatie tussen deskundigen in het werkveld van verschillende moderne vreemde talen, belemmeringen die voortkwamen uit de verschillen tussen onderwijsstelsels in Europa. Door het referentiekader kon het toekennen van een bepaald taalniveau internationaal gestandaardiseerd worden. Dit doet het ERK door taalniveaus op te delen in zes vaardigheidsniveaus (Nederlandse Taalunie, 2008), achtereenvolgens van 'laag' naar 'hoog' A1, A2, B1, B2, C1 en C2. De A-niveaus staan in het ERK voor basisgebruiker, de B-niveaus voor onafhankelijke gebruiker en de C-niveaus voor vaardige gebruiker. De globale beschrijving voor niveau B2 luidt als volgt:

Kan de hoofdgedachte van een ingewikkelde tekst begrijpen, zowel over concrete als over abstracte onderwerpen, met inbegrip van technische besprekingen in het eigen

vakgebied. Kan zo vloeiend en spontaan reageren dat een normale uitwisseling met moedertaalsprekers mogelijk is zonder dat dit voor een van de partijen inspanning met zich meebrengt. Kan duidelijke, gedetailleerde tekst produceren over een breed scala van onderwerpen; kan een standpunt over een actuele kwestie uiteenzetten en daarbij ingaan op de voor- en nadelen van diverse opties. (Nederlandse Taalunie, 2008, p. 26)

In deze beschrijving zijn alle vier de vaardigheden lezen, spreken, luisteren en schrijven opgenomen. Naast deze globale beschrijving zijn er zogenaamde ‘can-do’ descriptorren geformuleerd voor de communicatieve activiteiten receptie (luisteren en lezen), interactie en productie (spreken en schrijven) en voor strategieën die worden toegepast bij deze communicatieve activiteiten. De can-do descriptor van niveau B2 op het gebied van schrijven luidt als volgt:

Ik kan een duidelijke, gedetailleerde tekst schrijven over een breed scala van onderwerpen die betrekking hebben op mijn interesses. Ik kan een opstel of verslag schrijven, informatie doorgeven of redenen aanvoeren ter ondersteuning vóór of tégen een specifiek standpunt. Ik kan brieven schrijven waarin ik het persoonlijk belang van gebeurtenissen en ervaringen aangeef. (Nederlandse Taalunie, 2008, p. 28)

Dit is een algemene beschrijving van wat leerders op B2-niveau moeten kunnen op het gebied van schrijven. Per taalniveau zijn er verscheidene can-do descriptorren voor verschillende communicatieve activiteiten opgesteld. Zo wordt er ook onderscheid gemaakt tussen domeinen waarin de taalhandelingen plaatsvinden, namelijk persoonlijk, publiek, professioneel en educatief. Ook wordt er onderscheid gemaakt tussen communicatieve thema's zoals dagelijks leven, persoonlijke identificatie en vrije tijd. Er worden voor communicatieve activiteiten per niveau voorbeelden van activiteiten gegeven. Voorbeelden van communicatieve schrijfactiviteiten zijn een persoonlijke of zakelijke brief schrijven of creatief en fantasierijk schrijven. Sommige van deze activiteiten hebben ook nog toelichtende schalen. Zo worden voor de activiteiten ‘creatief schrijven’ en ‘verslagen en essays’ nog per niveau meer uitleg gegeven door middel van deze toelichtende schaal. De toelichtende schalen van creatief schrijven op niveau B2 zijn als volgt:

- Kan helder en gedetailleerd schrijven over waargebeurde of aan de verbeelding ontsproten gebeurtenissen en ervaringen, en daarbij het verband tussen ideeën aangeven in heldere samenhangende tekst en zich houden aan de aanvaarde conventies van het betrokken genre.

- Kan helder en gedetailleerd schrijven over een breed scala van onderwerpen die betrekking hebben op zijn of haar interessegebied. Kan een recensie schrijven over een film, boek of toneelstuk. (Nederlandse Taalunie, 2008, p. 60)

Het ERK besteedt daarnaast ook nog aandacht aan welke handelingen een taalleerder moet kunnen uitvoeren voor de verschillende vaardigheden. Voor schrijven zijn dit bijvoorbeeld de cognitieve en linguïstische vaardigheden ‘de boodschap kunnen organiseren en formuleren’ en de handmatige vaardigheden ‘de tekst kunnen opschrijven of typen. Samenvattend heeft het ERK voor iedere vaardigheid en voor subvaardigheden op verschillende vlakken, op verschillende niveaus en in verschillende categorieën verwachtingen geformuleerd in de vorm van can-do descriptoren. Aan de hand van deze can-do descriptoren kan er dus vastgesteld worden of taalleerders specifieke onderdelen van een tweede taal beheersen.

In 2018 is er een supplement van het ERK gepubliceerd (Council of Europe, 2018). Een Nederlandse vertaling hiervan is snel daarna verschenen (Nederlandse Taalunie, 2018). Een belangrijke wijziging ten opzichte van het document in 2001 is de toevoeging van het concept ‘bemiddeling’ als talige activiteit, naast de al eerder genoemde receptie, productie en interactie. Onder bemiddeling valt bijvoorbeeld ‘ruimte en voorwaarden creëren voor communicatie en leren’ en ‘samenwerken om nieuwe betekenissen te construeren’ (Nederlandse Taalunie, 2018, p. 96). Het gaat bij bemiddeling ook om het zich kunnen inleven in het perspectief van iemand anders en een situatie vanuit meerdere kanten kunnen bekijken. Daarnaast is er in het gereviseerde document aandacht voor leerders onder het A1-niveau, voor gebruikers van gebarentaal en jongeren en zijn er reeds bestaande schalen geherformuleerd. Wat betreft schrijven zijn er enkele wijzigingen doorgevoerd. Zo is de toelichtende schaal ‘correspondentie’ bijvoorbeeld uitgebreid met formele correspondentie; in het eerste document werd er alleen aandacht besteed aan persoonlijke correspondentie. Ook wordt er aandacht besteed aan online interactie. Ten slotte wordt er in het geüpdatete document ook een beoordelingsschaal voor schriftelijke productie gegeven voor de verschillende taalniveaus en voor de kenmerken algemeen, reikwijdte, coherentie, nauwkeurigheid, beschrijving en argument. Samenvattend is het een ERK, zeker na de update in 2018, een uitgebreid werk dat veel duidelijkheid en uniformiteit kan verschaffen bij het leren, onderwijzen en beoordelen van een bepaald taalniveau in een tweede taal.

## **1.2 Kritiek op het ERK**

Het ERK heeft sinds het jaar van publicatie een grote invloed gehad op de onderwijs- en toetspraktijk in Europa en daarbuiten. In het gereviseerde document (Council of Europe, 2018;

Nederlandse Taalunie, 2018) gaven de makers al aan dat er ook kritiek was op het oorspronkelijke document (Council of Europe, 2001; Nederlandse Taalunie, 2008). Zo was het oorspronkelijk bedoeld als een descriptieve collectie van taalvaardigheidsniveaus, maar werd het al snel normatief gebruikt. Daarnaast gaven de makers ook aan dat er kritiek was dat de ERK-niveaus onvoldoende empirische basis zouden hebben en dat ze niet concreet genoeg waren om in de praktijk te gebruiken. Een laatste kritiekpunt dat genoemd werd, was dat het ERK niet inspeelde op moderne communicatievormen. Als antwoord op deze kritiekpunten werd de eerdergenoemde revisie van het ERK uitgebracht (Council of Europe, 2018; Nederlandse Taalunie, 2018). Sommige punten zijn in het document in 2018 concreet aangepakt, zoals het inspelen op moderne communicatievormen. Het moet de komende jaren blijken of deze revisie inderdaad alle kritiekpunten weg heeft genomen.

Naast de kritiek die in het ERK-document zelf werd genoemd, hebben ook verschillende onderzoekers in hun afzonderlijke onderzoeken de afgelopen jaren kritiek geleverd op het ERK. Figueras (2012) stelde dat het ERK geen model van absolute perfectie is. Ze noemde dat taalcursussen nu vaak gericht zijn op de uitkomst en vooral aandacht besteden aan de can-do descriptoren van het ERK. Dit kan weer leiden tot ‘teaching to the test’, het fenomeen dat cursussen alleen gemaakt zijn om zo goed mogelijk op de test die de cursus afsluit te presteren (Hughes, 2013). Ook noemde ze andere tekortkomingen, zoals het niet definiëren van constructen als ‘lang’, ‘kort’ en ‘bekend zijn met’ en de incoherentie van terminologie (betekent ‘lezen’ bijvoorbeeld hetzelfde als ‘begrijpen’?). Bovendien stipte ze ook aan dat het ERK vaak oneigenlijk gebruikt wordt; het was oorspronkelijk bedoeld om te gebruiken bij het leren van een vreemde taal in een context van volwassenen, maar het wordt nu gebruikt in allerlei soorten onderwijs, zoals ook het middelbaar onderwijs en zelfs binnen de context van het leren van een moedertaal (Martyniuk & Noijons, 2007). Wisniewski (2017) had soortgelijke bevindingen en gaf aan dat de schalen van het ERK vaak gebruikt worden op een manier die van tevoren niet voorzien was. Volgens Harsch (2018) zou het ERK niet overal lukraak moeten worden ingezet. Voor iedere test of cursus zou een behoefteanalyse gedaan moeten worden wat betreft de taken en doelen van een test, omdat iedere test een individueel en uniek doel dient. Harsch (2018) gaf daarbij wel aan dat het niet kunnen vergelijken van testuitkomsten niet te verwijten is aan het ERK, maar wel aan de mensen die verschillende testen ontwerpen en interpreteren. Haar grootste bezwaar tegen het ERK is echter dat het vaak in contexten gebruikt wordt waarvoor het in eerste instantie niet bedoeld was, zoals ook hierboven beschreven. Green (2018) noemde daarbij dat ERK-niveaus niet als basis moeten worden gebruikt om high-stake beslissingen te nemen. Dat zijn beslissingen met mogelijk vergaande consequenties, zoals het wel of niet

toelaten van studenten tot een opleiding (Hughes, 2013).

Hulstijn (2007) ging verder in op twee dimensies die hij de kwalitatieve en kwantitatieve dimensie noemt. Volgens deze auteur houdt de kwantitatieve dimensie in ‘wat iemand moet kunnen’ en kwalitatieve dimensie ‘hoe goed iemand activiteit x moet kunnen’. Hij stelde dat er drie typen T2-leerders zijn: leerders die kwantitatief maar een paar dingen kunnen, maar van wie de prestaties kwalitatief hoogstaand zijn; leerders die kwantitatief veel dingen kunnen, maar van wie de prestaties kwalitatief niet van hoge kwaliteit zijn en leerders waarvan de kwaliteit en kwantiteit van hun taalgebruik overeenkomt, zoals ook het ERK veronderstelt. Volgens Hulstijn (2007) is de veronderstelling van het ERK dat alle leerders onder het laatste type vallen problematisch, omdat er dus ook andere typen leerders zijn. Daarnaast haalde Hulstijn (2011) het probleem aan dat in de meeste beoordelingsschalen, zo ook in die van het ERK, een hoger niveau van taalvaardigheid bij een persoon inhoudt dat ook het intellectueel functioneren beter zou moeten zijn. Dat zou impliceren dat personen met lagere intellectuele capaciteiten niet een hoog niveau van taalvaardigheid in een T2 zouden kunnen bereiken.

### **1.3 Toegang tot de universiteit**

Het ERK blijft echter nog steeds een document dat wijdverspreid gebruikt wordt. Zo moeten studenten die een andere moedertaal dan het Nederlands hebben niveau B2 in het Nederlands beheersen, om een opleiding aan een Nederlandse universiteit te mogen volgen (Dienst Uitvoering Onderwijs, n.d.). Dit kunnen ze doen door het Staatsexamen (STEX) Programma II te halen. Waarom dat precies het toegangsniveau voor de universiteit is, is niet altijd duidelijk. Deygers, Zeidler, Vilcu, en Carlsen (2018) vonden in hun onderzoek dat veel universiteiten al snel ‘doen wat anderen doen’ als het gaat om taaleisen stellen, in plaats van een behoefteanalyse doen over wat echt nodig is wat betreft taalniveau voor een bepaalde context. Dit ‘kopiëren’ lijkt dan ook vaak de reden dat B2 als toegangseis voor de universiteit gesteld wordt.

Harsch (2018) onderzocht of het ERK wel geschikt is om te gebruiken voor het instellen van universitaire toegangsnormen. Ze gaf aan dat het ERK ervoor kan zorgen dat entreetoetsen en toelatingseisen transparanter worden. Volgens haar is het echter wel zo dat verschillende testen verschillende zaken kunnen meten en taalvaardigheid op een andere manier operationaliseren, maar wel hetzelfde ERK-level kunnen ‘opleveren’. Dat kan problematisch zijn. Als taalvaardigheid immers op een andere manier geoperationaliseerd wordt, kunnen verschillende testen ook moeilijk met elkaar vergeleken worden via hetzelfde referentiekader. Het probleem is hier niet per definitie het ERK, maar wel de testen en de verschillende constructen die gemeten worden. Deygers, Van Gorp en Demeester (2018) stelden dan ook dat descriptoren eigenlijk aangepast moeten worden voor specifieke doeleinden, zoals toegang tot



de universiteit, omdat ze ongeschikt zijn voor onaangepast gebruik in beoordelingsschalen. Ze toonden in hun studie aan dat twee verschillende testen voor het meten van Nederlandse taalvaardigheid niet hetzelfde meten, maar wel hetzelfde ERK-niveau, namelijk B2, ‘opleveren’.

#### **1.4 B2-niveau als geschikte toegangseis voor de universiteit**

Verschillende studies toonden aan dat het eerdergenoemde B2-niveau wel geschikt is als toegangseis om aan een universitaire studie te beginnen. Harsch, Ushioda en Ladroue (2017) onderzochten de predicatieve validiteit van de TOEFL; ze wilden weten of het behalen van B2-niveau volgens de Engelstalige TOEFL-test ook betekende dat studenten goed mee konden komen in hun opleiding in het Verenigd Koninkrijk. Dit deden ze op een kwantitatieve manier, door middel van het analyseren van de relatie tussen de scores van de TOEFL en scores op de universiteit en op een kwalitatieve manier, door middel van vragenlijsten en interviews. Ook tutoren werden uitgenodigd om mee te doen in het kwalitatieve deel van de studie. Uit de studie kwam naar voren dat studenten die slaagden voor de toets goed mee konden komen met de talige eisen van een universitaire studie in het Verenigd Koninkrijk en dat de meerderheid ook daadwerkelijk afstudeerde aan de betreffende opleiding. Carlsen (2018) onderzocht het succes van buitenlandse studenten in het hoger onderwijs in Noorwegen in relatie tot het vaardigheidsniveau van de Noorse taal van deze studenten. Dit deed ze door middel van een vragenlijst. Ook zij vond dat studenten minimaal B2-niveau nodig hadden om mee te kunnen met de talige eisen van het Noorse hoger onderwijs. Bechger, Kuijper en Maris (2009) vroegen aan 57 belanghebbende personen binnen universiteiten en hogescholen of ze het minimale benodigde taalniveau van startende studenten aan konden geven. Deze personen moesten dit doen door aan te geven welk can-do descriptoren ze verwachtten van eerstejaarsstudenten als startkwalificatie. Gebaseerd op deze aanpak vonden de onderzoekers dat B2-niveau vereist is voor studenten die aan een studie aan de universiteit of hogeschool willen beginnen. Gebaseerd op deze studies zou B2-niveau dus een geschikte eis zijn om te beginnen aan een studie aan de universiteit.

Er zijn echter ook studies die het tegenovergestelde vonden. Deygers, Van den Branden en Van Gorp (2018) deden onderzoek naar NT2-studenten die een opleiding aan een universiteit in België volgden. Dit deden ze door resultaten van de studenten te onderzoeken en studenten én docenten te interviewen in focusgroepen. Ze vonden dat NT2-studenten met B2-niveau niet altijd klaar waren om lessen in het Nederlands te volgen, onder andere vanwege het tempo en het accent van de docenten. De T2-studenten gaven voor schrijven aan dat de docenten vaak enige coulance hanteerden vanwege hun taalachtergrond. Field (2011) bestudeerde T2-Engels

leerders op B2-niveau aan de universiteit, voornamelijk in de vaardigheid luisteren. Dit deed hij door bij het maken van luisteropgaves studenten hardop te laten nadenken en dit op te nemen. Hij concludeerde dat de studenten moeite hadden met de vaardigheid luisteren op B2-niveau.

Bovenstaande onderzoeken zijn dus niet eenduidig over of studenten met B2-niveau klaar zijn om te voldoen aan de eisen van de universiteit. Dat geeft ook aan dat er meer onderzoek nodig is. Verschillende onderzoekers die hierboven werden aangehaald pleitten ook voor meer onderzoek dat lokale onderwijscontexten meeneemt (o.a. Carlsen, 2018; Harsch, 2018). De onderzoeken die hierboven zijn aangehaald zijn namelijk uitgevoerd in bepaalde contexten. Met deze contexten wordt vooral bedoeld dat de onderzoeken zijn uitgevoerd in bepaalde landen met bepaalde talen. De resultaten en conclusies die uit deze onderzoeken voortvloeien kunnen daarom niet per definitie overgenomen worden in andere contexten. Om uitspraken te doen over het niveau dat in Nederland nodig is om naar de universiteit te gaan, moet er dan ook meer onderzoek worden gedaan dat zich toespitst op de situatie in Nederland. Het onderzoek van Bechger e.a. (2009) is een goed startpunt voor de Nederlandse context, maar dit onderzoek richtte zich op welke can-do descriptorren docenten verwachten van eerstejaarsstudenten. Er zijn echter meer invalshoeken nodig om vast te stellen welk taalniveau eerstejaarsstudenten moeten hebben. Daarom wordt er in het huidige onderzoek een andere invalshoek gekozen en zal er worden ingegaan op wat docenten van teksten vinden die geschreven zijn door NT2-ers, om zo meer te weten te komen over het gewenste taalniveau van eerstejaarsstudenten.

### **1.5 Huidig onderzoek**

In dit onderzoek wordt er gekeken naar de schrijfvaardigheid van studenten die Nederlands als tweede taal hebben geleerd bij Radboud in'to Languages, het expertisecentrum voor taal en communicatie van de Radboud Universiteit in Nijmegen. De hoofdvraag is dan ook of docenten van de Radboud Universiteit het taalniveau van studenten met minimaal B2-niveau hoog genoeg vinden voor eerstejaarsstudenten en dus of B2-niveau een goede toegangseis is om studenten toe te laten tot de universiteit. Het onderzoek focust zich op de Nederlandse context en meer specifiek op één universiteit, namelijk de Radboud Universiteit in Nijmegen. Daarnaast richt het onderzoek zich ook op specifiek op één vaardigheid, namelijk schrijfvaardigheid.

In dit onderzoek worden beoordelingen door docenten van teksten van NT2-ers met verschillende taalniveaus geanalyseerd. Er wordt verwacht dat teksten van moedertaalsprekers beter beoordeeld worden dan teksten van tweedetaalleerders. Binnen de groep tweedetaalleerders wordt er verwacht dat teksten van studenten met een niveau hoger dan B2 (B2+) beter beoordeeld worden dan teksten van studenten met een niveau lager dan

B2 (B2-). Of niveau B2 inderdaad een voldoende niveau is voor eerstejaarsstudenten aan de universiteit, is echter moeilijk in te schatten op basis van bovenstaande literatuur. Op basis van de toelatingseisen die de universiteiten in Nederland stellen, wordt er wel verwacht dat B2-niveau een voldoende niveau zou moeten zijn voor eerstejaarsstudenten.

De teksten worden beoordeeld op dezelfde aspecten als in het STEX II. In dit onderzoek wordt namelijk ook onderzocht welke van deze aspecten belangrijk zijn voor de docenten bij een algemene beoordeling over of de betreffende NT2-er voldoet aan het gewenste niveau.

Als laatste wordt er ook gekeken naar de correlatie tussen de toegekende scores van deze twee teksten. De verwachting is dat deze correlaties hoog, positief en significant zijn, omdat verwacht wordt dat het schrijfvaardigheidsniveau van één schrijver over twee teksten constant is.

## 2. Methode

### 2.1 Participanten

In dit onderzoek waren er twee soorten participanten: studenten die teksten schrijven en docenten die deze teksten beoordelen. Deze participanten worden hieronder apart besproken.

**2.1.1 Studenten.** Er werden drie soorten studenten benaderd met de vraag of ze teksten voor deze studie beschikbaar wilden stellen. De grootste groep studenten waren studenten die een cursus Nederlands van ERK-niveau B2 tot niveau C1 volgden bij Radboud in'to Languages, het expertisecentrum voor taal en communicatie van de Radboud Universiteit in Nijmegen. Deze studenten hadden allemaal al eerder via een taal cursus of via het Staatsexamen programma II B2-niveau behaald. Ook krijgen alle studenten voor aanvang van een taal cursus bij in'to Languages een intakegesprek om het taalniveau van de studenten vast te stellen. Er werd daarom aangenomen dat deze studenten inderdaad B2-niveau hadden. Uit deze groep deden er twee mannen en tien vrouwen mee. Van één vrouw is geen verdere achtergrondinformatie bekend. De gemiddelde leeftijd van deze participanten was 26.91 ( $SD = 5.50$ ) jaar. Ook werd de deelnemers naar hun land van herkomst en moedertaal gevraagd en werd gevraagd hoeveel jaar ze al in Nederland waren. De studenten kwamen uit Oekraïne, China, Irak, Engeland, Syrië, Libanon, Polen, Australië, Italië en Frankrijk en hun moedertaal was respectievelijk Oekraïns/Russisch, Chinees, Koerdisch, Engels, Arabisch, Libanees, Pools, Engels, Turks, Italiaans en Frans. Gemiddeld waren deze studenten al 5.36 ( $SD = 8.51$ ) jaar in Nederland.

Ook werden er studenten benaderd die nog geen B2-niveau hadden behaald, om als

controlegroep te dienen. Deze studenten zijn ook via Radboud in'to Languages geworven. Aan deze studenten werd ook gevraagd om dezelfde teksten te schrijven als de studenten uit de B2-C1 groep. Deze studenten kwamen allemaal uit een groep die van taalniveau A2 op weg was naar niveau B1. Vanuit deze groep deden er één man en twee vrouwen mee. De gemiddelde leeftijd van deze participanten was 38.00 ( $SD = 7.07$ ) jaar. Van één vrouw die teksten aanleverde is geen achtergrondinformatie bekend. De overige twee studenten kwamen uit Syrië. De moedertaal van deze studenten was Arabisch. Gemiddeld waren de twee studenten beiden al 3.00 jaar ( $SD = 0$ ) in Nederland.

Daarnaast werden er ook moedertaalsprekers van het Nederlands gevraagd om dezelfde teksten te schrijven als andere studenten, eveneens om als controlegroep te dienen. Vanuit deze groep deden er één man en vier vrouwen mee. De gemiddelde leeftijd van deze participanten was 24.60 jaar ( $SD = 0.55$ ). De studenten kwamen van verschillende opleidingen van de Radboud Universiteit: Rechten, Taalwetenschap, Communicatiewetenschap, Bedrijfskunde en Sociologie. Er is voor studenten van deze opleidingen gekozen omdat er ook docenten van deze opleidingen en van faculteiten waaraan deze opleidingen verbonden zijn meededen als beoordelaar van de teksten. In Tabel 1 is een overzicht te zien van het aantal studenten en de daarbij behorende informatie.

Tabel 1

*Aantal studenten die teksten aanleverden naar ERK-taalniveau, geslacht, leeftijd en aantal jaar in Nederland*

Niveau <sup>a</sup>	Aantal (waarvan vrouw)	Leeftijd (SD)	Aantal jaar in Nederland (SD)
B2-	3 (2)	38.00 (7.07)	3.00 (0)
B2+	12 (10)	26.91 (5.50)	5.36 (8.51)
NT1	5 (4)	24.60 (0.55)	n.v.t.

<sup>a</sup> B2- staat voor studenten met taalniveau lager dan B2, B2+ voor studenten met een taalniveau hoger dan B2 en NT1 staat voor moedertaalsprekers van het Nederlands.

**2.1.2 Docenten.** Docenten werden allemaal benaderd via e-mail. De e-mailadressen zijn verkregen via secretaresses van faculteiten en opleidingen en via de studiegidsen van de Radboud Universiteit. Uiteindelijk wilden er 22 docenten van vier verschillende faculteiten van de Radboud Universiteit meedoen: Rechten, Sociale Wetenschappen, Geesteswetenschappen en Managementwetenschappen. Alle docenten hadden ervaring met lesgeven in het eerste jaar van de betreffende studie. Uiteindelijk deden er veertien docenten mee aan dit onderzoek; twee

van de faculteit Rechten, vier van de faculteit Sociale Wetenschappen, twee van de faculteit Managementwetenschappen en zes van de faculteit Letteren. De gemiddelde leeftijd van de docenten van de verschillende faculteiten samen was 46.57 ( $SD = 14.83$ ). In Tabel 2 is een overzicht te zien van de deelnemende docenten naar faculteit, geslacht en leeftijd.

Tabel 2  
*Aantal docenten naar faculteit, geslacht en leeftijd*

Faculteit	Aantal (waarvan vrouw)	Leeftijd (SD)
Rechten	2 (0)	47.5 (10.61)
Sociale Wetenschappen	4 (3)	34.5 (14.39)
Letteren	6 (0)	57.5 (2.12)
Managementwetenschappen	2 (0)	50.7 (15.47)

## 2.2 Procedure

Deze studie heeft goedgekeuring gekregen van de Ethische Toetsingscommissie Geesteswetenschappen van de Radboud Universiteit Nijmegen. Alle studenten hebben toestemming gegeven voor het gebruik van hun teksten voor dit onderzoek. Alle studenten hebben ook de informatiebrief gelezen. Deze werd voor de B2-C1 en de A2-B1 groep uitgedeeld tijdens de taallessen. Voor de moedertaalsprekers werd deze brief via e-mail verstuurd.

Op het toestemmingsformulier werd er voor de niet-moedertaalsprekers van het Nederlands ook nog gevraagd naar hun geboortedatum, geslacht, moedertaal, land van herkomst en het aantal jaren dat ze in Nederland waren. Voor deelnemers die wel Nederlands als moedertaal hadden, werd er gevraagd naar hun geboortedatum en geslacht. De teksten werden allemaal online aangeleverd door de studenten zelf of (met toestemming) door de taaldocenten van de studenten. De teksten werden via Qualtrics verstuurd aan de beoordelende docenten. Dit gebeurde in mei 2019. Alle docenten kregen een e-mail met daarin een link om de beoordelingen over de teksten in te vullen. Dit was een herbruikbare link, zodat de docenten hem eventueel ook nog verder konden verspreiden onder collega's. De docenten konden de teksten beoordelen waar en wanneer ze zelf wilden. Ze hadden twee weken om dit te doen. Zeven dagen na het versturen van de teksten werd er een herinnering aan de docenten gestuurd.

## 2.3 Instrumenten

In dit onderzoek zijn twee instrumenten gebruikt. Ten eerste waren er de stimuli, namelijk de

teksten die de studenten hebben geschreven. Deze teksten zijn daarna via Qualtrics verstuurd naar de docenten, zodat zij de teksten konden beoordelen. De teksten en de ‘vragenlijst’ in Qualtrics om de teksten te beoordelen worden hieronder apart toegelicht.

**2.3.1 Teksten.** De meeste teksten die gebruikt zijn voor deze studie waren reeds geschreven in de eerdergenoemde B2-C1 cursus. Per student zijn er twee teksten voor dit onderzoek gebruikt. Er is gekozen voor twee teksten omdat één tekst te weinig werd bevonden om uitspraken te doen over het niveau van schrijfvaardigheid van een persoon (Schoonen, 2005). Het opnemen van meer dan twee teksten was niet haalbaar binnen de cursus waarin de participanten deelnamen en binnen de tijdspanne van dit onderzoek. Uiteindelijk is er van vier studenten uit de B2+ groep en één student uit de B2- groep maar één tekst opgenomen in plaats van twee, omdat zij maar één tekst hadden aangeleverd. Voor de B2+ groep waren er drie studenten die geen e-mailtekst hadden aangeleverd en één student die geen samenvatting had aangeleverd. De betreffende student met B2- niveau had geen samenvatting aangeleverd. Er is besloten om de teksten die deze studenten wel hadden geschreven toch op te nemen, omdat zo het aantal teksten verhoogd werd en de teksten wel door een groot deel van de deelnemende docenten gelezen werden, zodat er een duidelijk oordeel over geveld kon worden. De teksten die voor de B2-C1 cursus als huiswerk waren geschreven zijn leidend geweest voor de keuze van het aantal en het soort teksten. Dit was namelijk de grootste groep studenten in dit onderzoek. Daarnaast was het makkelijk om teksten te gebruiken die al geschreven waren voor de taal cursus. Als type tekst is er gekozen voor een samenvatting van een tekst over hoogbegaafdheid en een e-mail over een onderzoek dat de studenten zelf mochten kiezen. Er is voor deze soort teksten gekozen omdat deze het meest geschikt waren in vergelijking met andere teksten die in de cursus geschreven werden. Zo zou een open sollicitatie veel persoonlijke informatie van studenten kunnen bevatten en minder geschikt zijn binnen een universitaire context. Een klachtenbrief zou ook niet geschikt zijn om te gebruiken binnen een universitaire context. Een samenvatting maken van een tekst en het schrijven over een onderzoek zijn daarentegen realistische opdrachten binnen een universitaire context. De teksten die gebruikt werden in dit onderzoek zijn daarom ook representatief voor de ‘target language use’, wat belangrijk is volgens Bachman en Palmer (1996). In de taal cursus B2-C1 werd ook een betoog geschreven. Dit zou ook een geschikte tekstsoort voor dit onderzoek zijn geweest, maar het moment van schrijven in de taal cursus was te laat binnen de tijdspanne van dit onderzoek om deze tekstsoort op te nemen als tekst.

De opdracht ‘samenvatting’ was als volgt:

*'Je bent derdejaarsstudent aan de Academische Pabo, een combinatie van de bacheloropleiding Leraar Basisonderwijs (pabo) en de bacheloropleiding Pedagogische Wetenschappen. Voor een werkcollega krijg je de opdracht een voor jou interessant artikel over onderwijs samen te vatten en te delen met je medestudenten. Je hebt gekozen voor het artikel 'Beter omgaan met hoogbegaafde leerlingen'. De samenvatting mag maximaal tien procent van de originele tekst zijn. Maak er een goedlopend verhaal van.'*

Voor de docenten die de teksten gingen beoordelen werd de zin *'De samenvatting mag maximaal tien procent van de originele tekst zijn'* vervangen door *'De samenvatting mag maximaal 400 woorden zijn, ongeveer 1 A4'*. Dit om te voorkomen dat er verwarring zou ontstaan over deze zin; de zin zou immers kunnen betekenen dat de gehele samenvatting tien procent van het aantal woorden uit de originele tekst mocht zijn óf dat er tien procent van de (letterlijke) woorden uit de originele tekst in de samenvatting gebruikt mochten worden. Ook is er zo geprobeerd te voorkomen dat docenten zich te veel gingen richten op het aantal woorden waaruit de samenvatting bestond, wat niet de bedoeling was en alleen maar extra moeite voor de docenten zou opleveren. De originele tekst over hoogbegaafdheid is in bijlage 1 te vinden. De docenten konden de oorspronkelijke tekst over hoogbegaafdheid lezen via een link, maar dit was niet verplicht. Er werd de docenten dan ook gevraagd om de samenvattingen als zelfstandige teksten te beoordelen.

De opdracht 'e-mail over onderzoek' was als volgt:

*'Zoek een klein artikel (bijvoorbeeld op <http://www.scientias.nl> of <http://www.kennislink.nl>). Lees dit artikel goed door. Een collega of studiegenoot is geïnteresseerd in het onderzoek waarover je hebt gelezen. Je schrijft hem/haar een e-mail van 150-300 woorden waarin je uitlegt wat het onderzoek inhoudt. Probeer het artikel in eigen woorden na te vertellen. Wat is er precies onderzocht en op welke manier is het onderzoek gedaan? Geef ook een reactie op dit onderzoek: wat is jouw mening hierover?'*

De zin *'probeer het artikel in eigen woorden na te vertellen'* is voor docenten weggelaten. De reden hiervoor was dat er niet in iedere e-mail een link van het oorspronkelijke onderzoek vermeld werd. Als er geen link opgenomen was in de opdracht, zou het lastig zijn voor docenten om na te gaan of studenten het artikel inderdaad in hun eigen woorden hadden verteld. Daarnaast zou het nagaan van het gebruik van eigen woorden ook een extra last voor de docenten zijn geweest. Als er wel een bron van het onderzoek werd vermeld, is deze niet

weggehaald. De docenten mochten deze, indien aanwezig, ook bekijken, maar er werd ze ook hier gevraagd om de e-mails als zelfstandige teksten te beoordelen. Bij de e-mails zijn alle namen in de aanhef en afsluiting van de e-mail weggehaald en vervangen door '<naam>'. Als er geen aanhef of afsluiting in de tekst van de oorspronkelijke schrijver aanwezig was, werd deze ook niet toegevoegd. Aan de opmaak van beide teksten is verder niets veranderd.

**2.3.2 Vragenlijst.** Aan de docenten werden de teksten via Qualtrics verstuurd. Er waren twee versies: één versie bevatte alle samenvattingen en één versie bevatte alle e-mails over onderzoek. Hier is voor gekozen omdat de tijd die de docenten moesten investeren in het beoordelen van de teksten anders te veel zou worden (meer dan één uur) en er daardoor wellicht docenten zouden afhaken. De versies werden evenredig verdeeld over het aantal docenten dat mee wilde doen. Elf docenten ontvingen zo de versie over de samenvattingen en elf docenten de versie over de e-mails over onderzoek. Vijf docenten vulden de versie over de e-mails in, negen docenten de versie over de samenvattingen. Bij de e-mails was er één docent die maar de helft van de teksten had beoordeeld, maar vanwege het geringe aantal docenten die deze versie had ingevuld, is deze docent wel meegenomen in de resultaten. Aan het begin werd de informatiebrief voor docenten getoond. Daarna moesten de docenten toestemming geven om mee te doen aan het onderzoek. Vervolgens kregen de docenten enkele achtergrondvragen. Er werd gevraagd naar hun geslacht, leeftijd en opleiding waaraan ze les gaven. Daarna kregen ze uitleg over het onderzoek. Er werd verteld dat ze één soort tekst te zien kregen, dus ofwel de opdracht 'samenvatting', ofwel de opdracht 'e-mail over onderzoek'. Vervolgens werd de uitleg van de opdracht getoond. Er werd verteld dat de teksten ongecorrigeerd overgenomen waren en dat alle namen bij de aanhef en afsluiting van de e-mail waren weggehaald, om anonimiteit van de tekstschrijvers te waarborgen. Bovendien werd er vermeld dat er ook aan de opmaak van de teksten niets veranderd was. Bij de samenvattingen konden de docenten de oorspronkelijke tekst lezen via een link en in sommige e-mails stond ook een link naar het beschreven onderzoek. Er werd vermeld dat de docent deze oorspronkelijke teksten mocht lezen, maar de e-mails als zelfstandig leesbare teksten moesten beoordelen. Ook kregen de docenten nog een aanwijzing om de tekst te zien als een representatieve weergave van de schrijfvaardigheid van de student en om af te gaan op hun eerste ingeving bij het beoordelen. Daarna werd er vermeld dat er voor opmerkingen of eventuele uitleg van de keuze(s) van de docenten er bij iedere tekst een leeg tekstvak toegevoegd was. Als laatste werd er verteld dat er een balk in beeld verscheen die de voortgang liet zien en dat de docenten voor de samenvattingen en de e-mails respectievelijk achttien of zeventien teksten te zien kregen.

**2.3.2.1 Beoordeling van teksten.** Bij het beoordelen van de teksten kregen de



docenten alle teksten en bijbehorende vragen te zien. Aan de docenten werd niet verteld dat de teksten waren geschreven door studenten met verschillende taalniveaus, om ervoor te zorgen dat dit hun beoordelingen niet zou beïnvloeden. Alle docenten kregen alle samenvattingen of alle e-mails over onderzoek te zien, in gerandomiseerde volgorde. Dit werd gedaan om sequentie-effecten tegen te gaan. Bij een sequentie-effect heeft de voorgaande beoordeling invloed op de volgende beoordeling. Dit kan bijvoorbeeld beteken dat de eerste waarneming altijd te laag wordt beoordeeld. Andere vormen van het sequentie-effect kunnen optreden wanneer bijvoorbeeld na een reeks slechte producten een product van middelmatige kwaliteit moet worden beoordeeld en zo een te hoog cijfer krijgt of wanneer er een grote hoeveelheid producten moet worden beoordeeld en beoordelaars vermoeid worden, waardoor strengheid in beoordeling toe- of afneemt (Meuffels, 1994). Door voor iedere beoordelaar de teksten in een andere volgorde aan te bieden, kunnen deze mogelijke sequentie-effecten beperkt worden. Boven de tekst werd de opdracht voor de betreffende tekst herhaald. Boven iedere tekst stond ook nog de titel van de opdracht: ‘Samenvatting: beter omgaan met hoogbegaafde leerlingen’ of ‘E-mail over onderzoek’. Per tekst kregen de docenten eerst de vraag in hoeverre ze de volgende aspecten aanwezig in of kenmerkend voor deze tekst vonden: *Adequaatheid* (de tekst is een uitwerking van de opdracht), *Begrijpelijkheid* (de tekst is goed te begrijpen), *Grammaticale correctheid* (bv. gebruik van wel/geen lidwoord, congruentie van werkwoord en subject), *Correcte spelling* (woorden zijn op de juiste manier geschreven), *Samenhang* (er is een verband tussen zinnen en alinea’s, bijvoorbeeld door signaal- en/of verbindingswoorden), *Logische opbouw* (de tekst is logisch opgebouwd, bv. inleiding, kern, slot) en *Gepast woordgebruik* (het woordgebruik sluit aan bij het onderwerp van de opdracht). Deze vragen moesten docenten beantwoorden op een 7-puntslikertschaal. Er werd daarbij vermeld dat 1 stond voor helemaal niet, 4 voor net voldoende en 7 voor volledig. De keuze voor deze aspecten is gebaseerd op de beoordeling van het STEX Schrijven Programma II (Bureau ICE, 2009), het officiële nationale examen dat vaststelt of studenten niveau B2 hebben behaald en dus aan een hbo- of universitaire studie mogen beginnen. Op de aspecten in dit onderzoek worden studenten namelijk ook beoordeeld tijdens STEX Programma II. De aspecten zijn allemaal op een positieve manier geformuleerd, zodat een hoge score bij ieder aspect ook hetzelfde inhield. *Grammatica* werd daarom *grammaticale correctheid*, *spelling* werd *correcte spelling*, *opbouw* werd *logische opbouw* en *woordgebruik* werd *gepast woordgebruik*. De overige aspecten hebben hun oorspronkelijke naam behouden, omdat die al per definitie positief waren. Dit waren de aspecten *adequaatheid*, *begrijpelijkheid* en *samenhang*. De toelichting bij de aspecten zoals hierboven aangegeven werd ook bij iedere tekst getoond, om de aspecten zo duidelijk

mogelijk te maken voor de beoordeling. Er is voor gekozen om *adequaatheid* en *begrijpelijkheid* in dit onderzoek op te splitsen, hoewel dit in het beoordelingsvoorschrift van het STEX als één onderdeel wordt genoemd. Hier is voor gekozen omdat deze concepten feitelijk andere betekenissen hebben. Het kan voor docenten die niet gewend zijn aan dit correctiemodel daarom ook moeilijk zijn om één oordeel te geven over deze aspecten tezamen. Vervolgens werd er bij iedere tekst gevraagd of de betreffende tekst als geheel voldeed aan het niveau dat de docent verwachtte van een eerstejaarsstudent aan hun opleiding. Het antwoord op deze vraag werd gezien als een algemeen oordeel over het niveau van de tekst. Ook deze vraag moest op een zevenpuntsschaal beantwoord worden, waarbij ook hier 1 stond voor helemaal niet, 4 voor net voldoende en 7 voor volledig. De teksten worden dus zowel analytisch als globaal beoordeeld. Bij een analytische beoordeling wordt de kwaliteit van een opstel opgesplitst in een aantal verschillende deelaspecten, zoals hier onder andere spelling, grammatica en opbouw. Bij een globale beoordeling worden producten op grond van een ongedifferentieerde totaalindruk beoordeeld (Meuffels, 1994). Dat gebeurt hier dus door de laatste vraag. Als laatste was er voor de docenten nog een leeg tekstvak, waar ze indien gewenst hun keuzes konden toelichten of andere opmerkingen met betrekking tot de tekst konden achterlaten. Voorbeelden van de teksten van de opdracht ‘samenvatting’ en de opdracht ‘e-mail over onderzoek’ zijn per taalniveau te vinden in bijlages 2 tot en met 7. Van ieder taalniveau en tekstsoort is er één voorbeeld opgenomen. In bijlage 8 is ter illustratie te zien hoe de bijbehorende vragen bij iedere tekst eruit zagen in Qualtrics.

#### **2.4 Design**

De belangrijkste onafhankelijke variabele in deze studie was het veronderstelde taalniveau van de schrijvers. Deze was lager dan B2, hoger dan B2 of moedertaalniveau. De afhankelijke variabele is het oordeel dat de docenten gaven op de vraag of de tekst als geheel voldeed aan het niveau dat de docenten verwachtten van een eerstejaarsstudent op hun opleiding (variabele ‘algemene beoordeling’) en op de vragen in hoeverre ze de eerder genoemde aspecten aanwezig in of kenmerkend voor de betreffende tekst vonden (variabele ‘aspect X’).

#### **2.5 Data-analyse**

De data zijn geanalyseerd met SPSS, versie 25 (IBM Corp, 2017). Om de verschillen tussen de beoordelingen van de verschillende taalniveaus te analyseren, is er gebruik gemaakt van een One-Way ANOVA. Met behulp van Pearsons correlatiecoëfficiënt en multiple regressie is er gekeken in hoeverre de beoordeling van de afzonderlijke aspecten met het algemene oordeel in

verband staan. Als laatste is er ook nog gebruik gemaakt van Pearsons correlatiecoëfficiënt om de relatie tussen de beoordelingen van teksten van dezelfde schrijver te analyseren.

### 3. Resultaten

In deze sectie zullen de resultaten van deze studie besproken worden. Eerst zullen de resultaten van de tekst ‘samenvatting’ beschreven worden en daarna de resultaten van de tekst ‘e-mail over onderzoek’. Daarna is er nog gekeken of de beoordeling van de verschillende taken van dezelfde spreker met elkaar correleerden. Resultaten in deze studie werden significant bevonden bij een  $p$ -waarde kleiner dan .05, tweezijdig getoetst.

#### 3.1 Opdracht ‘samenvatting’

**3.1.1 Verskil tussen groepen.** De resultaten werden vergeleken door eerst per tekst de gemiddelde toegekende score en de standaarddeviatie te berekenen. Dit gemiddelde werd berekend voor alle eerder genoemde aspecten *adequaatheid*, *begrijpelijkheid*, *grammaticale correctheid*, *correcte spelling*, *samenhang*, *logische opbouw* en *gepast woordgebruik* én voor het algemene oordeel over de tekst, het antwoord op de vraag *Voldoet deze tekst als geheel aan het niveau dat u verwacht van een eerstejaarsstudent aan uw opleiding?*. De  $n$  per taalniveau was het aantal samenvattingen per taalniveau. De resultaten zijn te vinden in Tabel 3. Er is te zien dat de docenten alle aspecten van de teksten van de B2- groep (ruim) voldoende hebben beoordeeld en het algemene oordeel als net voldoende. De aspecten *adequaatheid*, *begrijpelijkheid* en *gepast woordgebruik* van de teksten van de B2+ groep zijn voldoende beoordeeld, maar de aspecten *grammaticale correctheid*, *correcte spelling*, *samenhang* en *logische opbouw* zijn als (net) onvoldoende beoordeeld. Het algemene oordeel over de B2+ teksten vonden de deelnemende docenten ook onvoldoende. Alle aspecten en het algemene oordeel over de NT1-teksten vonden de docenten (ruim) voldoende.

Voor alle aspecten en het algemeen oordeel werden de verschillende taalniveaus vergeleken met behulp van contrasten. Er werd namelijk verwacht dat de moedertaalsprekers (NT1-taalniveau) het beter zouden doen dan de studenten die Nederlands als tweede taal leerden. Daarnaast werd aangenomen dat de studenten die al een niveau hoger dan B2 in het Nederlands hadden (B2+ taalniveau) het beter zouden doen dan de studenten die nog geen B2-niveau hadden behaald (B2- taalniveau). Bij het nagaan van de assumpties, werd de assumptie van gelijke varianties bij geen enkel aspect geschonden. Dit is getoetst met Levenes test in SPSS (IBM Corp, 2017). Bij het rapporteren van de resultaten is daarom bij ieder aspect uitgegaan van gelijke varianties.

Tabel 3

Gemiddelde voor opdracht samenvatting per aspect per niveaugroep (schaal 1-7)

	Taalniveau			
	B2 <sup>-a</sup>	B2 <sup>+b</sup>	NT1 <sup>c</sup>	NT2 <sup>d,e</sup>
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Adequaatheid	4.68 (1.39)	4.29 (0.87)	5.64 (1.07)	4.35 (0.90)
Begrijpelijkheid	5.23 (0.61)	4.17 (0.90)	5.70 (0.83)	4.33 (0.93)
Grammaticale correctheid	4.83 (0.38)	3.26 (1.07)	5.54 (1.32)	3.51 (1.15)
Correcte spelling	4.97 (0.67)	3.77 (1.21)	6.22 (0.62)	3.96 (1.21)
Samenhang	4.81 (0.43)	3.92 (0.91)	5.52 (0.91)	4.05 (0.90)
Logische opbouw	4.49 (0.41)	3.86 (0.84)	5.36 (0.65)	3.96 (0.81)
Gepast woordgebruik	4.61 (0.71)	4.26 (0.66)	5.69 (0.60)	4.31 (0.65)
Algemeen oordeel	4.09 (0.97)	3.12 (0.78)	5.50 (0.99)	3.27 (0.85)

<sup>a</sup>  $n = 2$  voor B2<sup>-</sup>. <sup>b</sup>  $n = 11$  voor B2<sup>+</sup>. <sup>c</sup>  $n = 5$  voor NT1. <sup>d</sup> NT2 zijn de deelnemers van B2<sup>-</sup> en B2<sup>+</sup> tezamen. <sup>e</sup>  $n = 13$  voor NT2.

A priori testen lieten zien dat de NT1-groep ( $M = 5.64$ ,  $SD = 1.07$ ) in vergelijking met de NT2-groep ( $M = 4.35$ ,  $SD = 0.90$ ) hoger scoorde op *adequaatheid*. Het verschil was echter niet significant ( $t(15) = 2.02$ ,  $p = .061$ ). De NT2-groep bestond uit de studenten uit de B2<sup>+</sup> en de B2<sup>-</sup> groep tezamen. De B2<sup>+</sup> groep ( $M = 4.29$ ,  $SD = 0.87$ ) scoorde, tegen verwachting, niet hoger op *adequaatheid* dan de B2<sup>-</sup> groep ( $M = 4.68$ ,  $SD = 1.39$ ). De B2<sup>-</sup> groep scoorde juist hoger dan de B2<sup>+</sup> groep, maar het verschil tussen deze twee groepen was niet significant ( $t(15) = -0.53$ ,  $p = .602$ ).

Voor het aspect *begrijpelijkheid* lieten a priori testen zien dat de NT1-groep ( $M = 5.70$ ,  $SD = 0.83$ ) in vergelijking met de NT2-groep ( $M = 4.33$ ,  $SD = 0.93$ ) hoger scoorde, maar dat het verschil ook bij dit aspect niet significant was ( $t(15) = 1.96$ ,  $p = .068$ ). De B2<sup>+</sup> groep ( $M = 4.17$ ,  $SD = 0.90$ ) scoorde ook niet hoger op *begrijpelijkheid* dan de B2<sup>-</sup> groep ( $M = 5.23$ ,  $SD = 0.61$ ). De B2<sup>-</sup> groep scoorde ook hier onverwacht hoger dan de B2<sup>+</sup> groep, maar het verschil tussen deze twee groepen was ook bij dit aspect niet significant ( $t(15) = -1.60$ ,  $p = .130$ ).

A priori testen lieten zien dat, zoals verwacht werd, de NT1-groep ( $M = 5.54$ ,  $SD = 1.32$ ) in vergelijking met de NT2-groep ( $M = 3.51$ ,  $SD = 1.15$ ) hoger scoorde op *grammaticale correctheid*. Dit verschil bleek ook significant te zijn ( $t(15) = 2.28$ ,  $p = .038$ ). De B2<sup>+</sup> groep ( $M = 3.26$ ,  $SD = 1.07$ ) scoorde echter niet hoger op *grammaticale correctheid* dan de B2<sup>-</sup> groep ( $M = 4.83$ ,  $SD = 0.38$ ). De B2<sup>-</sup> groep behaalde hier weer een hogere score dan de

B2+ groep. Het verschil tussen deze groepen was echter ook niet significant ( $t(15) = -1.82, p = .088$ ).

Voor het aspect *correcte* spelling lieten a priori testen zien dat, naar verwachting, de NT1-groep ( $M = 6.22, SD = 0.62$ ) in vergelijking met de NT2-groep ( $M = 3.96, SD = 1.21$ ) een significant hogere score had. Dit verschil was ook significant ( $t(15) = 2.98, p = .009$ ). De B2+ groep ( $M = 3.77, SD = 1.21$ ) scoorde echter wederom niet hoger op *correcte spelling* dan de B2- groep ( $M = 4.97, SD = 0.67$ ). Ook hier was van het tegenovergestelde sprake: de B2- groep scoorde beter dan de B2+ groep. Het verschil tussen deze groepen was echter ook bij dit aspect niet significant ( $t(15) = -1.49, p = .157$ ).

A priori testen lieten zien dat, zoals verwacht werd, de NT1-groep ( $M = 5.52, SD = 0.91$ ) in vergelijking met de NT2-groep ( $M = 4.05, SD = 0.90$ ) een significant hogere score op *samenhang* kreeg. Dit verschil was ook significant. ( $t(15) = 2.22, p = .042$ ). De B2+ groep ( $M = 3.92, SD = 0.91$ ) scoorde echter niet hoger op *samenhang* dan de B2- groep ( $M = 4.81, SD = 0.43$ ). De B2- groep scoorde eveneens bij dit aspect juist hoger dan de B2+ groep. Het verschil tussen deze groepen was echter wederom niet significant ( $t(15) = -1.31, p = .211$ ).

Voor het aspect *logische opbouw* lieten a priori testen zien dat, zoals verwacht werd, de NT1-groep ( $M = 5.36, SD = 0.65$ ) in vergelijking met de NT2-groep ( $M = 3.96, SD = 0.81$ ) hoger scoorde. Het verschil tussen deze twee groepen was ook significant ( $t(15) = 2.62, p = .020$ ). De B2+ groep ( $M = 3.86, SD = 0.84$ ) scoorde niet hoger op *logische opbouw* dan de B2- groep ( $M = 4.49, SD = 0.41$ ). De B2- groep scoorde ook hier weer hoger dan de B2+ groep. Het verschil tussen deze groepen was niet significant ( $t(15) = -1.06, p = .307$ ).

A priori testen lieten zien dat, naar verwachting, de NT1-groep ( $M = 5.69, SD = 0.60$ ) in vergelijking met de NT2-groep ( $M = 4.31, SD = 0.65$ ) een hogere score behaalde op *gepast woordgebruik*. Dit verschil was ook significant ( $t(15) = 3.28, p = .005$ ). De B2+ groep ( $M = 4.26, SD = 0.66$ ) scoorde tegen verwachting wederom niet hoger op *gepast woordgebruik* dan de B2- groep ( $M = 4.61, SD = 0.71$ ). Ook bij dit aspect scoorde de B2- groep hoger dan de B2+ groep. Het verschil tussen deze groepen was niet significant ( $t(15) = -0.70, p = .498$ ).

Ten slotte lieten a priori testen lieten zien dat, zoals verwacht werd, de NT1-groep ( $M = 5.50, SD = 0.99$ ) in vergelijking met de NT2-groep ( $M = 3.27, SD = 0.85$ ) hoger scoorde op het algemene oordeel. Dit verschil was ook hier significant ( $t(15) = 3.76, p = .002$ ). De B2+ groep ( $M = 3.12, SD = 0.78$ ) scoorde echter niet hoger op het algemene oordeel dan de B2- groep ( $M = 4.09, SD = 0.97$ ). Ook bij het algemeen oordeel kregen de teksten van de B2- groep een betere score dan de teksten van de B2+ groep. Het verschil tussen deze groepen was echter ook voor het algemene oordeel niet significant ( $t(15) = -1.48, p = .159$ ).

Samenvattend zijn er, behalve bij adequaatheid en begrijpelijkheid, significante verschillen gevonden tussen de scores van de moedertaalsprekers en tweedetaalleerders, in het voordeel van de eerstgenoemde groep. Er zijn echter geen significante verschillen gevonden binnen de groep tweedetaalleerders. De B2- groep behaalde onverwacht betere scores dan de B2+ groep, maar deze verschillen bleken bij alle aspecten niet significant.

**3.1.2 Verband tussen aspecten en algemeen oordeel.** Om te kijken in hoeverre de scores van de verschillende aspecten met de score van het algemene oordeel in verband stonden, is er gekeken wat de correlatie was tussen deze scores. Dit is gedaan met behulp van Pearsons  $r$ , waarbij  $n = 18$ , het aantal samenvattingen. Uit deze analyse kwam dat alle afzonderlijke aspecten in hoge mate en significant positief correleerden met het algemene oordeel: *adequaatheid*  $r = .85$ , *begrijpelijkheid*  $r = .93$ , *grammaticale correctheid*  $r = .91$ , *correcte spelling*  $r = .89$ , *samenhang*  $r = .94$ , *logische opbouw*  $r = .92$ , *gepast woordgebruik*  $r = .96$ . Voor alle correlaties gold  $p < .001$ .

Met regressieanalyse is gekeken in welke mate de aspecten gezamenlijk samenhangen met het algemene oordeel over de samenvattingen. De zeven aspecten werden in een meervoudige regressie tegelijk ingevoerd als onafhankelijke variabelen, waarbij het algemene oordeel over de teksten de afhankelijke variabele was. De zeven aspecten *adequaatheid*, *begrijpelijkheid*, *grammaticale correctheid*, *correcte spelling*, *samenhang*, *logische opbouw* en *gepast woordgebruik* voorspelden gezamenlijk een groot deel van de variantie in het algemene oordeel,  $F(7,10) = 57.68$ ,  $R^2 = .98$ ,  $p = .000$ . en  $R^2_{Adjusted} = .96$ . In deze analyse is er ook nog gekeken welke aspecten de grootste inbreng hadden in de gezamenlijke voorspelling van het algemene oordeel over de tekst, dus als alle aspecten worden meegenomen. In Tabel 4 is te zien dat, als alle aspecten tezamen werden meegenomen, het aspect *adequaatheid* individueel een significante voorspeller van het algemene oordeel was. De aspecten hingen dus tezamen heel sterk samen met het algemene oordeel, maar behalve *adequaatheid* was er geen enkel aspect dat alleen het algemeen oordeel op een significante wijze voorspelde.

Tabel 4

*Uitkomst regressieanalyse tekst 'samenvatting', met als afhankelijke variabele het algemeen oordeel over de tekst*

Aspect	<i>b</i> (95% CI)	SE	Standardized $\beta$	<i>t</i>	df	<i>p</i>
Adequaatheid	.36 (0.08, 0.65)	0.13	.30	2.82	10	.018
Begrijpelijkheid	.07 (-0.55, 0.70)	0.28	.06	0.26	10	.800
Grammaticale correctheid	.32 (-0.01, 0.64)	0.15	.35	2.19	10	.053
Correcte spelling	.16 (-0.12, 0.44)	0.13	.18	1.27	10	.232
Samenhang	-.19 (-1.16, 0.79)	0.44	-.16	-0.43	10	.674
Logische opbouw	.20 (-0.58, 0.99)	0.35	.15	0.58	10	.575
Gepast woordgebruik	.32 (-0.42, 1.06)	0.33	.21	0.96	10	.360

### 3.2 Opdracht 'e-mail over onderzoek'

**3.2.1 Verschil tussen groepen.** De resultaten werden ook hier vergeleken door eerst per tekst de gemiddelde toegekende score en de standaarddeviatie te berekenen. Dit gemiddelde werd berekend voor alle aspecten *adequaatheid*, *begrijpelijkheid*, *grammaticale correctheid*, *correcte spelling*, *samenhang*, *logische opbouw* en *gepast woordgebruik* én voor het algemene oordeel over de tekst, het antwoord op de vraag *Voldoet deze tekst als geheel aan het niveau dat u verwacht van een eerstejaarsstudent op uw opleiding?*. De *n* per taalniveau was het aantal e-mails per taalniveau. De resultaten zijn te vinden in Tabel 5. Er is te zien dat de docenten de aspecten *adequaatheid*, *logische opbouw* en *gepast woordgebruik* van de e-mails van de B2-groep als (net) voldoende hebben beoordeeld. De aspecten *begrijpelijkheid*, *grammaticale correctheid*, *correcte spelling* en *samenhang* en het algemene oordeel van de teksten van de B2- groep waren volgens de docenten allemaal onvoldoende. Alle aspecten van de e-mails van de B2+ groep zijn voldoende beoordeeld. Tevens vonden de docenten het algemene oordeel over de tekst voldoende. Ook alle aspecten van en het algemeen oordeel over de e-mails van de NT1-groep werden (ruim) voldoende bevonden.

Tabel 5

Gemiddelde voor opdracht e-mail per aspect per niveaugroep (schaal 1-7)

	Taalniveau			
	B2 <sup>-a</sup>	B2 <sup>+b</sup>	NT1 <sup>c</sup>	NT2 <sup>d,e</sup>
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Adequaatheid	4.48 (0.73)	5.51 (1.04)	6.18 (0.39)	5.25 (1.05)
Begrijpelijkheid	3.80 (0.72)	5.18 (0.62)	6.34 (0.31)	4.83 (0.87)
Grammaticale correctheid	2.67 (0.38)	4.17 (0.98)	5.88 (0.35)	3.80 (1.09)
Correcte spelling	2.65 (1.39)	5.15 (0.87)	6.32 (0.28)	4.53 (1.48)
Samenhang	3.60 (0.53)	4.92 (0.95)	5.81 (0.37)	4.59 (1.03)
Logische opbouw	4.43 (0.59)	5.42 (0.75)	5.89 (0.41)	5.17 (0.82)
Gepast woordgebruik	4.02 (0.92)	5.07 (1.09)	6.06 (0.67)	4.81 (1.12)
Algemeen oordeel	2.99 (1.15)	4.72 (1.12)	5.93 (0.43)	4.29 (1.33)

<sup>a</sup>  $n = 3$  voor B2-. <sup>b</sup>  $n = 9$  voor B2+. <sup>c</sup>  $n = 5$  voor NT1. <sup>d</sup> NT2 zijn de deelnemers van B2- en B2+ tezamen. <sup>e</sup>  $n = 12$  voor NT2.

Voor alle aspecten en het algemeen oordeel werden de verschillende taalniveaus vergeleken met behulp van contrasten. Er werd namelijk verwacht dat de moedertaalsprekers (NT1-taalniveau) het beter zouden doen dan de studenten die Nederlands als tweede taal leerden. Daarnaast werd aangenomen dat de studenten die al een niveau hoger dan B2 in het Nederlands hadden (B2+ taalniveau) het beter zouden doen dan de studenten die nog geen B2-niveau hadden behaald (B2- taalniveau). Bij het checken van de assumpties werd de assumptie van gelijke varianties geschonden bij het aspect *correcte spelling*. Dit is getoetst met Levenes test in SPSS (IBM Corp, 2017). Voor het aspect *correcte spelling* zijn er daarom uitkomsten gerapporteerd die geen gelijke varianties aannemen. Voor de andere aspecten zijn wel gelijke varianties aangenomen bij het rapporteren van de resultaten.

A priori testen lieten zien dat de NT1-groep ( $M = 6.18$ ,  $SD = 0.39$ ), naar verwachting, in vergelijking met de NT2-groep ( $M = 5.25$ ,  $SD = 1.05$ ) hoger scoorde op *adequaatheid*. Het verschil tussen deze groepen was ook significant ( $t(14) = 2.48$ ,  $p = .027$ ). De NT2-groep bestond uit de studenten uit de B2+ en de B2- groep samen. De B2+ groep ( $M = 5.51$ ,  $SD = 1.04$ ) scoorde zoals verwacht hoger op *adequaatheid* dan de B2- groep ( $M = 4.48$ ,  $SD = 0.73$ ). Het verschil tussen deze twee groepen studenten was echter niet significant ( $t(14) = 1.79$ ,  $p = .095$ ).

Voor het aspect *begrijpelijkheid* lieten a priori testen zien dat de NT1-groep ( $M = 6.34$ ,  $SD = 0.31$ ) in vergelijking met de NT2-groep ( $M = 4.83$ ,  $SD = 0.87$ ) hoger scoorde en dat het



verschil ook bij dit aspect significant was ( $t(14) = 5.88, p = .000$ ). De B2+ groep ( $M = 5.18, SD = 0.62$ ) scoorde ook bij dit aspect hoger op *begrijpelijkheid* dan de B2- groep ( $M = 3.80, SD = 0.72$ ). Het verschil tussen deze twee groepen was hier wel significant ( $t(14) = 3.66, p = .003$ ).

A priori testen lieten zien dat, zoals verwacht werd, de NT1-groep ( $M = 5.88, SD = 0.35$ ) in vergelijking met de NT2-groep ( $M = 3.80, SD = 1.09$ ) hoger scoorde op *grammaticale correctheid*. Dit verschil bleek ook significant te zijn ( $t(14) = 5.68, p = .000$ ). De B2+ groep ( $M = 4.17, SD = 0.98$ ) scoorde ook hoger op *grammaticale correctheid* dan de B2- groep ( $M = 2.67, SD = 0.38$ ). Het verschil tussen deze groepen was ook hier significant ( $t(14) = 2.91, p = .011$ ).

Voor het aspect *correcte spelling* lieten a priori testen zien dat, naar verwachting, de NT1-groep ( $M = 6.32, SD = 0.28$ ) in vergelijking met de NT2-groep ( $M = 4.53, SD = 1.48$ ) een significant hogere score had. Dit verschil was ook significant ( $t(2.98) = 5.43, p = .012$ ). De B2+ groep ( $M = 5.15, SD = 0.87$ ) scoorde hoger op *correcte spelling* dan de B2- groep ( $M = 2.65, SD = 1.39$ ). Het verschil tussen deze groepen was echter niet significant ( $t(2.54) = 2.92, p = .075$ ).

A priori testen lieten zien dat, zoals verwacht werd, de NT1-groep ( $M = 5.81, SD = 0.37$ ) in vergelijking met de NT2-groep ( $M = 4.59, SD = 1.03$ ) een significant hogere score op *samenhang* kreeg. Dit verschil was ook significant. ( $t(14) = 3.61, p = .003$ ). De B2+ groep ( $M = 4.92, SD = 0.95$ ) scoorde ook hoger op *samenhang* dan de B2- groep ( $M = 3.60, SD = 0.53$ ). Het verschil tussen deze groepen was wel significant ( $t(14) = 2.56, p = .023$ ).

Voor het aspect *logische opbouw* lieten a priori testen zien dat, zoals verwacht werd, de NT1-groep ( $M = 5.89, SD = 0.41$ ) in vergelijking met de NT2-groep ( $M = 5.17, SD = 0.82$ ) hoger scoorde. Het verschil tussen deze twee groepen was ook significant ( $t(14) = 2.66, p = .019$ ). De B2+ groep ( $M = 5.42, SD = 0.75$ ) scoorde hoger op *logische opbouw* dan de B2- groep ( $M = 4.43, SD = 0.59$ ). Het verschil tussen deze groepen was wederom significant ( $t(14) = 2.27, p = .040$ ).

A priori testen lieten zien dat, naar verwachting, de NT1-groep ( $M = 6.06, SD = 0.67$ ) in vergelijking met de NT2-groep ( $M = 4.81, SD = 1.12$ ) een hogere score behaalde op *gepast woordgebruik*. Dit verschil was ook significant ( $t(14) = 2.82, p = .014$ ). De B2+ groep ( $M = 5.07, SD = 1.09$ ) scoorde hoger op *gepast woordgebruik* dan de B2- groep ( $M = 4.02, SD = 0.92$ ). Het verschil tussen deze groepen was echter niet significant ( $t(14) = 1.64, p = .123$ ).

Ten slotte lieten a priori testen zien dat de NT1-groep ( $M = 5.93, SD = 0.43$ ) in vergelijking met de NT2-groep ( $M = 4.29, SD = 1.33$ ) hoger scoorde op het algemene oordeel. Dit verschil was wederom significant ( $t(14) = 3.79, p = .002$ ). De B2+ groep ( $M = 4.72, SD =$

1.12) scoorde ook hoger op het algemeen oordeel dan de B2- groep ( $M = 2.99$ ,  $SD = 1.15$ ) en het verschil tussen deze groepen was ook significant ( $t(14) = 2.65$ ,  $p = .019$ ).

Samenvattend zijn er op alle aspecten ook bij deze tekstsoort significante verschillen gevonden tussen de scores van de moedertaalsprekers en tweedetaalleerders, in het voordeel van de eerstgenoemde groep. Bij deze tekstsoort zijn er ook significante verschillen gevonden binnen de tweedetaalleerders, in tegenstelling tot bij de samenvatting. De B2+ groep behaalde significant betere scores dan de B2- groep, behalve bij de aspecten adequaatheid, correcte spelling en gepast woordgebruik. Bij deze aspecten was er geen significant verschil tussen de scores van de B2+ en B2- groep.

**3.2.2 Verband tussen aspecten en algemeen oordeel.** Om te kijken in hoeverre de scores van de verschillende aspecten met de score van het algemene oordeel in verband stonden, is er gekeken wat de correlatie was tussen deze scores. Dit is gedaan met behulp van Pearsons  $r$ , waarbij  $n = 17$ , het aantal e-mails over onderzoek. Uit deze analyse kwam dat alle afzonderlijke aspecten in hoge mate en significant positief correleerden met het algemene oordeel: *adequaatheid*  $r = .87$ , *begrijpelijkheid*  $r = .93$ , *grammaticale correctheid*  $r = .89$ , *correcte spelling*  $r = .84$ , *samenhang*  $r = .96$ , *logische opbouw*  $r = .88$ , *gepast woordgebruik*  $r = .88$ . Voor alle correlaties gold  $p < .001$ .

Met regressieanalyse is gekeken in welke mate de aspecten samen invloed hebben op het algemene oordeel over de e-mails over onderzoek. De zeven aspecten zijn in een meervoudige regressie tegelijk ingevoerd als onafhankelijke variabelen, waarbij het algemene oordeel over de teksten de afhankelijke variabele was. De zeven aspecten *adequaatheid*, *begrijpelijkheid*, *grammaticale correctheid*, *correcte spelling*, *samenhang*, *logische opbouw* en *gepast woordgebruik* voorspelden gezamenlijk een groot deel van de variantie in het algemene oordeel,  $F(7,9) = 23.85$ ,  $R^2 = .95$   $p = .000$ . en  $R^2_{Adjusted} = .91$ . In deze analyse is er ook nog gekeken welke aspecten de grootste inbreng hadden in de gezamenlijke voorspelling van het algemene oordeel over de tekst, dus als alle aspecten worden meegenomen. In Tabel 6 is te zien dat, als alle aspecten werden meegenomen, geen enkele van de afzonderlijke aspecten een significante voorspeller van het algemene oordeel was; ieder aspect voorspelde het algemene oordeel een beetje. De aspecten hingen dus tezamen wel heel sterk samen met het algemene oordeel, maar er was geen enkel aspect dat alleen het algemeen oordeel op een significante wijze voorspelde.

Tabel 6

Uitkomst regressie-analyse tekst 'e-mail', met als afhankelijke variabele het algemeen oordeel over de tekst

Aspect	<i>b</i> (95% CI)	SE	Standardized $\beta$	<i>t</i>	df	<i>p</i>
Adequaatheid	.10 (-0.46, 0.66)	0.25	.07	0.41	9	.692
Begrijpelijkheid	.26 (-0.88, 1.40)	0.50	.19	0.51	9	.624
Grammaticale correctheid	-.05 (-0.58, 0.48)	0.23	-.05	-0.22	9	.832
Correcte spelling	.03 (-0.48, 0.54)	0.22	.03	0.14	9	.894
Samenhang	.88 (-0.16, 1.92)	0.46	.68	1.91	9	.087
Logische opbouw	-.20 (-1.07, 0.67)	0.39	-.12	-0.52	9	.617
Gepast woordgebruik	.23 (-0.34, 0.81)	0.26	.20	0.92	9	.384

### 3.3 Verband tussen beoordeling taken

Om te achterhalen of de schrijfvaardigheid van de schrijvers hetzelfde was op de twee taken, is er gekeken of er een verband was tussen de beoordelingen op de verschillende aspecten en het algemene oordeel op de twee verschillende taken. Dit is ook gedaan met Pearsons  $r$ , waarbij  $n = 15$ . Er waren namelijk 15 schrijvers die twee teksten aan hadden geleverd voor dit onderzoek. Er werd verwacht dat de correlaties positief, significant en hoog waren. Geen enkele correlatie voldeed hier echter aan. De correlatie tussen de beoordelingen op *adequaatheid* was  $r = .17$ ,  $p = .545$ , op *begrijpelijkheid*  $r = .37$ ,  $p = .175$ , op *grammaticale correctheid*  $r = .49$ ,  $p = .067$ , op *correcte spelling*,  $r = .48$ ,  $p = .070$ , op *samenhang*,  $r = .36$ ,  $p = .18$ , op *logische opbouw*  $r = .36$ ,  $p = .188$ , op *gepast woordgebruik*  $r = .48$ ,  $p = .073$ , en op het algemene oordeel  $r = .50$ ,  $p = .057$ .

### 3.4 Tekstvakken

Het was voor de docenten ook mogelijk om opmerkingen achter te laten bij het beoordelen van de teksten. De docenten hebben deze functie vooral gebruikt om hun keuze toe te lichten. Zo werden er opmerkingen als 'Op zich correct maar veel te lang', 'Niet voldaan aan de opdracht om er een goedlopend verhaal van te maken' en 'Eenvoudige zinnen schrijven gaat goed, wat complexere levert grote problemen op' achtergelaten. Daarnaast gaf een docent aan dat aan het

aspect ‘logische opbouw’ bij een samenvatting snel voldaan kan worden, omdat de student alleen maar de tekst hoeft te volgen. Ook gaven twee docenten bij de samenvatting aan dat de opdracht heel lang was en daarbij noemde één van deze twee docenten dat daardoor mogelijk vermoeidheid optrad bij het beoordelen de laatste teksten.

## 4. Discussie

In deze sectie zullen eerst de resultaten van de verschillende tekstsoorten en het verband tussen de beoordeling van deze twee tekstsoorten besproken worden. Daarna volgen de tekortkomingen van dit onderzoek, aanbevelingen voor de praktijk en aanbevelingen voor vervolgonderzoek. De opmerkingen in de tekstvakken zullen in deze discussie niet apart besproken worden, maar wel meegenomen worden in het bespreken van de andere onderdelen van de discussie.

### 4.1 Samenvatting

De verwachting was dat de moedertaalsprekers beter beoordeeld zouden worden dan de tweedetaalleerders en dat de studenten uit de B2+ groep beter beoordeeld zouden worden dan de B2- groep. Wat betreft de absolute beoordeling werden de teksten van moedertaalsprekers allemaal ruim voldoende beoordeeld. Bij de teksten van de B2+ groep werden de meeste aspecten echter onvoldoende beoordeeld. Alleen de aspecten *adequaatheid*, *begrijpelijkheid* en *gepast woordgebruik* werden voldoende beoordeeld. Ook het algemene oordeel vonden de docenten onvoldoende. Dit zou aangeven dat de docenten, algemeen gezien, het taalniveau van de studenten op basis van de beoordeling van deze tekst nog niet hoog genoeg vinden om aan een universitaire opleiding te beginnen. De aspecten van de teksten van de B2- groep zijn echter als voldoende beoordeeld, evenals het algemene oordeel over deze teksten. Het is opmerkelijk te noemen dat de teksten van de B2- groep beter beoordeeld werden dan de teksten van de B2+ groep. Bij het vinden van een verklaring voor dit gegeven is het belangrijk om te vermelden dat er bij de opdracht ‘samenvatting’ maar twee teksten zijn opgenomen van B2- taalniveau, in tegenstelling tot 11 teksten van B2+ taalniveau. Wat betreft het B2- taalniveau is dit veel te weinig en geeft dit mogelijk dus een vertekend beeld. Het is mogelijk dat alleen gemotiveerde studenten van B2- taalniveau een tekst hebben aangeleverd. Toch is ervoor gekozen om deze teksten op te nemen, om te kijken hoe de docenten deze teksten zouden beoordelen.

Statistisch gezien zijn de teksten van moedertaalsprekers op alle aspecten significant beter beoordeeld dan de teksten van tweedetaalleerders, behalve op de aspecten *adequaatheid* en *begrijpelijkheid*. *Adequaatheid* is dan ook een aspect waar verschillende taalniveaus wellicht

makkelijk goed op kunnen scoren; het gaat vooral om het voldoen aan de opdracht. Zo werd *adequaatheid* in deze opdracht immers beschreven. *Begrijpelijkheid* is ook een aspect waar wellicht makkelijk aan voldaan wordt, omdat een tekst nog steeds begrepen kan worden, ondanks dat aan andere aspecten niet voldaan wordt. De begrijpelijkheid van de B2- teksten was hier ook hoog beoordeeld, wat er ook voor zorgt dat er geen significant verschil bestond tussen de NT1-ers en NT2-ers. Daarnaast kan het zo zijn dat de docenten de samenvattingen al snel begrepen, omdat ze achttien samenvattingen over dezelfde tekst lazen. Dat de moedertaalsprekers het over het algemeen beter doen dan de tweedetaalleerders, is zoals verwacht werd. Er was echter geen significant verschil tussen de beoordelingen van de teksten van de twee groepen tweedetaalleerders. Hoewel verwacht werd dat de B2+ groep beter beoordeeld zou worden dan de B2- groep, is dit dus niet het geval. Dit kan te maken hebben met de dimensies die Hulstijn (2007) beschreef. Hij stelde dat er twee dimensies zijn, namelijk de kwantitatieve dimensie en de kwalitatieve dimensie. De kwantitatieve dimensie houdt in 'wat iemand moet kunnen' en de kwalitatieve dimensie 'hoe goed iemand activiteit x moet kunnen'. Volgens deze auteur veronderstelt het ERK dat deze dimensies altijd aan elkaar gelijk zijn. Er kunnen volgens de auteur echter wel degelijk discrepanties bestaan tussen deze twee dimensies. Het kan in dit geval voorkomen dat studenten uit de B2+ groep minder presteren op de kwalitatieve dimensie. De studenten kunnen bijvoorbeeld niet zo goed zijn in het schrijven van een samenvatting. Ook kan dit verschil met de verwachting komen, zoals eerder genoemd, door het geringe aantal teksten van de B2- groep of door het relatief hoge niveau van het kleine aantal teksten van de B2- groep. Concluderend worden de teksten van moedertaalsprekers dus als (ruim) voldoende beoordeeld, maar zijn de teksten van de tweedetaalleerders onder het niveau dat docenten verwachten van eerstejaarsstudenten van hun opleiding. Dit is in lijn met onderzoek van Deygers, Van den Branden en Van Gorp (2018) en Field (2011), die stelden dat studenten die B2-niveau hebben behaald, lang niet altijd klaar zijn om naar de universiteit te gaan.

De correlaties tussen de afzonderlijke aspecten en het algemene oordeel waren allemaal hoog, positief en significant. Dat wil zeggen dat er een samenhang is tussen deze aspecten en het algemene oordeel. Dit geldt dus voor ieder aspect. Het algemene oordeel hangt dus sterk samen met de beoordelingen van de afzonderlijke aspecten. Dat is volgens verwachting; de verschillende aspecten zijn allemaal onderdelen van een tekst, en zo ook van het algemeen oordeel. Daarnaast beoordeelden de docenten teksten één voor één en niet eerst alle teksten op *adequaatheid*, dan op *begrijpelijkheid* etc. Ze stemmen zo de beoordelingen binnen teksten wellicht op elkaar af. Uit de regressieanalyse kwam ook dat alle aspecten samen het algemene

oordeel bijna geheel voorspellen. Verder bleek dat alleen het aspect *adequaatheid* afzonderlijk een significante voorspeller van het algemene oordeel is, als de andere aspecten worden meegenomen in de analyse. Als alle aspecten worden meegenomen, kunnen de aspecten afzonderlijk echter wel gemist worden om het algemeen oordeel te voorspellen, met uitzondering van het aspect *adequaatheid*. De andere aspecten kunnen dan namelijk nog steeds gezamenlijk het algemene oordeel voorspellen. Samenvattend hebben de hebben de aspecten afzonderlijk dus een significante correlatie met het algemeen oordeel en hangen ze gezamenlijk ook sterk samen met het algemeen oordeel.

#### **4.2 E-mail**

De verwachting was ook hier dat de moedertaalsprekers beter beoordeeld zouden worden dan de tweedetaalleerders en dat de studenten uit de B2+ groep beter beoordeeld zouden worden dan de B2- groep. Wat betreft de absolute beoordeling werden de teksten van moedertaalsprekers allemaal ruim voldoende beoordeeld. Alle aspecten én het algemene oordeel van de teksten van de B2+ groep werden voldoende beoordeeld. Bij de B2- groep werden de aspecten *adequaatheid*, *logische opbouw* en *gepast woordgebruik* voldoende beoordeeld; de overige aspecten en het algemene oordeel werden onvoldoende beoordeeld.

Statistisch gezien was het niveau van moedertaalsprekers bij alle aspecten en het algemene oordeel beter dan het niveau van tweedetaalleerders. De aspecten *begrijpelijkheid*, *grammaticale correctheid*, *samenhang*, *logische opbouw* en het algemene oordeel werden voor de B2+ teksten beter beoordeeld dan voor de B2- teksten. Deze statistische uitkomsten kloppen dus ook met de hypothesen die er waren: de groep moedertaalsprekers deed het beter dan de tweedetaalleerders en de B2+ groep deed het in het algemeen beter dan de B2- groep. De teksten van moedertaalsprekers en B2+ leerders werden (ruim) voldoende beoordeeld en de teksten van B2- leerders als onvoldoende. Op basis van deze uitkomsten kan geconcludeerd worden dat studenten met B2+ niveau inderdaad een taalvaardigheid bezitten die nodig is in het eerste jaar van de universiteit en studenten van B2- niveau nog niet. Dit is in lijn met eerdere onderzoeken van o.a. Bechger e.a., (2009), Carlsen (2018) en Harsch e.a. (2017), die ook stelden dat B2 een voldoende niveau is om naar de universiteit te gaan.

Ook bij de e-mails waren alle correlaties tussen de afzonderlijke aspecten en het algemene oordeel hoog, positief en significant. Het algemene oordeel hangt dus ook hier sterk samen met de beoordelingen van de afzonderlijke aspecten. Uit de regressieanalyse bleek ook hier dat de verschillende aspecten tezamen het algemene oordeel bijna geheel voorspellen. Geen enkel aspect was echter afzonderlijk een significante voorspeller van het algemene oordeel, als alle aspecten werden meegenomen in de analyse. Als alle aspecten worden meegenomen,

kunnen individuele aspecten echter als het ware ‘gemist’ worden; dan vangen de andere aspecten de voorspellende waarde op. Samenvattend hebben de aspecten afzonderlijk dus een significante correlatie met het algemeen oordeel en hangen ze gezamenlijk ook sterk samen met het algemeen oordeel.

### **4.3 Verband tussen beoordeling taken**

De verwachting was dat er een positieve, hoge en significante correlatie gevonden zou worden tussen de beoordeling op de twee taken. Er was echter geen enkel significant verband tussen de beoordelingen van de verschillende aspecten en het algemene oordeel op de twee verschillende taken ‘samenvatting’ en ‘e-mail’. Dat kan betekenen dat de schrijvers geen constante schrijfvaardigheid hebben (laten zien) en/of dat de beoordelaars niet constant hebben beoordeeld. Wat het in ieder geval laat zien, is dat het laten uitvoeren van louter één schrijftaak niet genoeg is om uitspraken te doen over het niveau dat de schrijver van deze tekst heeft. In dit onderzoek worden de twee verschillende tekstsoorten immers verschillend beoordeeld. Voor een eerlijk oordeel over het taalniveau van een bepaalde schrijver, zouden er dus meer teksten door eenzelfde persoon moeten worden gemaakt.

### **4.4 Tekortkomingen**

In dit onderzoek is er gekozen voor een populatie van deelnemende studenten én docenten die heterogeen was. Zo was de achtergrond van de studenten wat betreft moedertaal, land van herkomst en aantal jaar in Nederland verschillend. De moedertaal kan bijvoorbeeld invloed hebben op bepaalde fouten die er worden gemaakt bij het schrijven van een tekst. De docenten kwamen van verschillende opleidingen van de Radboud Universiteit. Deze factoren zorgden voor veel variantie in de data en de resultaten, waardoor de data moeilijk te generaliseren zijn. De vraagstelling bij de beoordeling van de teksten was ook gebaseerd op de eigen opleiding. Er werd namelijk aan de docenten gevraagd *‘Voldoet deze tekst als geheel aan het niveau dat u verwacht van een eerstejaarsstudent op uw opleiding?’*. Het niveau dat docenten verwachten op hun opleiding van een eerstejaars, kan verschillend zijn per opleiding. Een te klein aantal docenten per faculteit zorgde er ook voor dat er niet gedifferentieerd kan worden naar resultaten van één bepaalde faculteit. Dit maakt het ook moeilijk om harde uitspraken te doen over het niveau dat verwacht wordt van tweedetaalleerders van het Nederlands. Overigens is het wel zo dat het STEX ook niet differentieert naar verschillende opleidingen.

Het aantal teksten van studenten die nog geen B2-niveau hebben behaald (B2-), was te klein. Bij de samenvattingen waren er maar twee teksten van deze studenten opgenomen, bij de e-mails drie. Er is aan een groep van vijftien studenten uit de A2-B1 gevraagd om mee te doen,

maar uiteindelijk hebben er dus maar drie studenten teksten aangeleverd voor dit onderzoek. De studenten die teksten hebben aangeleverd, zijn waarschijnlijk ook studenten die minder moeite hebben met het schrijven van teksten en gemotiveerd zijn om extra te oefenen. Als dat inderdaad zo is, geeft dat een vertekend beeld van het niveau van B2- studenten. Dat betekent dus ook dat vrijwillige deelname nadelig kan zijn voor de uitkomsten van dit (en toekomstig) onderzoek.

De tekstsoorten die gekozen zijn in dit onderzoek, waren wellicht niet zo geschikt als vooraf gedacht. Zo is een samenvatting op sommige punten wellicht makkelijker om te maken, omdat bijvoorbeeld de opbouw en samenhang van een tekst kan worden overgenomen van de oorspronkelijke tekst. Bij de e-mail was er meer ruimte om authentiek taalgebruik te laten zien, omdat er ook verlangd werd om een eigen mening over het gelezen onderzoek te delen. Dit zorgt wellicht beter voor een indruk van de schrijfvaardigheid van een tweedetaalleerder.

In dit onderzoek werd er alleen onderzoek gedaan naar schrijfvaardigheid. Deygers, Van den Branden en Van Gorp (2018) en Field (2011) gaven aan dat studenten in België en Engeland in een universitaire context vooral moeite hadden met het luisteren naar een tweede taal. Het is echter moeilijk om generaliserende uitspraken over het taalniveau van studenten te doen op basis van één taalvaardigheid, zoals in het huidige onderzoek, omdat studenten verschillend kunnen scoren op verschillende taalvaardigheden. Het is dan ook lastig om conclusies te verbinden aan dit onderzoek over schrijfvaardigheid met betrekking tot het algemene taalniveau van de deelnemende studenten.

#### **4.5 Aanbevelingen voor de praktijk**

De resultaten uit dit onderzoek laten zien dat B2-niveau niet per definitie het beste of meest geschikte niveau is om studenten toe te laten aan de universiteit. De uitkomsten zijn namelijk verschillend; bij de e-mail werd het schrijfniveau wel goed genoeg bevonden door docenten, bij de samenvatting niet. Voor universiteiten in Nederland is het daarom zaak om goed na te gaan waarom zij B2-niveau willen als toegangseis voor studenten die Nederlands als tweede taal leren. Toelating gebeurt nu vooral via het STEX II. Misschien zouden universiteiten echter niet te veel moeten uitgaan van het ERK en het STEX om studenten toe te laten tot de universiteiten. Zoals Figueras (2012), Harsch (2018) en Wisniewski (2017) aangeven, is het ERK namelijk ook geen model van absolute perfectie en wordt het (te) vaak gebruikt voor doeleinden die van te voren niet bedoeld of voorzien waren. Harsch (2018) geeft hierbij ook nog aan dat het goed zou zijn om een behoefteanalyse te doen voor iedere test of cursus die gegeven wordt. Een voldoende taalniveau is hoe dan ook heel belangrijk voor (aspirant)-studenten. Het ERK kan als hulpmiddel dienen om dit taalniveau te controleren, maar het moet



niet als doel op zich gezien worden. Het is vooral zaak dat dat studenten een voldoende taalniveau halen om mee te kunnen op de universiteit en het is niet zo dat ze daarvoor per definitie B2-niveau moeten halen. Een aanbeveling voor universiteiten, faculteiten en opleidingen zou dan ook zijn om zelf na te gaan wat zij vinden dat studenten moeten kunnen en welke eisen zij aan opdrachten die studenten krijgen stellen. Moeten ze bijvoorbeeld opstellen kunnen schrijven of presentaties kunnen houden? En hoeveel fouten zijn er in een tekst toegestaan? Wat is er belangrijk in/aan een tekst? Het is waarschijnlijk een tijdrovende en dure klus, maar het zorgt er wel voor dat studenten met een andere taalachtergrond dan het Nederlands beter voorbereid worden op de opleiding die ze willen volgen en dat docenten minder voor verrassingen komen te staan wat betreft het taalniveau van studenten.

#### **4.6 Aanbevelingen voor vervolgonderzoek**

In vervolgonderzoek is het mogelijk om de populatie deelnemers specifieker te maken; neem bijvoorbeeld alleen teksten van studenten met een bepaalde moedertaal op in de studie. Dat kan zorgen voor een betere interne validiteit. In lijn met deze aanpassing is het dan ook goed om de docentenpopulatie homogener te maken; zo kunnen er duidelijke patronen ontdekt worden in de beoordelingen van docenten van bepaalde faculteiten. Aan de andere kant zorgt het homogener maken van de populatie docenten en studenten ervoor dat de uitkomsten minder gegeneraliseerd kunnen worden. Het is in ieder geval belangrijk om de populatie van deelnemers, zowel studenten als docenten, in de toekomst uit te breiden. Ook kan er in de toekomst voor gekozen worden om meer universiteiten mee te nemen, om generaliserende uitspraken te doen over de nationale context. De keuzes voor de huidige populatie tekstschrijvers én docenten zijn vooral gemaakt uit praktische overwegingen. Docenten gaven ook nu al aan dat het veel werk was om de teksten te beoordelen. Eén docent beoordeelde waarschijnlijk daardoor ook maar de helft van de teksten. Om meer docenten van één opleiding of één faculteit te vinden, zal in de toekomst dan ook niet makkelijk zijn. Dit kan dus wellicht wel opgelost worden door het onderzoek op nationaal niveau uit te voeren.

Om achter de redenen te komen voor het toekennen van bepaalde scores en om erachter te komen wat docenten verwachten van (eerstejaars)studenten, is het voor vervolgonderzoek ook goed om een kwalitatieve insteek te gebruiken. Een aanpak van zowel kwalitatieve als kwantitatieve data kan meer inzicht geven in waarom bepaalde scores worden toegekend en wat docenten van teksten vinden. Zo kan er bijvoorbeeld achterhaald worden of docenten minder streng zijn voor NT2-studenten wat betreft schrijfvaardigheid, zoals Deygers, Van den Branden en Van Gorp (2018) vonden in hun onderzoek.

Om meer studenten van lagere niveaus op te nemen in soortgelijke onderzoeken is het

een idee om (meer) samenwerking te zoeken met taalscholen. Teksten zouden dan opgegeven kunnen worden als huiswerk voor alle niveaugroepen en (met toestemming) gebruikt kunnen worden voor onderzoek. Naast dat er meer studenten van het niveau B2- opgenomen moeten worden in vervolgstudies, is het ook belangrijk om tekstsoorten of opdrachten op te nemen die voor een brede populatie te schrijven zijn, zodat er zo ook meer studenten van verschillende, en in ieder geval lagere, taalniveaus opgenomen kunnen worden. In toekomstig onderzoek moet er daarom beter nagedacht worden over de moeilijkheidsgraad van een tekstsoort en de taalniveaus van de studenten die de teksten moeten schrijven. Het opnemen van een tekstsoort zoals een betoog zou bijvoorbeeld een goede inkijk geven in het taalniveau van verschillende studenten, maar het kan studenten van een lager niveau afschrikken om mee te doen aan onderzoeken zoals hierboven is uitgevoerd.

Ten slotte richtte dit onderzoek zich alleen op schrijfvaardigheid. Om meer uitspraken te doen over het toelaten van NT2-ers aan de universiteit is het zaak dat in de toekomst ook onderzoek wordt gedaan naar de andere taalvaardigheden lezen, luisteren en spreken.

## **5. Conclusie**

Op de hoofdvraag van dit onderzoek, of B2-niveau een voldoende niveau is om naar de universiteit te gaan in Nederland, is geen eenduidig antwoord te geven. De beoordelingen van de twee teksten laten namelijk verschillende resultaten zien. Op basis van de samenvatting kan gezegd worden dat tweedetaalleerders met B2+ niveau nog niet voldoende niveau hebben om naar de universiteit te gaan. De resultaten van de e-mails laten echter zien dat deze studenten wel voldoende niveau hebben, volgens de docenten die de teksten beoordeelden. Deze uitkomsten laten in ieder geval wel zien dat op basis van de één geschreven tekst geen uitspraak gedaan kan worden over het taalniveau van studenten. Er is enerzijds meer dan één schrijfproduct nodig, en er zijn anderzijds gevarieerde schrijfproducten nodig om het niveau van studenten te bepalen. De resultaten van deze studie laten namelijk zien dat de beoordeling van teksten afhankelijk kan zijn van de taak.

## Referenties

- Bachman, L. F., & Palmer, A. S. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.
- Bechger, T. M., Kuijper, H., & Maris, G. (2009). Standard Setting in Relation to the Common European Framework of Reference for Languages: The Case of the State Examination of Dutch as a Second Language. *Language Assessment Quarterly*, 6(2), 126–150.  
<https://doi.org/10.1080/15434300802457521>
- Bureau ICE. (2009). Beoordelingsmodel schrijven II: 2009 voorbeeldexamen. Geraadpleegd van <https://www.staatsexamensnt2.nl/item/voorbeeldexamens-programma-ii-2009>
- Carlsen, C. H. (2018). The Adequacy of the B2 Level as University Entrance Requirement. *Language Assessment Quarterly*, 15(1), 75–89.  
<https://doi.org/10.1080/15434303.2017.1405962>
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching and Assessment*. Strasbourg: Council of Europe.
- Council of Europe. (2018). *Common European Framework of Reference for Languages: Learning, Teaching Assessment. Companion Volume with New Descriptors*. Strasbourg: Council of Europe.
- Deygers, B., Van den Branden, K., & Van Gorp, K. (2018). University entrance language tests: A matter of justice. *Language Testing*, 35(4), 449–476.  
<https://doi.org/10.1177/0265532217706196>
- Deygers, B., Van Gorp, K., & Demeester, T. (2018). The B2 Level and the Dream of a Common Standard. *Language Assessment Quarterly*, 15(1), 44–58.  
<https://doi.org/10.1080/15434303.2017.1421955>

- Deygers, B., Zeidler, B., Vilcu, D., & Carlsen, C. H. (2018). One Framework to Unite Them All? Use of the CEFR in European University Entrance Policies. *Language Assessment Quarterly*, 15(1), 3–15. <https://doi.org/10.1080/15434303.2016.1261350>
- Dienst Uitvoering Onderwijs. (n.d.). Staatsexamen NT2. Geraadpleegd 25 juni 2019, van <https://duo.nl/particulier/staatsexamen-nt2/hoe-het-staatsexamen-nt2-werkt.jsp#>
- Field, J. (2011). Into the mind of the academic listener. *Journal of English for Academic Purposes*, 10(2), 102–112. <https://doi.org/10.1016/j.jeap.2011.04.002>
- Figueras, N. (2012). The impact of the CEFR. *ELT Journal*, 66(4), 477–485. <https://doi.org/10.1093/elt/ccs037>
- Green, A. (2018). Linking Tests of English for Academic Purposes to the CEFR: The Score User's Perspective. *Language Assessment Quarterly*, 15(1), 59–74. <https://doi.org/10.1080/15434303.2017.1350685>
- Harsch, C. (2018). How Suitable Is the CEFR for Setting University Entrance Standards? *Language Assessment Quarterly*, 15(1), 102–108. <https://doi.org/10.1080/15434303.2017.1420793>
- Harsch, C., Ushioda, E., & Ladroue, C. (2017). Investigating the Predictive Validity of *TOEFL iBT*® Test Scores and Their Use in Informing Policy in a United Kingdom University Setting: Investigating the Predictive Validity of *TOEFL iBT*® Test Scores. *ETS Research Report Series*, 2017(1), 1–80. <https://doi.org/10.1002/ets2.12167>
- Hughes, A. (2013). *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- Hulstijn, J. H. (2007). The Shaky Ground Beneath the CEFR: Quantitative and Qualitative Dimensions of Language Proficiency<sup>1</sup>. *The Modern Language Journal*, 91(4), 663–667. [https://doi.org/10.1111/j.1540-4781.2007.00627\\_5.x](https://doi.org/10.1111/j.1540-4781.2007.00627_5.x)

- Hulstijn, J. H. (2011). Language Proficiency in Native and Nonnative Speakers: An Agenda for Research and Suggestions for Second-Language Assessment. *Language Assessment Quarterly*, 8(3), 229–249. <https://doi.org/10.1080/15434303.2011.565844>
- IBM Corp. (2017). *IBM Statistics for Windows, Version 25*. Armonk, NY: IBM Corp.
- Martyniuk, W., & Noijons, J. (2007). *Executive Summary of Results of a Survey on the Use of the CEFR at National Level in the Council of Europe Member States*. Geraadpleegd van [http://www.coe.int/t/dg4/linguistic/Source/Survey\\_CEFR\\_2007\\_EN.doc](http://www.coe.int/t/dg4/linguistic/Source/Survey_CEFR_2007_EN.doc)
- Meuffels, B. (1994). *De verguisde beoordelaar: Opstellen over opstelbeoordeling*. Amsterdam: Thesis Publishers.
- Nederlandse Taalunie. (2008). *Gemeenschappelijk Europees referentiekader voor moderne vreemde talen: Leren, onderwijzen beoordelen*. Den Haag: Nederlandse Taalunie.
- Nederlandse Taalunie. (2018). *Gemeenschappelijk Europees referentiekader voor moderne vreemde talen: Leren, onderwijzen, beoordelen. Supplement met nieuwe descriptoren*. Den Haag: Nederlandse Taalunie.
- Schoonen, R. (2005). Generalizability of writing scores: An application of structural equation modeling. *Language Testing*, 22(1), 1–30. <https://doi.org/10.1191/0265532205lt295oa>
- Wang, H., Choi, I., Schmidgall, J., & Bachman, L. F. (2012). Review of Pearson Test of English Academic: Building an assessment use argument. *Language Testing*, 29(4), 603–619. <https://doi.org/10.1177/0265532212448619>
- Wisniewski, K. (2017). Empirical Learner Language and the Levels of the *Common European Framework of Reference: The CEFR Levels and Learner Language*. *Language Learning*, 67(S1), 232–253. <https://doi.org/10.1111/lang.12223>

## **Bijlagen**

### **Bijlage 1: Originale tekst over hoogbegaafdheid**

#### **Beter omgaan met hoogbegaafde leerlingen**

*Interview met Jan Kuipers*

“Vroeger was de zorg vooral op de zwakke leerlingen gericht. Hoogbegaafde leerlingen konden zichzelf wel redden, was de overheersende opinie. Het aangeboren talent werd als een luxe gezien, waarin je geen extra tijd hoefde te steken. Bovendien was het een typisch Nederlands cultuurverschijnsel dat je je hoofd niet te ver boven het maaiveld uit mocht steken. De laatste tien jaar is onder druk van de ouders het denken over deze groep kinderen radicaal veranderd.”

Aan het woord is onderwijsadviseur Jan Kuipers (Cedin). Hij begeleidt directies en leerkrachten in de juiste aanpak van hoogbegaafde leerlingen. “Want inmiddels staat wel vast dat ook hoogbegaafde leerlingen behoorlijk wat extra aandacht en vooral een eigen benadering behoeven”, weet Kuipers uit ervaring. “Als ze op een verkeerde manier benaderd worden, kunnen ze zich erg ongelukkig en geïsoleerd voelen.” Verveling, perfectionisme (inclusief faalangst), onzekerheid en hyperactiviteit zijn veelgehoorde problemen. Uit onderzoek blijkt ook dat ze dikwijls te maken hebben met psychosociale problemen, zoals gepest worden op school, eenzaamheid, overgevoeligheid en een zwakke sociaal-emotionele ontwikkeling.

In onderlinge samenhang kunnen deze factoren gemakkelijk leiden tot onderpresteren. Kuipers: “Onderpresteren lijkt daarmee één van de meest raadselachtige en frustrerende problemen van hoogbegaafde leerlingen. Het lijkt wel of ze twee gezichten hebben. Op sommige momenten zie je de onbegrensde mogelijkheden, een briljante opmerking, een goed doordachte visie, maar daarna is het voorbij en wordt er een muur van apathie en onverschilligheid opgetrokken.”

Niet alleen voor het kind is onderpresteren een grote frustratie, ook voor ouders en leerkrachten is onderpresteren vaak een lastig en moeilijk te hanteren probleem. Hoogbegaafde kinderen hebben vaak twee gezichten. Ze zitten vol passie en zitten als de juiste snaar geraakt wordt op het puntje van hun stoel. Anderzijds kunnen ze ook zeer lusteloos en doelloos gedrag vertonen en zijn ze soms tegendraads en onbenaderbaar, of juist volledig aangepast en daardoor onzichtbaar.

#### **Lat hoger leggen**

Kuipers stelt onomwonden dat in het basisonderwijs onderpresteren het gevolg is van een ontoereikend lesaanbod gecombineerd met een te laag ambitieniveau van de ouders en de leerkrachten. Kuipers: “In de VS en de Aziatische landen ligt de lat veel hoger. Een topprestatie leveren is daar de normaalste zaak van de wereld. In Nederland, maar ook in de rest van Europa, zie je zelden dat een school er prat op gaat dat ze het maximale van leerkracht en leerlingen vragen. Daarin zie je nu langzaam maar zeker een kentering optreden, we krijgen steeds meer aandacht voor talent en prestaties.”

Volgens Kuipers vergt dit echter wel een totaal andere pedagogisch-didactische aanpak van de leerkracht. Hij weet dat enkel en alleen het creëren van een plusklas geen optie is. “Niet één onderwijsaanpassing is zaligmakend. Daar zijn de leerlingen nu eenmaal te verschillend voor. Maar inmiddels is wel duidelijk dat leerkrachten zich veel meer moeten focussen op het leerproces en niet zozeer op het aspect van kennisoverdracht. Ze moeten het pedagogische talent hebben deze kinderen goed te kunnen begeleiden.”

Kuipers benadrukt dat hoogbegaafde kinderen – hij spreekt overigens liever over begaafde, hoogbegaafde en excellent lerende kinderen – anders denken en leren en dus een totaal andere aanpak nodig hebben. “Het verschil met het traditionele lesgeven is best groot: je moet deze kinderen bijvoorbeeld leren reflecteren op hun eigen denken en handelen. Je moet ze planmatig leren werken. Je moet ze leren om het beste uit zichzelf te halen. Waarbij succes én falen hoort. Je moet ze triggeren en motiveren, ze laten inzien dat het leveren van prestaties moeite en discipline vergt. Veel hoogbegaafde kinderen kunnen de slag van leren-met-tweevingers-in-de-neus naar leren-dat-moeite-kost maar moeilijk maken. Ook veel leerkrachten

vinden deze verandering in didactische aanpak heel lastig, maar mijn ervaring is ook dat zij die draai wel kunnen maken.”

## **De juiste omgeving**

Volgens Kuipers moet een leerkracht beseffen dat hoogbegaafde leerlingen tachtig tot negentig procent van de basisstof die in een instructieles wordt aangeboden, binnen vijf minuten beheersen. Het gevaar van desinteresse en demotivatie liggen levensgroot op de loer. “Een leerkracht die hier bewust mee bezig is stimuleert daarom grotere denk- en verwerkingsstappen, biedt andere, vervangende leerstof aan. Dat moet regel zijn en geen uitzondering. Wie bij een onderpresterende leerling de motor weer op gang wil krijgen, zal zich heel gericht moeten bezighouden met de motivatie van het kind. Het blijkt dat hoogbegaafde kinderen erg veel baat hebben bij het zelf kiezen van de onderwerpen. Daar moeten ze een of twee uur per week zelfstandig aan mogen werken”, aldus Kuipers. De onderzoekers Philips en Lindsey noemen

dit de x-factor, de sleutel tot presteren, waarbij uitdaging (passende leerstof, samenwerken met gelijkgestemden, eigen onderwerpkeuze, zelfsturing), support (aanmoediging, waardering, reële verwachtingen) en acceptatie (door leerkrachten en medeleerlingen) centraal staan. Toch onderstrepen zij ook dat uiteindelijk de motivatie van binnenuit moet komen.

## **Organisatie**

Omdat deze extra aandacht in het reguliere onderwijs voor hoogbegaafde kinderen door de bank genomen een lastenverzwaring voor de leerkracht betekent, stelt Kuipers voor om ‘clusters van slimme kinderen te maken die in een soort plusuur of plusmiddag onder begeleiding van een leerkracht bovengroeps of zelfs bovenschools opereren’. Hij vindt overigens ook dat een goede leerkracht binnen zijn eigen lessen allerlei verrijkende opdrachten aan moet bieden. Volgens hem zijn daar inmiddels voldoende speciale materialen en leerlijnen voor ontwikkeld. “Ik pleit er niet voor om deze kinderen alleen bij elkaar te zetten. Dat vind ik een verschraling

van het reguliere onderwijs. Maar door ze op gezette tijden extra aandacht te geven, heb je een prima alternatief. Bovendien denk ik dat het voor hun sociaal-emotionele ontwikkeling ook veel beter is als ze deel uitmaken van een heterogene setting. Alleen voor kinderen die een IQ van 145 of meer hebben,

wil ik een uitzondering maken. Door de ‘gewone slimme kinderen’ van een school te halen, verlaag je het hele ambitieniveau van een school. Dat vind ik geen goede zaak.”

## Oorzaken van onderpresteren

Enkele schoolfactoren die de kans op onderpresteren vergroten, zijn:

Een ontoereikende wijze van signalering: aandacht voor (hoog)begaafdheid begint met een goede wijze van signalering. Begaafde kinderen worden lang niet altijd herkend. Bij een oppervlakkige beoordeling wordt bijna de helft van de (hoog)begaafde kinderen onderschat. Een ontoereikend leeraanbod: het reguliere onderwijssysteem verwacht dat alle leerlingen in dezelfde tijd, in dezelfde ruimte, dezelfde leerstof kunnen doorlopen om uiteindelijk dezelfde minimumprestatie te halen. Vijftien procent van de leerlingen spant zich erg in en haalt geen goede resultaten, met als gevolg demotivatie en weerspanning tegen leren. Vijftien procent spant zich nauwelijks in en behaalt toch een goed resultaat, met soms als gevolg demotivatie en weerspanning tegen leren. In het hedendaagse onderwijs vindt leren plaats in kleine leerstappen met veel oefeningen. Dit is de geëigende werkwijze van veel leerkrachten en deze werkwijze is voor veel leerlingen ook uitstekend geschikt. (Hoog)begaafde leerlingen leren echter in grote stappen met weinig oefentijd. Zij kunnen bovendien in minder tijd veel meer en grotere leersprongen maken dan van hen verwacht wordt. Ook dat leidt onherroepelijk tot onderpresteren en omdat het plezier in het leren ontbreekt, tot verveling, toberigheid en bezorgdheid.

Te weinig ruimte voor creatief denken: (hoog)begaafde kinderen beschikken naast een hoge intelligentie vaak ook over een hoge vorm van creatief of divergent denken. Creatieve denkers zien problemen, zien oplossingen en zien verbanden die anderen niet zien. Hun opvattingen druisen vaak in tegen de gevestigde opvattingen en worden als zodanig niet positief gewaardeerd. Ook de strakke hantering van huidige onderwijsmethodes staat het creatieve proces van kinderen maar moeizaam toe. Daardoor werken veel scholen de stimulering van creatief denken eerder tegen dan dat het wordt aangemoedigd, en dit werkt onderpresteren in de hand.

Te weinig aandacht voor metacognitieve vaardigheden: metacognitieve kennis omvat de kennis en opvattingen die mensen hebben over hun eigen cognitief functioneren, vaardigheden die het leerproces aansturen, zoals plannen en reflecteren (vrij vertaald: nadenken over het eigen denken). Metacognitieve kennis is noodzakelijk om nieuwe kennis te verwerven. De leerkracht krijgt in dit kader de rol van coach, waarbij hij zeer individueel en op belangstellende wijze het beste uit de leerling naar boven weet te halen. Overigens zijn lang niet alle (hoog)begaafde kinderen in metacognitief opzicht vaardig: regelmatig oriënteren deze kinderen zich te kort op taken, werken ze te snel of vergeten ze bijvoorbeeld het werk na te kijken. Onderpresteerders zijn hierin juist zwak.

Ontoereikende pedagogisch-didactische kwaliteiten van de leerkracht: het spreekt vanzelf dat de rol van de leerkracht een belangrijke – zo niet de belangrijkste – factor is voor een succesvolle ontwikkeling van alle leerlingen. Gedifferentieerd lesgeven wordt in het gehele onderwijs als belangrijke werkwijze erkend. Iedere leerling heeft immers zijn eigen leerstijl. Daarom moeten leraren, om zo effectief mogelijk te onderwijzen, tijdens het lesgeven rekening houden met aan wie en wat zij onderwijzen. Het doel van de gedifferentieerde klas is actief plannen en voortdurend elke leerling helpen om zo ver en zo snel mogelijk te vorderen op een leercontinuüm.

Voor een juiste begeleiding van (hoog)begaafde leerlingen is het nodig dat de leerkracht in ieder geval erkent dat deze kinderen een andere aanpak vragen. Veel leerkrachten hebben echter moeite om



leerlingen ruimte te geven, ze willen 'in control' blijven. Dit uit zich bijvoorbeeld in de angst om leerstof te schrappen, waardoor het vergroten van de leerstappen wordt belemmerd. Maar ook veel leerlingen aarzelen om de ruimte te nemen. Ze hebben zich afhankelijk gemaakt van de controle van de leerkracht. Onderpresteerders waarbij onzekerheid speelt, bijvoorbeeld als gevolg van perfectionisme, hebben veel behoefte aan het oordeel van de leerkracht. Ook de rol van klasgenoten is van invloed; het streven naar aanpassen aan de groepsnorm is belangrijker dan het streven naar het leveren van hoge prestaties.

Bron: [pulseprimaironderwijs.nl](http://pulseprimaironderwijs.nl)

[eureka.inos.nl/?q=node/127](http://eureka.inos.nl/?q=node/127)

## **Bijlage 2: Tekst ‘samenvatting’, niveau B2-**

### **Samenvatting “Beter omgaan met hoogbegaafde leerlingen”**

Ik heb een interview met Jan Kuipers gelezen, hij werkt als onderwijsadviseur, het interview gaat over hoe kunnen we met hoogbegaafde leerlingen behandelen?

Vroeger was de zorg vooral op de zwakke leerlingen gericht. Hoogbegaafde leerlingen konden zichzelf wel redden, was de overheersende opinie. Het aangeboren talent werd als een luxe gezien, waarin je geen extra tijd hoefde te steken.

Kuipers benadrukt dat hoogbegaafde kinderen – hij spreekt overigens liever over begaafde, hoogbegaafde en excellent lerende kinderen – anders denken en leren en dus een totaal andere aanpak nodig hebben. “Het verschil met het traditionele lesgeven is best groot.

Voor een juiste begeleiding van (hoog)begaafde leerlingen is het nodig dat de leerkracht in ieder geval erkent dat deze kinderen een andere aanpak vragen. Veel leerkrachten hebben echter moeite om leerlingen ruimte te geven, ze willen ‘in control’ blijven. Dit uit zich bijvoorbeeld in de angst om leerstof te schrappen, waardoor het vergroten van de leerstappen wordt belemmerd. Maar ook veel leerlingen aarzelen om de ruimte te nemen.

### **Bijlage 3: Tekst ‘samenvatting’, niveau B2+**

#### **Samenvatting “Beter omgaan met hoogbegaafde leerlingen”**

Niet alleen de zwakke leerlingen maar ook hoogbegaafde leerlingen hebben wat extra aandacht nodig. Ze hebben hun eigen benadering nodig om niet geïsoleerd te voelen en om te gaan met hun psychosociale problemen. Met deze problemen onderpresteren de kinderen vaak. Hoogbegaafde leerlingen kunnen briljant zijn op een moment, maar dan kunnen ze onverschillig worden. Dit veroorzaakt frustratie voor het kind en ook voor ouders and leerkrachten.

Onderwijsadviseur Jan Kuipers denkt dat onderpresteren het gevolg is van een ontoereikende lesaanbod met een laag ambitieniveau. Hij suggereert dat een andere pedagogisch-didactische aanpak is belangrijk door de leerkracht. Leerkrachten moeten de kinderen goed te begeleiden door ze te motiveren en laten begrijpen dat het leveren van prestaties moeite vergt. Het blijkt uit onderzoek dat het goed is voor hoogbegaafde kinderen als ze van de onderwerpen zichzelf kunnen kiezen. Uitdaging, support en acceptatie zijn ook centraal voor goed presteren.

Kuipers vindt dat hoogbegaafde kinderen in een heterogene setting zouden moeten blijven, ook voor hun sociaal-emotionele ontwikkeling. Maar ze kunnen in een soort plusuur onder begeleiding van een boven groeps leerkracht extra aandacht krijgen.

Er zijn enkele factoren die oorzaken van onderpresteren kunnen zijn. Bijvoorbeeld, begaafde kinderen worden niet altijd herkend. Begaafde leerlingen leren veel meer in minder tijd, en dit leidt tot onderpresteren en demotivatie. Er is vaak ook te weinig ruimte voor creatief denken, die is aanwezig in begaafde kinderen, en voor metacognitieve vaardigheden. Ten slotte, de leerkracht is soms niet goed genoeg voor deze moeilijke taak, hoewel hij centraal is voor de succesvolle ontwikkeling van alle leerlingen. Zijn pedagogisch-didactische vaardigheden kunnen ontoereikende zijn.

Leerkrachten moeten een andere aanpak gebruiken met deze kinderen, zelfs als ze moeite hebben om leerlingen ruimte te geven. Maar ook leerlingen aarzelen soms om ruimte te nemen en worden beïnvloed door zijn klasgenoten om aan de groepsnorm aan te passen.

## **Bijlage 4: Tekst ‘samenvatting’, niveau NT1**

### **Samenvatting ‘Beter omgaan met hoogbegaafde leerlingen’**

In het artikel staat het beter omgaan met hoogbegaafde leerlingen centraal. Vroeger werd gedacht dat hoogbegaafde leerlingen zichzelf wel konden redden, daarin heeft een omslag plaatsgevonden en nu komt er naar voren dat er extra aandacht voor deze groep nodig is om onderpresteren te voorkomen. Er wordt een onderscheid gemaakt tussen 1) begaafde, 2) hoogbegaafde en 3) excellent lerende kinderen. Er worden vier onderwerpen genoemd die correleren met onderpresteren.

#### 1) Ontoereikend onderwijsaanbod

Hierin komt naar voren dat ontoereikend lesaanbod in het basisonderwijs gecombineerd met een te laag ambitieniveau van de ouders onderpresteren tot gevolg heeft. In andere landen ligt de lat veel hoger, hoewel daarin in Nederland nu ook een omslag in plaatsvindt. Echter, eist deze omslag wel een andere pedagogisch-didactische aanpak waarin er meer op het leerproces gefocust wordt in plaats van het centraal stellen van kennisoverdracht.

#### 2) Omgeving

In dit gedeelte komt naar voren dat (hoog)begaafde kinderen 80-90 procent van basisstof binnen vijf minuten beheersen. Hierdoor ligt gevaar van desinteresse en demotivatie op de loer. Er moet daarom gedacht worden aan vervangende leerstof. Zelfdeterminatie is deugdzaam voor de (hoog)begaafde kinderen. In het onderzoek van Philips en Lindsey komt een x-factor naar voren die zelfdeterminatie ondersteunen, hierbij staan uitdaging, support en acceptatie centraal. centraal staan.

#### 3) Organisatie

Het kopje organisatie stelt dat (hoog)begaafde kinderen baat hebben bij op gezette tijden extra aandacht in clustergroepen te krijgen. Desalniettemin is het voor de sociaal-emotionele ontwikkeling van belang om de rest van de tijd in een heterogene setting les te krijgen.

#### 4) Oorzaken van onderpresteren

Oorzaken van onderpresteren liggen met name in ontoereikende signalering van (hoog)begaafdheid, demotivatie bij de leerlingen, te weinig ruimte voor creatief denken en te weinig aandacht voor metacognitieve vaardigheden. Daarnaast komt er naar voren dat het bij veel leerkrachten aan toereikende pedagogisch-didactische kwaliteiten ontbreekt.

Concluderend is het nodig, om beter te kunnen omgaan met (hoog)begaafde leerlingen, dat de leerkracht erkent dat er een andere aanpak nodig is voor (hoog)begaafde leerlingen, waarin zelfdeterminatie meer ruimte krijgt. De crux in dit alles ligt in aan de ene kant de angst van leraren om de manier van onderwijsgeven aan te passen en aan de andere kant de aarzeling van leerlingen om de benodigde ruimte in te nemen.

## **Bijlage 5: Tekst ‘e-mail over onderzoek’, niveau B2-**

Beste <naam>,

Ik weet dat je heel erg geïnteresseerd bent met de cyberbeveiliging, ik heb een artikel gelezen, Er zijn heel veel belangrijkste informatie over onze digitale erfgoed. Het aantal ongekomen ‘gebruikers’ stagen naar 4.9 miljard het einde van deze eeuw, dus het belangrijkste vraag is, wie heeft de recht op al deze gegevens?

Er zijn veel analyse verwacht dat ongeveer 1.4 miljard Facebook leden zullen doden voor 2100, in een andere woorden kunnen we verwachten dat 4.9 miljard Facebook ‘gebruikers’ zullen doden, en dit is echt niet normaal.

Bovendien is er een aanmerkelijke vraag, hoe deze profielen van doodgegaan moet beheren? Wie heeft deze recht? De onderzoeker David Watson zie:” Het beheer van dit archief heeft dus ook gevolgen voor de controle op onze geschiedenis.”

Tegenwoordig heeft Facebook een Historisch archief! Dit archief helpt archeologen om onze erfgoed te verzamelen en beschermen als we komen te overleden.

Naar mijn mening dit is een significant onderzoek, want we gebruiken social media in dagelijkse leven, dus we moeten over onze digitale erfgoed overdenken!.

Je kunt meer informatie bij deze link vinden: ( <https://www.scientias.nl/over-vijftig-jaar-hebben-meer-doden-dan-levenden-een-facebook-account/> )

Groetjes,

<naam>

## **Bijlage 6: Tekst ‘e-mail over onderzoek’, niveau B2+**

Beste <naam>,

Ik heb net een interessant artikel gelezen. De titel van het artikel is “Marslander Insight meet voor het eerst een mogelijke ‘marsbeving’”. Ik weet dat je geïnteresseerd bent in allerlei onderzoeken over Mars. Daarom wil ik jou uitleggen wat het artikel over gaat.

Je weet dat Marslander Insight al een tijdje op Mars blijft om de rode planeet te onderzoeken en op de 128<sup>e</sup> Marsdag kwam een groot nieuws van de lander: De eerste marsbeving is gemeten! Het was toch geen grote trilling maar het was wel een beving. Het signaal werd gemeten door de seismometer. En de beving wordt waarschijnlijk niet veroorzaakt door bijvoorbeeld wind van boven het oppervlak. Dus het betekent dat de beving uit het binnenste van Mars kwam. En deze beving toont ook aan dat Mars nog steeds seismisch actief is.

We weten dat verschillende en bewegende platen van de aardkorst ontstaan van beving van het oppervlak van de aarde. Maar er is op Mars geen sprake van platentektoniek. Volgens de onderzoekers is het toch mogelijk om het oppervlak van Mars te beven. Als dit soort bevingen kunnen worden waargenomen, kunnen de onderzoekers meer begrijpen hoe Mars en andere rotsachtige planeten zijn gevormd.

Nu is het tijd om verder de data die onderzoekers over Marsbeving verzamelden te onderzoeken zodat ze kunnen proberen te begrijpen of de signalen van de seismometer over het binnenste van de planeet juiste informatie geven.

Hoewel deze beving niet groot genoeg is om het diepe binnenste van Mars te onderzoeken, vind ik het nog heel belangrijke mijlpaal want we kunnen steeds meer informatie over de planeten krijgen. En ik geloof dat de mensen binnenkort gemakkelijk naar Mars reizen om daar te wonen! Ik hoop dat jij het artikel ook interessant vindt.

Groeten,

<naam>

## **Bijlage 7: Tekst ‘e-mail over onderzoek’, niveau NT1**

Hee <naam>,

Je vertelde me vanmiddag dat je geïnteresseerd bent in het artikel dat ik aan het lezen was: [‘deze airconditioning die CO2 vangt en omzet in synthetische brandstof’](#). Ik zal je er wat meer over vertellen.

Zoals je weet, zit onze atmosfeer veel te vol met CO<sub>2</sub>, waardoor onze planeet opwarmt. In een nieuwe Duitse studie bekeken onderzoekers kantoorpanden, supermarkten en huizen en kwamen tot de ontdekking dat bestaande airconditionings omgebouwd kunnen worden zodat ze CO<sub>2</sub> en water in de lucht direct omzetten in synthetische brandstoffen. Zo slaat deze airconditioning twee vliegen in één klap: het neemt CO<sub>2</sub> weg uit de atmosfeer én maakt er milieuvriendelijke brandstof van.

Hoewel de onderzoekers wel wat getallen noemen van het potentiële effect van deze airconditionings, zetten ze dit niet af tegen bijvoorbeeld de doelen in het Parijsakkoord. Het is dus onduidelijk in welke mate deze oplossing bijdraagt aan het terugdringen van CO<sub>2</sub> in de atmosfeer.

De onderzoekers maken naar mijn mening wel een ander terecht punt: hoewel er zowel publieke als wetenschappelijke belangstelling is om de hoeveelheid CO<sub>2</sub> in de atmosfeer terug te brengen, lijken de industrie en de politiek minder haast te hebben. Ik denk, net als de onderzoekers, dat daar het grootste verschil gemaakt kan worden!

Veel leesplezier, mocht je besluiten het zelf te lezen!

Groetjes,

<naam>

## Bijlage 8: Overzicht vragen in Qualtrics

In hoeverre vindt u de volgende aspecten aanwezig in of kenmerkend voor deze tekst?

Waarbij 1 = helemaal niet, 4 = net voldoende, 7 = volledig aanwezig

	1	2	3	4	5	6	7
Adequaatheid (de tekst is een uitwerking van de opdracht)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Begrijpelijkheid (de tekst is goed te begrijpen)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Grammaticale oorrectheid (bv. gebruik van wel/geen lidwoord, congruentie van werkwoord en subject)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Correcte spelling (woorden zijn op de juiste manier geschreven)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Samenhang (er is een verband tussen zinnen en alinea's, bijvoorbeeld door signaal- en/of verbindingswoorden)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Logische opbouw (de tekst is logisch opgebouwd, bv. inleiding, kern, slot)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Gepast woordgebruik (het woordgebruik sluit aan bij het onderwerp van de opdracht)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	1	2	3	4	5	6	7
Voldoet deze tekst als geheel aan het niveau dat u verwacht van een eerstejaars student op uw opleiding? (waarbij 1 = Helemaal niet - 4 = net voldoende - 7 = Volledig)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

U kunt hieronder uw keuze(s) toelichten of andere opmerkingen met betrekking tot de tekst achterlaten