|  |
| --- |
|  |
| Validatie in tekstverwerking |
| Het effect van de correctheid van informatie op informatieverwerking en het onthouden ervan |

|  |
| --- |
| M.H.W. (Milot) Tubben  Studentnr: 6141765  Begeleider: M.L. (Marloes) van Moort  1-1-2021 |



“Nobody is going to pour truth into your brain. It’s something you have to find out for yourself”

A.N. Chomsky (2000)



# Voorwoord

Gedurende mijn bachelor Communicatie- en Informatiewetenschappen ben ik altijd gefascineerd geweest door cognitieve mechanismen, waaronder de processen die een rol spelen bij tekstverwerking. Dit scriptieonderwerp was mij dan ook op het lijf geschreven en ik heb mij er met veel plezier in verdiept.

Mijn scriptie had niet tot stand kunnen komen zonder hulp van enkele mensen. Mijn begeleider Marloes van Moort heeft mij erg geholpen in het formuleren van een adequate onderzoeksvraag en heeft door het beantwoorden van al mijn vragen vele inzichten geboden. Daarnaast wil ik alle participanten, die ik tot in den treure heb achternagezeten, bedanken voor hun deelname aan het experiment.

# Inhoudsopgave

# Samenvatting p. 4

1. Inleiding p. 5

2. Theoretisch kader p. 6

3. Probleemstelling p. 8

4. Methode p. 10

5. Resultaten p. 14

6. Discussie p. 15

7. Conclusie p. 17

Literatuuroverzicht p. 18

# Samenvatting

# Tijdens het lezen van teksten zijn lezers voortdurend bezig met het ontwikkelen van tekstbegrip. Dit doen zij door het vormen en updaten van hun situatiemodel (O’Brien et al., 1998). Om het situatiemodel te beschermen tegen incorrecte informatie controleren lezers de binnenkomende informatie op correctheid via het proces van validatie (Singer, 2006). Dit validatieproces kan van invloed zijn op online tekstverwerking en de offline mentale representatie van de gelezen tekst. Deze studie onderzoekt de invloed van de correctheid van informatie op leestijd en het onthouden van die informatie. Dit is gedaan door een experiment uit te voeren onder HBO-/WO-studenten van 18-35 jaar. De participanten lazen teksten die ofwel correcte, ofwel incorrecte informatie bevatten en de dag daarna maakten zij een geheugentaak. Uit de resultaten is gebleken dat participanten langer doen over het lezen van incorrecte informatie dan correcte informatie (het zogeheten *inconsistency effect* (Cook & O’Brien, 2014; Van Moort et al., 2018; Van Moort et al., n.d.)). Daarnaast werd de incorrecte informatie minder goed onthouden dan de correcte informatie. Dit ondersteunt eerdere theorieën over tekstverwerking en validatie, namelijk dat lezers meer moeite hebben met incorrecte informatie dan met correcte informatie (Van Moort et al., n.d.). Hoewel het wel verwacht werd, is er geen relatie gevonden tussen de grootte van het inconsistency effect en de mate van onthouden. Omdat deze relatie vanuit theorie bezien aannemelijk is, is er waarschijnlijk een methodologische verklaring voor het niet vinden ervan. Vervolgonderzoek zou hier aandacht aan moeten besteden.

# 1. Inleiding

Mensen komen overal teksten tegen; in kranten, tijdschriften en op sociale media. Hoewel men zich er tijdens het lezen waarschijnlijk niet bewust van is, zijn er op de achtergrond allerlei mentale processen bezig om de tekst die gelezen wordt te begrijpen. Een essentieel proces in het ontwikkelen van tekstbegrip is het maken van een mentaal model van de inhoud van de tekst, ook wel het *situation model* (hierna: situatiemodel) genoemd. In dit situatiemodel wordt gelezen informatie uit de tekst geïntegreerd met voorkennis van de lezer, om zo een mentale representatie van de tekst te maken en uiteindelijk tot begrip van de tekst te komen (O’Brien et al., 1998).

In de hedendaagse maatschappij is ook een ander aspect van tekstbegrip van groot belang, namelijk het omgaan met incorrecte informatie. Op sociale media, in populaire media en soms zelfs in statements van publieke figuren komt misinformatie voor (Shu et al., 2017). Vaak is dit in de vorm van *fake news*, een bewust misleidend nieuwsbericht (Gelfert, 2018). Het situatiemodel dat de lezer gevormd heeft kan verstoord worden door misinformatie. Daarnaast kan in het situatiemodel geïntegreerde incorrecte informatie verder tekstbegrip negatief beïnvloeden (Singer, 2019).

Het is dus zaak dat lezers zichzelf beschermen tegen incorrecte informatie. Dit doen zij door middel van een proces dat is gedefinieerd als *validation* (hierna: validatie). Tijdens validatie controleren lezers de binnenkomende informatie op correctheid, alvorens zij de informatie opnemen in het situatiemodel. (Singer, 2006). Dit proces kan invloed hebben op hoe de verwerking van de gelezen tekst plaatsvindt en op hoe de gelezen informatie onthouden wordt. Dit onderzoek poogt daar meer inzicht in te krijgen en daarom staat de volgende vraag centraal: “Wat is het effect van de correctheid van informatie op informatieverwerking en het onthouden van de gelezen informatie?”

# 2. Theoretisch kader

Tijdens het lezen van een tekst zijn lezers continu bezig met het ontwikkelen van tekstbegrip. Dit doen zij onder meer door het creëren en updaten van hun situatiemodel. In het situatiemodel maken lezers representaties van de tekst op vijf dimensies: tijd, ruimte, personages, causaliteit en motivatie. Op die manier maken de lezers een mentale voorstelling van de tekst (Zwaan et al., 1998; Zwaan et al., 1995). Tijdens het lezen blijven zij continu monitoren of de informatie die zij tegenkomen al in het situatiemodel is opgenomen. Als dat niet het geval is moet het situatiemodel geüpdatet worden zodat de nieuwe informatie ook geïntegreerd kan worden. Zo is de lezer constant bezig met het reguleren en integreren van binnenkomende informatie (Zwaan et al, 1998).

Het situatiemodel is van grote invloed op de manier waarop informatie uit een tekst onthouden wordt. Zo bleek uit een serie experimenten dat mentale modellen (zoals het situatiemodel) ervoor zorgen dat informatie op een structurele manier wordt opgeslagen in het langetermijngeheugen, waardoor het beter onthouden wordt (Radvansky & Zacks, 1991).

Het RI-Val model van tekstbegrip biedt een raamwerk om de processen van situatiemodel-constructie in te plaatsen. Het model is het resultaat van een combinatie van verschillende modellen, waaronder het C-I model van Kintsch (1988), en enkele experimenten, zoals die van Cook en O’Brien (2014). Het RI-Val model gaat uit van drie ‘fasen’ in tekstbegrip; activatie, integratie en validatie. In de activatiefase worden concepten uit het geheugen geactiveerd en betrokken bij de binnenkomende informatie. Dit gebeurt via het proces van *spread of activation*, wat wordt getriggerd wanneer men informatie leest. Spread of activation houdt in dat de gelezen informatie de activatie van allerlei mentale concepten in het geheugen van de lezer, die te maken hebben met de gelezen informatie, in gang zet (Van den Broek, Rapp & Kendeou, 2005). In de integratiefase worden verschillende concepten aan elkaar gelinkt, zowel de concepten die geactiveerd zijn in het geheugen tijdens de activatiefase, als nieuwe gelezen informatie. De integratie gebeurt op basis van hoe goed de concepten bij elkaar passen. In de integratiefase wordt dus informatie uit het geheugen gecombineerd met tekstuele informatie, wat resulteert in de constructie van een situatiemodel. (O’Brien & Cook, 2016a; O’Brien & Cook, 2016b). In de laatste fase, de validatiefase, worden de gelinkte concepten gevalideerd tegen achtergrondkennis of eerder gelezen tekst om zo te beoordelen of de aan elkaar gelinkte concepten ‘kloppen’ met de eigen kennis van de lezer. (Cook & O’Brien, 2014; O’Brien & Cook, 2016a; O’Brien & Cook, 2016b). Tijdens validatie wordt de binnenkomende informatie dus ‘gecontroleerd’ op correctheid voordat het definitief in het situatiemodel wordt opgenomen. (Singer, 2006).

Validatie kan zoals gezegd op twee manieren plaatsvinden; door informatie te vergelijken met eigen achtergrondkennis over het onderwerp, of door de informatie te vergelijken met eerdere informatie uit de tekst. Bij de eerste variant worden contradicties tussen de informatie in de tekst en de achtergrondkennis van de lezer gedetecteerd en bij de tweede variant let men vooral op inconsistenties binnen de tekst (Singer, 2013; Van Moort et al., 2018). In dit onderzoek zal alleen gekeken worden naar de validatie tegen de eigen achtergrondkennis.

Valideren ten opzichte van achtergrondkennis is essentieel in het voorkomen dat incorrecte informatie in het situatiemodel terechtkomt. Dit is belangrijk omdat incorrecte informatie die in het situatiemodel is opgenomen verder tekstbegrip kan vermoeilijken (Singer, 2019).

Er zijn drie belangrijke assumpties in het RI-Val model. De eerste is dat de drie processen in het RI-Val model zowel asynchroon als parallel zijn. De processen beginnen niet tegelijk; integratie kan pas beginnen als ten minste twee concepten zijn geactiveerd in de activatiefase, anders valt er immers niets te integreren. Daarnaast kan validatie pas beginnen als er reeds een aantal concepten aan elkaar zijn gelinkt in de integratiefase. Maar de processen kunnen wel overlappen; zo kan integratie nog niet volledig zijn afgelopen als validatie al is begonnen. Daarnaast zijn alle drie de processen in dit model passief en worden dus niet bewust door de lezer geïnitieerd. Dit heeft als gevolg dat de processen op de achtergrond altijd ‘volledig’ worden doorlopen; de processen gaan door totdat zij geen verdere invloed meer hebben op tekstbegrip. De laatste assumptie houdt in dat de lezer pas doorleest als de drie processen, en in het bijzonder de validatiefase, hebben gezorgd voor een voldoende coherent situatiemodel van de binnengekomen informatie. (Cook & Brien, 2014; O’Brien & Cook, 2016a; O’Brien & Cook, 2016b).

Naast het RI-Val model kan validatie ook gedefinieerd worden in termen van *plausibility* (hierna: plausibiliteit). Zo beschrijft Richter (2015) validatie als een proces waarbij de lezer een oordeel vormt over de plausibiliteit van de tekst. Plausibiliteit wordt hierbij gedefinieerd als de aanvaardbaarheid en/of waarschijnlijkheid van een situatie of van een zin die die situatie beschrijft (Matsuki et al., 2011). Na afloop van het validatieproces heeft de lezer dan een oordeel over hoe aanvaardbaar en waarschijnlijk de informatie uit de tekst is. Lombardi, Sinatra en Nussbaum (2013) toonden vervolgens aan dat de waargenomen plausibiliteit van de lezer van sterke invloed is op hoe de gelezen informatie wordt geïntegreerd in het situatiemodel.

Onderzoek naar validatie maakt veelal gebruik van het inconsistentie paradigma (Van den Broek, Rapp & Kendeou, 2005). Het paradigma houdt in dat lezers een stukje tekst lezen waar een contradictie in staat (een veelgebruikt voorbeeld: “*Mary is a vegetarian.”* (…) “*Mary ordered a hamburger and fries.”* (Albrecht & O’Brien, 1993; Myers, O’Brien, Albrecht & Mason, 1994; Cook & O’Brien, 2014)). Door het effect van deze inconsistentie op bijvoorbeeld online verwerking (in de vorm van leestijd) te bestuderen kan dan een theorie gevormd worden over hoe validatie plaatsvindt (Van den Broek, Rapp & Kendeou, 2005). Een veelvuldig gevonden resultaat is het *inconsistency effect*, dit is het effect dat leestijd langer wordt als lezers informatie tegenkomen die niet overeenkomt met hun eigen kennis. Men vermoedt dat dit het gevolg is van validatie, omdat lezers in het geval van inconsistenties -hetzij met eerdere tekstuele informatie, hetzij met eigen kennis- meer cognitieve moeite moeten doen om de incorrecte informatie te verwerken. (Cook & O’Brien, 2014; Van Moort et al., n.d.; Van Moort et al., 2018). Ook vanuit het RI-Val model kan verwacht worden dat validatie voor langere leestijd zorgt, omdat de lezer volgens dat model pas doorleest als de drie processen -activatie, integratie en validatie- hebben gezorgd voor een coherent situatiemodel en voldoende tekstbegrip. Incorrecte informatie zou ervoor zorgen dat het moeilijker wordt om een coherent situatiemodel te vormen, en daardoor zou de leestijd langer worden. Deze verlengde leestijd is vooral het resultaat van een langere validatiefase. (O’Brien & Cook, 2016a; O’Brien & Cook, 2016b).

De algemene aanname is dat processen die tijdens het lezen plaatsvinden, waaronder validatie, ook invloed hebben op de offline representatie van de tekst, hetgeen dat lezers onthouden van een tekst (Van Moort et al., n.d.). Informatie die als incorrect wordt bestempeld door het validatieproces wordt, niet in het situatiemodel wordt opgenomen, waardoor het minder goed wordt onthouden (Schroeder et al., 2008; Van Moort et al., 2020; Van Moort et al., n.d.). Andersom is ook het geval dat informatie die als gevolg van het validatieproces als correct wordt aangenomen en dus wel in het situatiemodel wordt opgenomen ook beter wordt onthouden. Incorrecte informatie wordt aldus minder goed onthouden dan correcte informatie. (Schroeder et al., 2008; Van Moort et al., 2020; Radvansky & Zacks, 1991).

# 3. Probleemstelling

Vraagstelling

Uit bovenstaand theoretisch kader wordt duidelijk dat er al veel onderzoek is gedaan naar validatie. Veel van deze onderzoeken hebben echter gekeken naar ofwel online verwerking, ofwel offline mentale representatie. Er is nauwelijks onderzoek dat online verwerking en offline representatie combineert in één studie. Dit onderzoek poogt dit gat te dichten. Daarnaast tracht dit onderzoek de theorieën en modellen, zoals het RI-Val model, die het resultaat zijn van eerdere onderzoeken te ondersteunen.

Dit zal gedaan worden door antwoord te geven op de volgende onderzoeksvraag: “Wat is het effect van de correctheid van informatie op informatieverwerking en het onthouden van deze informatie?” Deze vraag is beantwoord door gebruik te maken van een inconsistentie paradigma (Van den Broek, Rapp & Kendeou, 2005). Participanten lazen korte stukjes tekst die targetzinnen bevatten die ofwel correct waren, ofwel incorrect. Tijdens het lezen is leestijd bijgehouden om online verwerking te monitoren. Ook is het geheugen van de participanten over de gelezen informatie de volgende dag getest om offline representatie van de tekst te onderzoeken. Deze combinatie van online verwerking en offline representatie biedt de mogelijkheid om verschillende verwachtingen te toetsen.

Op basis van eerder onderzoek wordt verwacht dat lezers een inconsistency effect zullen laten zien en dus langer zullen doen over het lezen van incorrecte informatie dan over het lezen van correcte informatie (Hypothese 1) (Cook & O’Brien, 2014; Van Moort et al., 2018; Van Moort et al., n.d.). Daarnaast is de verwachting dat lezers incorrecte informatie minder goed zullen onthouden dan correcte informatie (hypothese 2), omdat de incorrecte informatie waarschijnlijk niet in het situatiemodel wordt opgenomen, waardoor het minder goed wordt onthouden. Correcte informatie wordt daarentegen wel in het situatiemodel opgenomen, wat ervoor zorgt dat deze informatie beter wordt onthouden.  (Schroeder et al., 2008; Van Moort et al., 2020; Van Moort et al., n.d.; Radvansky & Zacks, 1991).

Tot slot zal gekeken worden naar de relatie tussen de online verwerking en de offline mentale representatie. De verwachting is dat er een relatie is tussen de grootte van het inconsistency effect en de mate waarin de informatie onthouden wordt. Dit is aannemelijk, omdat online verwerking en offline mentale representatie niet los van elkaar staan. Het zijn immers allebei factoren in het proces van tekstbegrip.

Het RI-Val model stelt dat een langere leestijd -dus een groter inconsistency effect- duidt op het feit dat de drie processen activatie, integratie en validatie bezig zijn een coherent situatiemodel van de tekst te vormen (Cook & O’Brien, 2014; O’Brien & Cook, 2016a; O’Brien & Cook, 2016b). Een meer coherent situatiemodel zorgt er in principe voor dat informatie beter wordt onthouden (Schroeder et al., 2008; Van Moort et al., 2020; Van Moort et al., n.d.; Radvansky & Zacks, 1991). Als dit klopt, zullen de resultaten van dit onderzoek een positieve relatie laten zien tussen de grootte van het inconsistency effect en de mate van onthouden. Een langere leestijd -dus groter inconsistency effect- zal dan tot beter onthouden van de informatie leiden.

Resultaten van ander onderzoek (Van Moort et al., n.d.) lieten zien dat hoewel lezers langer deden over het lezen van incorrecte informatie, zij deze informatie niet beter onthielden dan correcte informatie. Dit werd verklaard vanuit het idee dat er meer interferentie was tussen de gelezen informatie en de eigen kennis over het onderwerp. Daardoor zou het lastiger worden om de gelezen informatie te onderhouden. Een andere mogelijke verklaring was dat de correcte informatie wellicht niet goed geïntegreerd kon worden in het situatiemodel, waardoor het minder goed kon worden onthouden. (Van Moort et al., n.d.). Als deze theorieën kloppen, zullen de resultaten van dit onderzoek laten zien dat een kleiner inconsistency effect leidt tot het beter onthouden van de gelezen informatie.

Relevantie

Dit onderzoek zal bijdragen aan het ondersteunen van reeds bekende wetenschappelijke theorieën over tekstverwerking en validatie, waaronder het inconsistency effect en het RI-Val model. Daarnaast zijn er nauwelijks onderzoeken geweest die aandacht hebben besteed aan zowel online verwerking als offline representatie van informatie in relatie tot de correctheid van informatie. Door het effect van incorrecte informatie op leestijd en het onthouden van die informatie te onderzoeken kan meer duidelijkheid verkregen worden over hoe het proces van validatie plaatsvindt. Daarnaast kan kennis over hoe mensen omgaan met incorrecte informatie van waarde zijn in het ontwikkelen van kennis en adviezen over hoe om te gaan met misleidende informatie en fake news.

# 4. Methode

# Dit onderzoek is deel van een groter onderzoek. Dit grotere onderzoek heeft een 2 (historisch correcte informatie/historisch incorrecte informatie) x 2 (wel/niet congruent met eerder gelezen tekst) x 2 (instructie: aangeven of zin klopt met voorkennis/aangeven of zin past in de tekst) design. Voor dit deelonderzoek is enkel naar de manipulatie historisch correcte/historisch incorrecte informatie gekeken. In het restant van deze methodebeschrijving zal daarom alleen het deelonderzoek worden beschreven.

Participanten

In totaal hebben 79 HBO/WO-studenten tussen de 18-35 jaar deelgenomen aan dit onderzoek. Bij drie participanten ontbraken een aantal datapunten, waardoor zij weggelaten zijn uit de data-analyse. Daarmee komt het aantal geanalyseerde participanten op 76 (waarvan 35 vrouw en 41 man). De gemiddelde leeftijd is 24. Alle participanten zijn Nederlandssprekend. De participanten zijn verzameld door in eigen kringen rond te vragen en daarbij gebruik te maken van een sneeuwbalsteekproef (Boeije, 2016). Participanten hebben geen gediagnosticeerde lees- of leerproblemen. Voor deelname hebben de participanten een *informed consent* ondertekend.

Variabelen

Voor het toetsen van de hypotheses zijn een aantal variabelen gebruikt. De onafhankelijke variabele die van belang is voor dit onderzoek is de correctheid van de informatie in de targetzin.

Voor de eerste hypothese is gekeken naar de leestijden per zin als maat van online verwerking. De leestijd per zin betreft de tijd tussen het moment dat een zin verschijnt tot het moment dat de participant doorklikt naar de volgende zin.

Voor de tweede hypothese is het percentage goed beantwoorde vragen op de geheugentest gebruikt. Deze variabele test de offline geheugenrepresentatie, oftewel wat de participanten hebben onthouden van de tekst.

Bij de relatie tussen online verwerking en offline representatie zijn de eerder genoemde variabelen gecombineerd. Voor deze derde hypothese zijn dus geen aparte manipulaties of uitkomstmaten gebruikt.

Materialen

De participanten hebben tijdens het experiment zestien teksten gelezen. Deze teksten gingen over historische onderwerpen die bekend zouden moeten zijn voor de participanten. Op basis van de uitkomsten van een onderzoek waarin de onderwerpen van de teksten genormeerd zijn in een vergelijkbare groep deelnemers kunnen de onderwerpen van de teksten als bekend worden verondersteld (zie voor meer informatie over de normering Van Moort et al, 2018). Voorbeelden van onderwerpen die zijn gebruikt zijn het zinken van de Titanic, de deelname van Conchita Wurst aan het Eurovisie Songfestival en de afschaffing van de basisbeurs voor studenten. De teksten bevatten elk een targetzin, die ofwel historisch correct ofwel historisch incorrect was. De teksten met correcte en incorrecte targetzinnen zijn willekeurig verdeeld onder de participanten op zo’n manier dat alle participanten evenveel correcte als incorrecte targetzinnen lazen.

De teksten waren allemaal tien zinnen lang. De eerste twee zinnen dienden als introductie in het onderwerp van de tekst. De vijf zinnen daarna gaven context over het onderwerp van de tekst. De achtste zin was altijd de targetzin. Deze zin was, zoals eerder genoemd, ofwel historisch correct ofwel historisch incorrect. De laatste twee zinnen dienden als conclusie. De teksten telden gemiddeld 121 woorden. Daarnaast is ervoor gezorgd dat de lengte van de correcte en incorrecte targetzinnen hetzelfde was, zodat leestijd betrouwbaar kon worden gemeten (Van Moort et al., 2018).

# In tabel 1 staat een voorbeeld van twee tekstversies over hetzelfde tekstonderwerp. De linker tekst bevat een historisch correcte targetzin, de rechter tekst bevat een historisch incorrecte targetzin.

*Tabel 1: Voorbeeldtekst met een correcte en een incorrecte targetzin*

|  |  |
| --- | --- |
| **Targetzin correct** | **Targetzin incorrect** |
| Lance Armstrong is een internationaal bekende atleet. | Lance Armstrong is een internationaal bekende  atleet. |
| Hij heeft meegedaan aan vele mondiale wielrencompetities op het hoogste niveau. | Hij heeft meegedaan aan vele mondiale wielrencompetities op het hoogste niveau. |
| Tijdens zijn carrière werd hij achtervolgd door beschuldigingen van dopinggebruik. | Tijdens zijn carrière werd hij achtervolgd door beschuldigingen van dopinggebruik. |
| Lance heeft dit vaak afgedaan als een heksenjacht en hield vol dat hij onschuldig was. | Lance heeft dit vaak afgedaan als een heksenjacht en hield vol dat hij onschuldig was. |
| Hij noemde zichzelf vaak ‘de meest geteste atleet ter wereld’. | Hij noemde zichzelf vaak ‘de meest geteste atleet ter wereld’. |
| Hiermee doelde hij op de talloze doping testen die bij hem werden afgenomen. | Hiermee doelde hij op de talloze doping testen die bij hem werden afgenomen. |
| Vele onderzoeken naar bewijs voor deze beschuldigingen liepen op niets uit. | Vele onderzoeken naar bewijs voor deze beschuldigingen liepen op niets uit. |
| **Targetzin:** Lance Armstrong werd schuldig bevonden aan dopinggebruik. | **Targetzin:** Lance Armstrong werd niet schuldig bevonden aan dopinggebruik. |
| Hij is tijdens zijn actieve wielercarrière betrokken geweest bij de Livestrong foundation. | Hij is tijdens zijn actieve wielercarrière betrokken geweest bij de Livestrong foundation. |
| Deze organisatie zet zich in voor kankerpatiënten en mensen die kanker overleefd hebben. | Deze organisatie zet zich in voor kankerpatiënten en mensen die kanker overleefd hebben. |

Meetinstrumenten

**Leestaak**

De leestaak opende met een korte instructie. Na de instructie volgde één voorbeeldtekst, waarna de participanten begonnen met de echte taak. Participanten kregen de teksten één voor één te lezen. De teksten werden zin voor zin aangeboden en participanten moesten zelf doorklikken naar de volgende zin. Halverwege de leestaak konden participanten pauze nemen. Tijdens de leestaak is de leestijd per zin bijgehouden. De leestijd werd gemeten vanaf het moment dat een zin verscheen tot het moment dat de participant doorklikte naar de volgende zin.

**Geheugentaak**

Tijdens de geheugentaak kregen participanten zinnen te zien en moesten zij voor iedere zin aangeven of zij deze zin wel of niet hadden gelezen tijdens de leestaak. Daarna moesten zij op een VAS-schaal van 0-100 (helemaal niet zeker - heel zeker) aangeven hoe zeker zij waren van hun antwoord. In totaal bestond de geheugentaak uit 128 zinnen.

Procedure

Participanten zijn benaderd via sociale media. Zij werden slechts oppervlakkig geïnformeerd over het onderzoek. Hun werd verteld dat ze mee zouden doen aan een experiment over tekstbegrip en dat ze achteraf gedebrieft zouden worden over het precieze doel van het onderzoek. Voorafgaand aan deelname aan het onderzoek werden de participanten ook geïnformeerd over hun rechten tijdens het onderzoek en hebben zij een informed consent ondertekend.

Het experiment bestond uit twee sessies, die plaatsvonden op twee achtereenvolgende dagen. Op de eerste dag maakten de participanten de leestaak, die zij vanaf hun eigen computer konden uitvoeren. De participanten kwamen bij deze leestaak door op een link te klikken die naar hen was verstuurd via e-mail. Op de internetpagina waar zij via de link terecht kwamen moesten zij een aantal algemene gegevens invullen, zoals het proefpersoonnummer (die was bijgevoegd in de informatieve e-mail) en de laatste vier cijfers van hun telefoonnummer. Ook werd op deze pagina het opleidingsniveau en de leeftijd ingevuld. Vanuit deze pagina konden participanten vervolgens de leestaak downloaden. De leestaak werd uitgevoerd in het computerprogramma E-prime Go. In dit programma moest nogmaals het proefpersoonnummer en de laatste vier cijfers van het telefoonnummer worden ingevoerd. Daarna konden participanten beginnen met de leestaak.

Na afloop van de leestaak moesten participanten hun databestand uploaden naar de internetpagina waar zij hun gegevens hadden ingevuld. Hierna volgde nog een korte afsluitende vragenlijst, die onder andere bevroeg of participanten moeilijkheden hadden ondervonden tijdens het uitvoeren van de taak. In totaal duurde de leestaak ongeveer dertig minuten.

Op dag twee van het experiment maakten de participanten de geheugentaak. Ook deze taak konden zij vanuit huis uitvoeren. Via een tweede link in de naar hen verstuurde e-mail kwamen zij bij de geheugentaak terecht. De link leidde wederom naar een internetpagina waar dezelfde gegevens dienden te worden ingevuld als bij de leestaak. Vervolgens downloadden de participanten de geheugentaak, die -net als de leestaak- via E-prime Go werd afgenomen. Na het invullen van het proefpersoonnummer en de laatste vier cijfers van het telefoonnummer konden de participanten beginnen met de geheugentaak.

Na afronding van de geheugentaak moesten de participanten hun databestanden weer uploaden naar de internetpagina. Daarna volgde dezelfde korte vragenlijst die eventuele moeilijkheden tijdens het uitvoeren van de taak bevroeg. De participanten waren in totaal ongeveer twintig minuten bezig met deze tweede sessie van het experiment.

Na afloop van het tweede deel van het experiment kregen participanten een link naar een debriefing over het onderzoek. Daar konden zij het ware doel van het experiment lezen. In het geval van eventuele vragen werden zij doorverwezen naar de proefleider.

Data-analyse

De data die is verzameld (leestijd en score op de geheugentest) is geanalyseerd met behulp van het statistische computerprogramma SPSS Statistics (versie 26).

De eerste en tweede hypothese zijn getest aan de hand van een gepaarde t-toets. Dit is de meest passende statistische toets aangezien er bij beide hypotheses sprake is van één categorische predictor met twee categorieën en één numerieke uitkomstmaat. Daarnaast is het zo dat de gepaarde t-toets gemiddelden van groepen met elkaar kan vergelijken, wat nodig was voor deze eerste twee hypotheses. Er gelden twee assumpties voor de gepaarde t-toets. De eerste assumptie is dat er sprake moet zijn van een normale verdeling. Deze assumptie kan getoetst worden aan de hand van beschrijvende statistiek (zoals de Kolmogorov-Smirnov test). Daarnaast moet er sprake zijn van een aselecte steekproef. Er is in het geval van dit onderzoek sprake van een aselecte steekproef, omdat de correcte en incorrecte targetzinnen willekeurig zijn verdeeld onder de participanten.

Bij de eerste hypothese was de onafhankelijke variabele de correctheid van de targetzin die als predictor is gebruikt voor de afhankelijke variabele leestijd per zin (op de targetzin). Bij de tweede hypothese is dezelfde onafhankelijke variabele gebruikt, maar dan als predictor voor de score op de geheugentest (het percentage correct beantwoorde vragen).

De relatie tussen de grootte van het inconsistency effect en de mate van onthouden is getest aan de hand van een Repeated Measures ANCOVA. De verwachting was dat de afhankelijke variabele (de score op de geheugentest) afhankelijk is van zowel een onafhankelijke variabele (correctheid van de targetzin) en een covariaat (de grootte van het inconsistency effect). Een Repeated Measures ANCOVA is de meest passende manier om dit te toetsen.

Om de assumpties van de Repeated Measures ANCOVA te toetsen moest eerst een korte analyse worden uitgevoerd. Er moest gecontroleerd worden of de populaties normaal zijn verdeeld, met behulp van beschrijvende statistiek (zoals de Kolmogorov-Smirnov test). De tweede assumptie is dat de metingen onafhankelijk moeten zijn. Dit is het geval, omdat er allemaal aparte datapunten zijn die apart van elkaar getoetst kunnen worden.

De onafhankelijke variabele bij deze Repeated Measures ANCOVA is zoals genoemd de correctheid van de targetzin. De covariaat die is toegevoegd is de grootte van het inconsistency effect. Dit is een numerieke covariaat die is berekend door de leestijd op de correcte targetzin van de leestijd op de incorrecte targetzin af te trekken. De afhankelijke variabele is de score op de geheugentest.

# 5. Resultaten

# Er zijn verschillende analyses uitgevoerd om de data te doorgronden. Zo zijn de eerste twee hypotheses getoetst door middel van een gepaarde t-toets en de verwachte relatie tussen online verwerking en offline mentale representatie met een Repeated Measures ANOVA. Voordat deze analyses werden uitgevoerd is eerst gekeken naar de beschrijvende statistiek over de data. In tabel 2 zijn de resultaten daarvan te zien.

# 

*Tabel 2: beschrijvende data*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Leestijd (in ms)** | **Goed onthouden** | **Inconsistency effect (in ms)** |
| **Algemeen** |  | *M =* 74% (*SD* = 43.8%) | *M* = 258.68 (*SD* = 608.49) |
| **Correcte targetzin** | *M* = 2376.98 (*SD* = 627.60) | *M* = 75.5% (*SD* = 15.3%) |  |
| **Incorrecte targetzin** | *M* = 2635.66 (*SD* = 659.83) | *M* = 51% (*SD* = 20.4%) |  |

# Tijdens het analyseren van de data bleek dat vier participanten buiten het inclusiecriterium van leeftijd vielen. Het inclusiecriterium was een leeftijd van 18-35 jaar oud, in verband met de *norming study* die was uitgevoerd onder deze leeftijdsgroep (Van Moort et al., 2018). Er moest dus allereerst gekeken worden of deze participanten geëxcludeerd moesten worden. Dit is gedaan door het gemiddelde van de score op de voorkennis vragen over alle participanten te berekenen, alsmede de standaarddeviatie (*M* = 57.53, *SD* = 9.93). De participanten zouden geëxcludeerd moeten worden als hun score op de voorkennis vragen drie standaarddeviaties van het gemiddelde zouden zitten. Dit bleek voor geen van de vier participanten het geval te zijn. Deze participanten zijn dus meegenomen in de verdere analyses.

De assumpties van de statistische toetsen zijn bekeken en de data bleek bij benadering normaal verdeeld te zijn. Er is dus voldaan aan de assumpties.

Online verwerking

Om te onderzoeken of participanten langer doen over het lezen van incorrecte informatie dan over het lezen van correcte informatie (Hypothese 1) is een gepaarde t-toets uitgevoerd. Uit de resultaten van deze t-toets bleek dat participanten langer doen over het lezen van incorrecte targetzinnen (*M* = 2635.66, *SD* = 659.83) dan over het lezen van correcte targetzinnen (*M* = 2376.98, *SD* = 627.60), t(75) = 3.706, p < .000). Hypothese 1 wordt hiermee ondersteund.

Offline representatie

Om te onderzoeken of incorrecte informatie minder goed wordt onthouden dan correcte informatie is wederom een gepaarde t-toets uitgevoerd. De resultaten van de t-toets lieten zien dat lezers de incorrecte targetzinnen (*M* = .510, *SD* = .204) minder goed onthouden dan de correcte targetzinnen (*M* = .755, *SD* = .153), (t(75) = . -8.088, p < .000). Hiermee wordt hypothese 2 ondersteund.

Relatie online-offline

Voor het toetsen van de relatie tussen de grootte van het inconsistency effect en de score op de geheugentaak is een Repeated Measures ANCOVA uitgevoerd met de score op de geheugentest als afhankelijke variabele, de correctheid van de targetzin als onafhankelijke variabele en grootte van het inconsistency effect als covariaat. Allereerst laten de resultaten een hoofdeffect zien voor correctheid van de targetzin. De correcte targetzinnen (*M* = .755, *SD* = .153) werden beter onthouden dan de incorrecte targetzinnen (*M* = .510, *SD* = .204), (F(1, 74) = 58.129, p < .000). Dit is overeenkomstig met de resultaten van de t-toets die is uitgevoerd om hypothese 2 te toetsen. Er bleek geen hoofdeffect voor de grootte van het inconsistency effect te zijn (F(1,74) = .048, p = .828). Daarnaast is er ook geen interactie-effect gevonden tussen de grootte van het inconsistency effect en de mate van onthouden (F(1,74) = .318, p = .574). De verwachte relatie is dus niet gevonden.

# 6. Discussie

# Het doel van dit onderzoek was te onderzoeken hoe de correctheid van gelezen informatie van invloed is op informatieverwerking en het onthouden van informatie. Dit is gedaan door een experiment uit te voeren waarin leestijd en het percentage correct onthouden informatie werden gebruikt als variabelen om hier inzicht in te krijgen. In lijn met de verwachtingen lieten de resultaten een inconsistency effect zien: leestijden op incorrecte targetzinnen waren langer dan de leestijden op de correcte targetzinnen. Ook lieten de resultaten zien dat incorrecte targetzinnen minder goed werden onthouden dan correcte targetzinnen. Een relatie tussen de grootte van het inconsistency effect en het onthouden van informatie werd echter niet gevonden.

# De langere leestijden op de incorrecte targetzinnen ondersteunen hypothese 1 en bieden ondersteuning voor het inconsistency effect. Veel eerdere onderzoeken vonden al bewijs voor het inconsistency effect (o.a. Cook & O’Brien, 2014; Van Moort et al., 2018; Van Moort et al., n.d.). Dit onderzoek kan zich daarbij aansluiten. Dit ondersteunt dus het idee dat lezers tijdens het lezen continu bezig zijn met het valideren van de informatie die zij lezen. Daarnaast kunnen de langere leestijden op de incorrecte informatie gezien worden als ondersteuning voor het RI-Val model. Een van de assumpties van dit model is dat de lezer pas doorleest als er op basis van drie processen -activatie, integratie en validatie- een voldoende coherent situatiemodel is gevormd. Incorrecte informatie zou ervoor zorgen dat deze processen langer duren, omdat het de lezer dan meer moeite kost om een acceptabel situatiemodel te vormen. (O’Brien & Cook, 2016a; O’Brien & Cook, 2016b). De resultaten van dit onderzoek sluiten aan bij dit idee. Ook zijn er andere verklaringen mogelijk voor de langere leestijd op incorrecte informatie. Zo kan het zijn dat de gelezen informatie minder snel gerelateerd kan worden aan de eigen kennis van de lezer, omdat de incorrecte gelezen informatie en de eigen kennis van de lezer niet genoeg op elkaar lijken. Onderzoek toonde eerder al aan dat de mate waarin de gelezen informatie lijkt op de eigen kennis van de lezer van invloed is op de grootte van het inconsistency effect (Cook & O’Brien, 2014). De resultaten van het huidige onderzoek kunnen geen uitsluitsel geven over de precieze oorzaak van de langere leestijd op incorrecte informatie, dus vervolgonderzoek zou hier aandacht aan moeten besteden.

Het feit dat de incorrecte informatie minder goed werd onthouden dan de correcte informatie ondersteunt hypothese 2. Eerder onderzoek wees al uit dat incorrecte informatie minder goed in het situatiemodel wordt opgenomen, waardoor het minder goed wordt onthouden. (Schroeder et al., 2008; Van Moort et al., 2020, in: Van Moort et al., n.d.; Radvansky & Zacks, 1991). De resultaten van dit onderzoek liggen in lijn met de resultaten uit deze eerdere onderzoeken. Het minder goed onthouden van de incorrecte informatie kan verschillende verklaringen hebben. Zo zou het kunnen zijn dat er tijdens het lezen van incorrecte informatie meer interferentie is met de eigen kennis van de lezer over het onderwerp, waardoor de gelezen informatie minder goed in het situatiemodel wordt geïntegreerd. Daarnaast is het mogelijk dat de moeilijkheid zit in het ophalen van de informatie uit het geheugen tijdens de geheugentaak. De gelezen informatie kan overschaduwd worden door kennis die de lezer zelf al heeft over het onderwerp, waardoor de incorrecte gelezen informatie tijdens de geheugentaak meer naar de achtergrond verdwijnt (Van Moort et al., n.d.). Aan de mogelijke oorzaken van het minder goed onthouden van incorrecte informatie is in dit onderzoek geen aandacht besteed, dus vervolgonderzoek is nodig om hier meer inzicht in te krijgen.

De verwachte relatie tussen de grootte van het inconsistency effect en het onthouden van informatie, die niet gevonden werd, biedt een interessante insteek voor vervolgonderzoek. Er is reden om een relatie te verwachten tussen deze twee variabelen, zoals reeds beschreven in het theoretisch kader. Wat men doet tijdens het lezen is namelijk zonder twijfel verbonden aan wat men kan onthouden van de gelezen informatie (Van Moort et al., n.d.). Dit komt echter uit de resultaten van dit onderzoek niet naar voren. Omdat het niet aannemelijk is dat de online verwerking en de offline mentale representatie niet aan elkaar gerelateerd zijn, is er waarschijnlijk een meer methodologische verklaring voor het feit dat de relatie niet gevonden is.

Allereerst zou het kunnen zijn dat de grootte van het inconsistency effect een dusdanig kleine voorspeller is dat het effect in dit onderzoek niet naar voren kwam. Er zijn namelijk tal van andere voorspellers denkbaar voor het onthouden van informatie, zoals leesinstructie (Van den Broek et al., 2001; Van Moort et al., n.d.), interesse in het onderwerp waarover gelezen wordt (Schiefele, 1996) en andere lezerskenmerken zoals algemene cognitieve capaciteit en voorkennis over het onderwerp (Shing & Brod, 2016; Umanath & Marsh, 2014; Schneider, Körkel & Weinert, 1990). Vervolgonderzoek zou zich meer moeten verdiepen in deze verschillende voorspellers van het onthouden van informatie, om hier meer duidelijkheid over te verschaffen.

Daarnaast kan het zo zijn dat de gebruikte onderzoeksopzet niet geschikt is om een relatie tussen online verwerking en offline representatie te onderzoeken. Wellicht dat een andere voorspeller, zoals een meting van oogbewegingen, meer geschikt is.

Er zijn enkele aandachtspunten die van belang kunnen zijn voor toekomstige studies. Allereerst is het zo dat enkele proefpersonen tijdens het onderzoek hebben aangegeven moeilijkheden te hebben ondervonden tijdens het werken met het programma waarin het experiment werd afgenomen. Dit kan mogelijk de leestijden en/of de resultaten op de geheugentaak hebben beïnvloed. Het is dus van belang dat toekomstige studies hier aandacht aan besteden, om eventuele moeilijkheden te ondervangen. Daarnaast moest dit onderzoek volledig digitaal worden afgenomen, waardoor minder zicht was op hoe de participanten het onderzoek uitvoerden. Het kan zijn dat participanten hierdoor sneller waren afgeleid, wat leestijden kan hebben beïnvloed. Toekomstige studies zouden hier rekening mee moeten houden.

# 7. Conclusie

# De conclusie die getrokken kan worden uit dit onderzoek is dat de correctheid van informatie van invloed is op zowel online tekstverwerking als op de offline mentale representatie van de gelezen tekst. Uit dit onderzoek kan opgemaakt worden dat lezers meer moeite hebben met het verwerken van incorrecte informatie dan met het verwerken van correcte informatie. Ook is gebleken dat lezers meer problemen hebben met het vormen en/of updaten van hun situatiemodel als zij incorrecte informatie hebben gelezen. Hoewel het wel verwacht werd, is er geen verband gevonden tussen de online verwerking en de offline mentale representatie van tekst. Dit onderzoek heeft inzicht geboden in hoe lezers omgaan met incorrecte informatie, wat kan helpen bij het opstellen van adviezen over misinformatie.

# Literatuur

Albrecht, J. E., & O’Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(5), 1061–1070. https://doi.org/10.1037/0278-7393.19.5.1061

# Boeije, H. R. (2016). *Analyseren in kwalitatief onderzoek*. Boom Lemma.

Chomsky, A. N. (2000, 10 december). *Propaganda and indoctrination*. chomsky.info. https://chomsky.info/20001210/

Cook, A. E., & O’Brien, E. J. (2014). Knowledge activation, integration, and validation during narrative text comprehension. *Discourse Processes*, *51*(1–2), 26–49. https://doi.org/10.1080/0163853x.2013.855107

Gelfert, A. (2018). Fake news: A definition. *Informal Logic, 38* (1), 84–117.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. https://doi.org/10.1037/0033-295x.95.2.163

Lombardi, D., Sinatra, G. M., & Nussbaum, E. M. (2013). Plausibility reappraisals and shifts in middle school students’ climate change conceptions. *Learning and Instruction*, *27*, 50–62. https://doi.org/10.1016/j.learninstruc.2013.03.001

Matsuki, K., Chow, T., Hare, M., Elman, J. L., Scheepers, C., & McRae, K. (2011). Event-based plausibility immediately influences on-line language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 913–934. <https://doi.org/10.1037/a0022964>

Myers, J. L., O’Brien, E. J., Albrecht, J. E., & Mason, R. A. (1994). Maintaining global coherence during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 876–886. https://doi.org/10.1037/0278-7393.20.4.876

O’Brien, E.J, Albrecht, J.E., Rizzella, M.L., & Halleran, J.G. (1998). Updating a situation model: A memory-based tekst processing view. *Journal of Experimental Psychology, 24*(5), 1200-1210.

O'Brien, E. J., & Cook, A. E. (2016a). Coherence threshold and the continuity of processing: The RI-Val model of comprehension. *Discourse Processes, 53*(5-6), 326-338. https://doi.org/10.1080/0163853

O'Brien, E. J., & Cook, A. E. (2016b). Separating the activation, integration, and validation components of reading. In H. R. Brian (Ed.), Psychology of Learning and Motivation(Vol. Volume 65, pp. 249-276): Academic Press.

Radvansky, G. A., & Zacks, R. T. (1991). Mental models and the fan effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(5), 940–953. <https://doi.org/10.1037/0278-7393.17.5.940>

Richter, T. (2015). Validation and comprehension of text information: Two sides of the same coin. *Discourse Processes*, *52*(5–6), 337–355. <https://doi.org/10.1080/0163853x.2015.1025665>

Schiefele, U. (1996). Topic interest, text representation, and quality of experience. *Contemporary Educational Psychology*, *21*(1), 3–18. <https://doi.org/10.1006/ceps.1996.0002>

Schneider W., Körkel J., Weinert F.E. (1990) Expert knowledge, general abilities, and text processing. In: Schneider W., Weinert F.E. (eds) Interactions Among Aptitudes, Strategies, and Knowledge in Cognitive Performance. Springer, New York, NY. <https://doi.org/10.1007/978-1-4612-3268-1_17>

Schroeder, S., Richter, T., & Hoever, I. (2008). Getting a picture that is both accurate and stable: Situation models and epistemic validation. *Journal of Memory and Language, 59*(3), 237–255. doi:10.1016/j.jml.2008.05.001

Shing, Y. L., & Brod, G. (2016). Effects of prior knowledge on memory: Implications for education. *Mind, Brain, and Education*, *10*(3), 153–161. <https://doi.org/10.1111/mbe.12110>

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media. *ACM SIGKDD Explorations Newsletter*, *19*(1), 22–36. https://doi.org/10.1145/3137597.3137600

Singer, M. (2006). Verification of text ideas during reading. *Journal of Memory and Language, 54*, 574-591.

Singer, M. (2013). Validation in reading comprehension. *Current Directions in Psychological Science, 22* (5), 361-366.

Singer, M. (2019). Challenges in processes of validation and comprehension. *Discourse processes, 56*(5-6), 465-483.

Umanath, S., & Marsh, E. J. (2014). Understanding how prior knowledge influences memory in older adults. *Perspectives on Psychological Science*, *9*(4), 408–426. https://doi.org/10.1177/1745691614535933

Van den Broek, P., Lorch, R. F., Linderholm, T., & Gustafson, M. (2001). The effects of readers’ goals on inference generation and memory for texts. *Memory & Cognition*, *29*(8), 1081–1087. <https://doi.org/10.3758/bf03206376>

Van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse Processes*, *39*(2–3), 299–316. https://doi.org/10.1080/0163853x.2005.9651685

Van Moort, M.L., Koornneef, A., & Van den Broek, P.W. (2018). Validation: Knowledge- and text-based monitoring during reading. *Discourse processes, 55*(5-6), 480-496

Van Moort, M.L., Koornneef, A., & Van den Broek, P.W. (2020). Differentiating tekst-based and knowledge-based validation processes during reading: Evidence from eye movements. *Discourse processes*, 1-20

Van Moort, M.L., Koornreef, A., & Van den Broek, P.W. (n.d.). Purposeful validation: Are validation processes and the construction of a mental representation influenced by reading goal? Manuscript in voorbereiding.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, *6*(5), 292–297. https://doi.org/10.1111/j.1467-9280.1995.tb00513.x

Zwaan, R. A., Radvansky, G. A., Hilliard, A. E., & Curiel, J. M. (1998). Constructing multidimensional situation models during reading. *Scientific Studies of Reading*, *2*(3), 199–220. <https://doi.org/10.1207/s1532799xssr0203_2>