

De relatie tussen het vooraf lezen van teksten en het beantwoorden van begripsvragen

Twee experimenten onder 5-vwo-leerlingen

P. Rooijackers, G. van Silfhout, U. Schuurs, en H. van den Bergh

Samenvatting

In Nederland traint men vaak tekstbegrip met gebruikmaking van teksten met vragen: leerlingen bestuderen vooraf een tekst en beantwoorden vervolgens tekstafhankelijke vragen, waarbij de tekst raadpleegbaar blijft. Daarbij wordt doorgaans verondersteld dat wie een tekst vooraf intensief bestudeert, de vragen vaker correct beantwoordt.

In deze studie is de relatie tussen initiële tekstbestudering en antwoordscore nader onderzocht, onder 5-vwo-leerlingen, bij twee zakelijke teksten, met begripsvragen, via twee experimenten. Experiment I onderzocht de mate waarin de *beschikbaarheid* van de tekst tijdens vraagbeantwoording de antwoordscore beïnvloedt, nadat deelnemers ($N = 174$) vooraf teksten intensief hadden bestudeerd. Experiment II onderzocht in hoeverre de *intensiteit* van vooraf-tekstbestudering de antwoordscore beïnvloedt, als deelnemers ($N = 398$) bij de vragen de tekst kunnen raadplegen. Resultaten: in experiment I beantwoorden leerlingen die bij de vragen de tekst wél kunnen raadplegen, de vragen beduidend vaker correct dan leerlingen die deze niet kunnen raadplegen. In experiment II kan niet worden aangetoond dat leerlingen die de tekst vooraf intensief bestuderen, hoger scoren dan leerlingen die de tekst vooraf minder of niet-intensief bestudeerden. Het antwoordproces wordt kennelijk niet zozeer beïnvloed door intensiteit van vooraf lezen, als wel door herlezen tijdens het antwoordproces. De vraag is daarmee in hoeverre deze taak leerlingen ertoe traint zelfstandig een tekst te bestuderen.

Kernwoorden: tekstbegrip, vwo-leerlingen, tekst met vragen, schoolvak Nederlands, zakelijke teksten

1 Inleiding

1.1 Aanleiding

In het schoolvak Nederlands wordt in de leesles doorgaans gewerkt met teksten met vragen: leerlingen bestuderen een tekst en beantwoorden vervolgens begripsvragen, waarbij de tekst raadpleegbaar blijft (Linthorst & De Gloppe, 2015; SLO, 2017). Op de nogal dominante positie van deze taak in de leesles Nederlands is de afgelopen decennia vaak door neerlandici en onderwijskundigen in Nederland en Vlaanderen kritiek geuit (bijv. Hoogveen & Bonset, 1998; Hoogveen, 2018; Van den Branden, 2003; Van de Wetering & Groenendijk, 2015). Het markantst uitte zich in Nederland deze kritiek rondom het Centraal Examen Nederlands havo/vwo, dat vooral gebruikmaakt van vragen bij teksten en dat volgens critici niet zozeer de leesvaardigheid van leerlingen zou toetsen als wel ‘vraagvaardigheid’, de mate waarin ‘je goed bent in examen doen’ (Grezel, 2014).

Een belangrijke veronderstelling achter de didactiek bij deze taak is doorgaans dat wie de tekst vooraf intensief bestudeert, de vragen vervolgens vaker correct beantwoordt – een veronderstelling die wortelt in opvattingen over tekstbegrip in de neerlandistiek uit de jaren tachtig van de vorige eeuw (vgl. Drop, 1983; Wesdorp, 1981; zie Rooijackers, Van Silfhout & Van den Bergh, 2021a). Deze ‘simpele’ relatie tussen de initiële tekstbestudering en vraagbeantwoording vinden we impliciet terug in oefenboeken voor het Centraal Schriftelijk Examen Nederlands (bijv. ExamenOverzicht, 2020) en in een aantal Nederlandse leergangen voor het schoolvak Nederlands. Zo vraagt de bovenbouweditie vwo van de in Nederland veelgebruikte leergang *Nieuw Nederlands* (6^e editie) van leerlingen vaak dat ze eerst een tekst via een uitgebreid “stappenplan tekstanalyse” door-

werken, voordat ze de vragen over de tekst gaan beantwoorden (Frank et al., 2019).

De empirische onderbouwing bij deze veronderstelling is echter schaars en onvolledig. Wanneer een onderscheid gemaakt wordt tussen wel en niet vooraf lezen van de tekst, dan lijkt wél vooraf lezen bij deze taak inderdaad te leiden tot hogere resultaten dan direct starten met de vragen (Cerdán, Vidal-Abarca, Martínez, Gilabert, & Gil, 2009; De Milliano & Van Gelderen, 2020; Salmerón et al., 2015). Maar daarmee is de vraag nog niet beantwoord of de *intensiteit* van de voorafbestudering ook van invloed is op de antwoordscore, zoals vaak verondersteld. Studies hiernaar ontbreken naar ons weten.

Deze veronderstelling spreekt in elk geval niet vanzelf. Vooral tijdens het antwoordproces blijken leerlingen een mentale representatie van (delen van) de tekst op te bouwen om antwoord te kunnen geven, en niet zozeer tijdens het vooraf lezen (Rooijackers, Van Silfhout, Schuurs, & Van den Bergh, 2020; Schroeder, 2011): doordat leerlingen tijdens de vraagbeantwoording de tekst nog kunnen raadplegen, kunnen ze ontbrekend of inadequaat tekstbegrip repareren en hun antwoord controleren. Een intensieve voorafbestudering heeft voor leerlingen dan op het oog weinig meerwaarde. Die situatie verandert als leerlingen bij het beantwoorden van de vragen *niet* meer over de tekst kunnen beschikken: in zo'n taaksetting blijken hogere-ordedenkprocessen tijdens het vooraf lezen, zoals conceptuele integratie en coherente representatie van de tekstinhoud, grotendeels verschillen in antwoordscore van leerlingen te kunnen verklaren (Schroeder, 2011; vgl. Ferrer, Vidal-Abarca, Serrano, & Gilabert, 2017; Ozuru, Best, Bell, Witherspoon, & McNamara, 2007). Doordat leerlingen niet langer ontbrekend of onvoldoende tekstbegrip tijdens de antwoordfase kunnen repareren, zijn ze gedwongen hun antwoord grotendeels te baseren op hun vooraf-opgebouwde mentale representatie van de tekst, en daarmee krijgt een intensieve vooraf-tekstbestudering in deze taaksetting een duidelijk meerwaarde. Deze onderzoeksresultaten houden echter niet in dat een intensieve vooraf-tekstbestudering tot meer correcte ant-

woorden leidt als leerlingen bij het beantwoorden van de vragen nog wél over de tekst kunnen beschikken.

In deze studie willen we daarom nagaan in hoeverre een intensieve vooraf-tekstbestudering tot een hogere tekstbegripsscore op de vragen leidt, zoals vaak bij het schoolvak Nederlands verondersteld. Deze veronderstelling vormt immers een belangrijk argument om de 'tekst met vragen' veelvuldig als oefenmateriaal in de les in te zetten. Om deze vraag te kunnen beantwoorden, hebben we in deze studie getracht de relatie tussen initieel leesproces en antwoordscore nader bloot te leggen. Dat doen we door twee factoren te manipuleren: a) de beschikbaarheid van de tekst bij het beantwoorden van de tekstafhankelijke vragen, en b) de intensiteit waarmee lezers vooraf een tekst bestuderen.

1.2 Theorie: tekstbegrip en intentionaliteit

Allereerst: wat is tekstbegrip? Een aantal breed geaccepteerde theoretische modellen van vóór de millenniumwisseling voorspellen en verklaren inmiddels de cognitieve processen tijdens tekstbegrip, zoals de Constructie-Integratie-Theorie (CI-Theorie: Kintsch & Van Dijk, 1978; Kintsch, 1998) en het Landschapsmodel (Van den Broek, Young, Tzeng, & Linderholm, 1999). Een belangrijk kernconcept hierin is het onderscheid tussen tekstbasis en situatiemodel. Een tekst wordt in deze theorieën beschouwd als een netwerk van hiërarchisch samenhangende, betekenisvolle beweringen. Hierin zijn lezers steeds op zoek naar samenhang, in eerste instantie tussen woorden en zinnen (microstructuur) en in tweede instanties tussen alinea's (macrostructuur). Al lezend bouwen ze zo gaandeweg een 'tekstbasis' op, een vrij letterlijke mentale representatie van de tekst, en een 'situatiemodel', een mentaal model waarin ze de tekstuele wereld integreren in hun eigen voorkennis, opvattingen en ideeën. Pas op dit laatste niveau ontstaat diepere betekenis: de lezer heeft zich de tekstinhoud eigengemaakt.

In deze twintigste-eeuwse theoretische modellen bleef nog goeddeels onuitgewerkt hoe specifieke leesdoelen in studie- en arbeidscontexten leesgedrag kunnen beïnvloeden, zoals in onze studie bij het

beantwoorden van begripsvragen over een tekst. Sinds de millenniumwisseling is in internationaal leesonderzoek steeds meer deze intentionaliteit achter leesgedrag in kaart gebracht, de situationele en functionele inbedding van tekstbegrip. Uiteenlopende leesdoelen en leessituaties blijken bij lezers tot deels afwijkende begripsprocessen en een afwijkende inzet van vaardigheden en strategieën te leiden (bijv. McCrudden, Magliano, & Schraw, 2011; Van den Broek & Helder, 2017; Yeari, Van den Broek & Oudega, 2015). Al naargelang het doel dat lezers zichzelf stellen, bouwen ze een tekstuele representatie op die kwalitatief en kwantitatief afdoende lijkt om aan het doel van de taak te voldoen (bijv. Van den Broek, Ridsen, & Husebye-Hartmann, 1995). Geruggenstund door deze studies naar de intentionaliteit achter leesgedrag wil het recente RESOLV-model (Britt, Rouet, & Durik, 2018; zie ook Rouet, Britt & Durik, 2017) ook het antwoordgedrag bij leestaken verklaren en voorspellen, een dimensie van leesgedrag die in de modellen van vóór 2000 nog grotendeels onbesproken bleef.

Het RESOLV-model beschouwt lezen voor een taak (*Reading*) als het oplossen van een probleem (*Solving*). Bij een taak over een of meerdere teksten vertrekken lezers daarom vanuit een statisch *contextmodel*, dat zaken verdisconteert als de taakinterpretatie, de status van de ‘opdrachtgever’, de hulpmiddelen en belemmeringen bij de taak, evenals een zelfbeoordeling bij deze taak door de lezers. Vanuit dit contextmodel stellen lezers vervolgens een dynamisch *taakmodel* op: taakspecifieke doelen en scenario’s, waarin ze interpreteren wat ze moeten doen om de doelen te bereiken, en welke procedures ze daarbij moeten hanteren. In tegenstelling tot het contextmodel, dat in principe na het starten met de taak vastligt, is dit taakmodel dynamisch: lezers kunnen het gaandeweg de taak aanpassen en bijstellen. Afhankelijk van dit taakmodel bouwen ze van de tekst(en) een passende mentale representatie op en zetten ze zich tot een eventuele vervolftaak.

Het RESOLV-model staat ons in deze studie toe verwachtingen te formuleren over lees- en antwoordgedrag van lezers bij een

tekst met vragen. Twee van de vijf basishypothesen bij dit model verdienen vooraf nog nadere uitwerking. Volgens de ‘*minimal task-elaboration hypothesis*’ plannen lezers vooraf niet uitvoerig hoe ze een taak moeten aanpakken. Ze vertrekken weliswaar in hun taakmodel vanuit een hogere-orde-doel, maar ze werken vooraf enkel in zeer algemene lijnen uit hoe ze een taak willen aanpakken; al naargelang de uitvoeringsproblemen die ze ervaren, passen ze min of meer noodgedwongen gaandeweg hun werkwijze aan. Daarnaast: volgens de ‘*parsing the context hypothesis*’ beslissen lezers vanuit hun contextmodel vooraf niet alleen of ze de taak zullen uitvoeren, maar ook de mate waarin ze dit doe. Zo zullen lezers zich bij een zwaarwegende leestoets tot een diepere, nauwkeurigere en meer gelaagde werkwijze zetten dan bij een ‘alledaagse’ leestaak van gelijke strekking. In de volgende paragraaf werken we deze twee hypothesen nader uit door ze toe te passen op de taak ‘tekst met vragen’.

1.3 Taakmodel bij een tekst met vragen

Vragen stellen bij een tekst is waarschijnlijk wel de bekendste internationale vorm van tekstbegripsmeting (Pearson & Hamm, 2005). Ook in Nederland is het een veel gebruikte toets in het basis- en voortgezet onderwijs, die ook vaak ingezet wordt om in de leesles tekstbegrip te trainen (bijv. Canton, Aler, Heemskerk, Van der Westen, & Willemsen, 2012; Bogaerds-Hazenberg, Evers-Vermeul, & Van den Bergh, [ingediend]). Als argument voor deze inzet in de leesles kan gelden dat het stellen van vragen bij een tekst een diepere verwerking van de tekstinhoud genereert (Rouet, Vidal-Abarca, Erboul, & Millogo, 2011; vgl. Hendrix & Hulshof, 1994). Bovendien staat de taak een zeer gerichte bevraging van tekstkenmerken en -elementen toe. Opvallend aan de courante taakinrichting is dat een leerling vooraf doorgaans een nogal generieke, tweeledige opdracht ontvangt: ‘Bestudeer de tekst goed en beantwoord daarna de vragen’.

Nemen we aan dat een vaardige lezer de tekst nog bij de vragen beschikbaar heeft én dat hij zich vanuit zijn contextmodel ten volle engageert met de taak (omdat het bijvoor-

beeld een toets betreft): welk taakmodel bouwt deze lezer dan bij bovenstaande opdracht hoogstwaarschijnlijk op? We stellen nu een hypothetisch scenario op, vertrekkend vanuit de ideeën van Britt, Rouet en Durik (2018), dat we daarna onderzoeksmatig nader zullen onderbouwen. Een vaardige lezer ervaart vanuit de *minimal task-elaboration hypothesis* het eerste doel achter de opdracht ('de tekst goed begrijpen') hoogstwaarschijnlijk als weinig concreet; het tweede doel (het 'correct beantwoorden van de vragen') is voor hem aanmerkelijk concreter en genereert bovendien een eindproduct dat naderhand via correctie meetbaar op kwaliteit is te beoordelen. Vooraf zal hij de tekst daarom tamelijk oppervlakkig doorlezen, alleen al omdat het hem wordt verzocht maar vooral omdat het hem kan helpen de vragen uiteindelijk beter te duiden en antwoord-relevante tekstelementen beter te lokaliseren. Bij het beantwoorden van de vragen zal hij zich dan, naargelang een vraag daartoe noodzaakt, tot een nauwkeuriger en dieper bestudering van relevante tekstelementen zetten.

Dit hypothetische taakmodel van een vaardiger lezer sluit aan bij de resultaten van de meeste onderzoeken naar lees- en antwoordgedrag bij teksten met vragen. Allereerst: de initiële bestudering van de tekst. Deze draagt inderdaad tot op enige hoogte bij aan een hogere score op de vragen (Cerdán et al., 2009; Salmerón et al., 2015). Leerlingen die de tekst vooraf hebben gelezen, bouwen vooraf een globaal tekstbeeld op, waardoor ze beter de vraag interpreteren en beter de vraag-relevante tekstpassages selecteren en interpreteren (Cataldo & Oakhill, 2000; O'Reilly et al., 2018; Payne & Reader, 2006; Salmerón et al., 2015).

Dan: het antwoordproces. Vooral hier blijken vaardige lezers zich te onderscheiden van minder vaardige lezers. Het beantwoorden van een tekstafhankelijke vraag dwingt een lezer namelijk tot een veeleisende, stapsgewijze zelfregulatie. Zo dient hij allereerst de vraag te begrijpen: hij moet de mentale representatie van de vraag in lijn brengen met de eerder opgebouwde representatie van de tekstinhoud. Dit kan minder vaardige lezers aanzienlijke problemen opleveren (Brunfaut

& McCray, 2015; Cerdán et al., 2019; Vidal-Abarca, Mañá & Gil, 2010). Vervolgens dient hij zijn 'opzoekbeslissing' te reguleren: raadpleegt hij de tekst wel of niet? Minder vaardige lezers lijken vaker hun eigen begrip van de tekst en van de vraag te overschatten dan vaardige lezers (Dunlosky & Rawson, 2005; Maki, 1998; Maki, Shields, Wheeler, & Zachilli, 2005). Besluit een leerling de tekst te raadplegen, dan dient hij vervolgens het zoekproces te reguleren: hij moet de relevante informatie vinden en adequaat verwerken. Minder vaardige lezers besteden gemiddeld minder tijd aan relevante informatie in de tekst dan vaardige lezers (Bax, 2013; Cataldo & Oakhill, 2000; Rooijackers et al., 2020). Ten slotte moet een lezer zijn antwoord reguleren en een antwoord formuleren of, bij een meerkeuzevraag, een antwoordalternatief selecteren (Cain & Oakhill, 2006; Traub, 1993). Het type vraag en de complexiteit van de vraag hebben op dit gehele proces een grote impact. Daardoor kan het onderscheid in antwoordgedrag tussen vaardige en minder vaardige lezers zich per vraag op zeer uiteenlopende wijze uiten: vaardige lezers lijken bijvoorbeeld niet per vraag gemiddeld minder tijd te besteden aan het antwoordproces dan minder vaardige lezers, ze kijken bijvoorbeeld korter naar de vraag en bouwen bij de tekstraadpleging meer controlestappen in (vgl. Bax, 2013; Brunfaut & McCray, 2015; Rooijackers et al., 2020).

Ten slotte: de antwoordscore – een hypothetische verwachting. In welke mate de antwoordscore de kwaliteit van de initiële tekstbestudering of die van het antwoordproces weerspiegelt, hangt hoogstwaarschijnlijk voor belangrijk deel af van de complexiteit en 'lokaliteit' van de gestelde vragen. Bij eenvoudige of meer globale vragen zal een vaardiger lezer het antwoord vrijwel direct vanuit zijn reeds opgebouwde tekstrepresentatie kunnen construeren, terwijl vragen over zeer lokale tekstelementen of complexere vragen hem eerder zullen noodzaken om tamelijk omstandig de tekst te raadplegen: zijn uiteindelijke antwoord weerspiegelt dan vooral de kwaliteit van dit 'uitgebouwde' antwoordproces. Deze verwachting aangaande de antwoordscore kan de nogal uiteenlopende

resultaten in internationale studies op dit vlak verklaren. Er is, zoals eerder besproken, een samenhang aantoonbaar tussen vooraf lezen en de antwoordscore (bijv. Cerdán et al., 2009). Tegelijk lijkt in een aantal studies het antwoordproces een sterk op zichzelf staand complex, waarin ‘opzoekvaardigheid naar aanleiding van de vraag’ doorslaggevend is: deze zoekbeslissingen kunnen soms duidelijk losstaan van het initiële tekstbegrip en ‘*test-management-strategies*’ betreffen (Cohen & Upton, 2006; deels: Ozuru, Rowe, O’Reilly & McNamara, 2008; O’Reilly et al., 2018; Rooijackers et al., 2020; Schroeder, 2011; Vidal-Abarca et al., 2010). Een enkele studie trekt zelfs de conclusie dat in een tekst met vragen vrijwel uitsluitend toetsvaardigheden worden aangesproken (Rupp, Ferne & Choi, 2006). Die conclusie wordt door andere studies weer sterk gerelativeerd: mentale processen die tijdens de ‘*response decision phase*’ spelen, lijken wel degelijk samen te hangen met het tekstbegrip dat leerlingen vooraf hebben opgebouwd (Gil, Martinez, & Vidal-Abarca, 2015; Ozuru et al., 2008; vgl. Ardoin, Binder, Zawoyski, Nimrocks, & Foster, 2019).

Maar wat als de tekst bij de vragen niet langer raadpleegbaar is? In het taakmodel van een vaardige lezer, zoals we dat hierboven beschreven en met onderzoeksresultaten onderbouwen, is de verwachting dat de tekst beschikbaar blijft bij de vragen, cruciaal: tijdens de antwoordfase weet hij immers dat hij onvolledig opgebouwd tekstbegrip altijd nog kan repareren, indien noodzakelijk. Weet hij vooraf dat hij bij de vragen de tekst niet langer beschikbaar zal hebben, dan ontstaat er een drastisch andere situatie.

We schetsen voor deze situatie weer een hypothetisch taakmodel. Wanneer vooraf bekend is dat de tekst bij de vragen niet langer raadpleegbaar is, dan rijst voor een lezer, volgens de ‘*parsing the context hypothesis*’, allereerst de vraag in hoeverre hij zich vanuit zijn contextmodel wil engageren met deze taak. Het doel ‘de tekst goed begrijpen’ is immers vaag en vergt van hem, al naargelang de lengte en complexiteit van de tekst, een mogelijk grote mentale inspanning. Vindt hij de taak voldoende relevant of hangt er voor

hem veel van de taak af, dan engageert hij zich en kan hij niet anders dan tijdens de tekstbestudering aan de instructie optimaal trachten te voldoen; de vragen zal hij vervolgens moeten beantwoorden vanuit zijn optimaal opgebouwde representatie van de tekstinhoud. In hoeverre hij de vragen goed beantwoordt, is in hoge mate afhankelijk van de kwaliteit van zijn eerdere tekstbestudering.

Ook dit hypothetische taakmodel laat zich onderbouwen door bevindingen in bestaand onderzoek. Leerlingen die voorafgaande aan de leestaak weten dat ze bij de vragen de tekst niet langer beschikbaar zullen hebben, bestuderen de tekst vooraf aanmerkelijk uitgebreider, met aanmerkelijk meer herlezingen, dan leerlingen die weten dat ze de tekst tijdens de vragen nog wel beschikbaar zullen hebben (Ferrer et al., 2017; Higgs, Magliano, Vidal-Abarca, Martinez, & McNamara, 2017; vgl. Agarwal & Roediger, 2011). Bovendien lijkt in deze situatie de antwoordscore wél vooral de kwaliteit van het vooraf lezen te weerspiegelen (Ferrer et al., 2017; Ozuru et al., 2007; Schaffner & Schiefele, 2013; Schroeder, 2011): “When students are unable to refer back to the text, they have to access their situation model to answer the comprehension questions. Thus, the more effort readers put into organizing and integrating the text during [initial] reading, the better their comprehension performance” (Schroeder 2011, p. 892). Of en in welke mate de context achter deze taak – *high-stakes test* vs. normale leestaak – hierop van invloed is, is bij ons weten overigens (nog) niet onderzocht.

1.4 Onderzoekshypothesen

Welke impact heeft de bestudering vooraf van een tekst op de antwoordscore van een vaardige leerling bij de tekst met vragen? In deze studie willen we de relatie tussen vooraf lezen en vragen beantwoorden nader blootleggen door twee variabelen te manipuleren: a) de beschikbaarheid van de tekst tijdens de antwoordfase, b) de intensiteit waarmee leerlingen een tekst vooraf bestuderen. Daarbij nemen we aan dat zowel de ‘tekstraadpleegbaarheid’ als de intensieve initiële tekstbestudering bij de vragen de antwoordscore van leerlingen aantoonbaar zal verhogen.

We verwachten:

a) dat leerlingen die de tekst vooraf intensief bestuderen én beschikbaar hebben bij de vragen, aanmerkelijk hoger zullen scoren dan leerlingen die de tekst intensief bestuderen maar niet langer beschikbaar hebben bij de vragen;

b) dat naarmate leerlingen een tekst *intensiever* bestuderen, ze hoger zullen scoren bij de vragen, indien de tekst beschikbaar blijft tijdens de antwoordfase.

Deze twee onderzoekshypothesen onderzoeken we bij twee zakelijke teksten, in twee experimenten met in totaal vijf verschillende condities. Dat doen we bij vaardiger lezers in het Nederlandse voortgezet onderwijs: 5-vwo-leerlingen (Expertgroep Doorlopende Leerlijnen Taal, 2008; Gubbels, Van Langen, Maasen, & Melissen, 2019). Omdat we vooral in begrip op het situatiemodelniveau geïnteresseerd zijn, stellen we in het experiment enkel vragen die op het situatiemodel betrekking hebben en die vaardiger lezers, met of zonder de tekst beschikbaar, vanuit een adequaat situatiemodel zouden moeten kunnen beantwoorden.

2 Methode

2.1 Experimentopzet en participanten

In het eerste experiment (I) manipuleren we de tekstschikbaarheid tijdens de vraagbeantwoording: hebben leerlingen die de tekst bij de vragen nog kunnen raadplegen, een hogere score dan leerlingen die de tekst niet meer kunnen raadplegen? In dit experiment is er sprake van twee experimentele condities: conditie IA ('Raadpleegbaar'), waarin de tekst bij de vragen raadpleegbaar blijft, en conditie IB ('Niet-raadpleegbaar'), waarin de tekst bij de vragen niet raadpleegbaar blijft.

In het tweede experiment (II) manipuleren we de intensiteit van de initiële tekstbestudering: in hoeverre verhoogt een intensieve tekstbestudering vooraf de score op de vragen? Hier is er sprake van drie experimentele condities: IIA ('Intensief-lezen'), waarin de leerlingen vooraf de tekst intensief moeten bestuderen; IIB ('Lezen'), waarin de leerlingen vooraf de tekst moeten lezen maar zelf de intensiteit daarvan bepalen; en IIC ('Vra-

gen'), waarin de leerlingen niet de tekst vooraf hoeven te bestuderen en naar eigen inzicht kunnen starten met het beantwoorden van de vragen of met het lezen van de tekst.

We nemen beide experimenten af onder vaardige lezers, 5-vwo-leerlingen. Van 5-vwo-leerlingen mogen we verwachten dat ze een volwaardig taakmodel hebben opgebouwd rondom de bekende leestaak tekst met vragen en dat zij aanpassingen in de taakinrichting adequaat in hun taakmodel kunnen verdisconteren. In alle condities krijgen de deelnemers hetzelfde materiaal onder ogen: twee zakelijke teksten, elk gevolgd door negen situatiemodelvragen. Wel variëren we (noodgedwongen¹) in toetsvorm: experiment I wordt afgenomen als diagnostische toets, experiment II als summatieve toets.

Experiment I. In experiment I, waarin we de impact van tekstbeschikbaarheid op de antwoordscore onderzoeken, hebben we twee condities ingericht: IA ('Raadpleegbaar') en IB ('Niet-raadpleegbaar'). Dit experiment is afgenomen in tien 5-vwo-klassen ($N = 174$) op vier voortgezet onderwijscholen in het midden en zuiden van Nederland. Deelnemende docenten Nederlands boden leerlingen de test aan als diagnostische toets in de reguliere les Nederlands, in twee zittingen van 50 minuten. Omdat afname in klassenverband plaatsvond, werden klassen gerandomiseerd toegekend aan een van de twee experimentcondities. Leerlingen wisten vooraf niet aan welke van de twee condities hun klas deelnam. Daarnaast ontving elke klas gerandomiseerd twee versies van de toets, toetsboekjes waarin enkel de volgorde van de twee sets teksten en vragen werd omgewisseld.

Bij aanvang van de toets ontvingen de deelnemende leerlingen in beide condities de eerste tekst (lengte: 800-900 woorden), voorafgegaan door een uitgebreide instructie. Daarbij werden ze geïnstrueerd de algemene gedachtegang erin zo te bestuderen dat ze later tekstafhankelijke begripsvragen erover zouden kunnen beantwoorden zonder dat ze de tekst daarbij nog raadpleegbaar hadden: de tekst zou na bestudering worden ingenomen, waarna de vragen zouden worden verstrekt.

Onderstrepingen en aantekeningen in de tekst waren toegestaan, alsmede aantekeningen op een apart blad, die echter zouden worden ingenomen met de teksten. Voor deze tekstbestudering hadden leerlingen per tekst een behoorlijke hoeveelheid tijd beschikbaar: 15 minuten. Leerlingen konden de tekst niet eerder inleveren, noch konden ze de vragen eerder opvragen. Met deze instructie hoopten we de leerlingen in experiment I te dwingen tot een zeer intensieve bestudering van de tekst. Immers, lezers lezen vooraf een tekst doorgaans aanmerkelijk langer en met meer aandacht voor de hoofdzaken wanneer ze weten dat ze een tekst niet beschikbaar zullen hebben bij de vragen (Ferrer et al., 2017; Schroeder, 2011).

In conditie IA ('Raadpleegbaar') werden de leerlingen met deze instructie echter misleid: bij het overhandigen van de vragen werd de tekst uiteindelijk *niet* ingenomen. De surveillant nam dus na vijftien minuten de tekst niet in en overhandigde enkel de vragen, waarbij deze aangaf dat ze tekst wel degelijk konden raadplegen. Deze procedure werd in de tweede les herhaald, met de andere set tekst en vragen.²

In conditie IB ('Niet-raadpleegbaar') nam de surveillant na vijftien minuten de tekst wél in en overhandigde de vragen. Hier beschikten de leerlingen dus bij het beantwoorden van de vragen, conform instructie, niet langer over de tekst. Deze procedure werd in de tweede les herhaald, met de andere set tekst en vragen.

Leerlingen dienden voorafgaand aan het beantwoorden van de vragen als eerste op het antwoordblad aan te geven hoeveel minuten ze de tekst vooraf hadden bestudeerd. De surveillant controleerde of ze dit daadwerkelijk noteerden.

Experiment II. In experiment II, waarin we de intensiteit van de tekstbestudering vooraf manipuleren, hebben we drie condities ingericht: de condities IIA ('Intensief-lezen'), IIB ('Lezen') en IIC ('Vragen'). Dit experiment is afgenomen in zestien 5-vwo-klassen ($N = 398$) op vier voortgezet onderwijs scholen in het midden en zuiden van Nederland. Hier maakte de toets onderdeel uit van de summa-

tieve toetsing in de toetsweek, waarbij deze werd afgenomen in één centrale zitting van 100 minuten. Docenten van de deelnemende scholen surveilleerden bij de toets. Omdat afname steeds plaatsvond op een centraal toetsmoment, konden leerlingen (en dus niet klassen, zoals bij experiment I) binnen scholen gerandomiseerd worden toegekend aan een van de drie experimentcondities. Leerlingen wisten vooraf niet aan welke conditie ze deelnamen, noch dat er sprake was van onderling afwijkende afnamecondities. In elke afzonderlijke testruimte had telkens slechts één conditie zitting. Daarnaast ontving elke groep gerandomiseerd twee versies van de toets, waarin enkel de volgorde van de sets afweek.

Conditie IIA ('Intensief-lezen') is in opzet, instructie en procedure gelijk aan conditie IA ('Raadpleegbaar') uit experiment I: we herhalen daarom hier enkel de hoofdzaken. In conditie IIA ('Intensief-lezen') ontvingen de leerlingen allereerst een tekst met daarbij de instructie de algemene gedachtegang erin zodanig te bestuderen dat ze straks begripsvragen erover zouden kunnen beantwoorden zonder de tekst te kunnen raadplegen: de tekst zou dan worden ingenomen. Daarvoor hadden ze per tekst 15 minuten beschikbaar. Uiteindelijk bleken leerlingen hiermee te zijn misleid, want bij het overhandigen van de vragen werd de tekst uiteindelijk niet ingenomen. Zo hoopten we een intensieve voorafbestudering van de tekst bij leerlingen af te dwingen. Na vijftig minuten herhaalde zich deze procedure.

In conditie IIB ('Lezen') wilden we leerlingen ertoe aanzetten dat ze minimaal de hele tekst gelezen hadden voordat ze starten met het beantwoorden van de vragen. In deze conditie ontvingen de leerlingen eerst een tekst, met de instructie de algemene gedachtegang zo te bestuderen dat ze straks begripsvragen erover zouden kunnen beantwoorden. Die begripsvragen zouden na 15 minuten worden uitgereikt; de leerlingen vernamen hier dus expliciet dat de tekst bij de vragen raadpleegbaar zou blijven. Desgewenst kon een leerling bij de surveillant na lezing wel eerder om de vragen verzoeken. Doordat deze leerlingen wisten dat de tekst raadpleegbaar

bleef bij de vragen en ze de optie hadden de vragen eerder op te vragen, verwachtten we dat deze leerlingen vooraf de tekst zouden bestuderen maar daarbij een minder rijk gespecificeerd begrip van de tekst zouden opbouwen dan de leerlingen in de vorige conditie.

In conditie IIC ('Vragen') verliep de procedure grotendeels zoals leerlingen doorgaans bij leestoetsen in het schoolvak Nederlands gewend zijn. Hier ontvingen de leerlingen bij de start van het experiment een vragenboekje, waarin achterin de tekst als bijlage was opgenomen. De instructie op de toets was in deze conditie minimaal: leerlingen werd slechts geïnstrueerd de begripsvragen te beantwoorden, en daarbij werd vermeld dat ze niet noodzakelijkerwijs de tekst vooraf hoefden te bestuderen. Wanneer leerlingen klaar waren, konden ze het materiaal inleveren; na vijftig minuten haalde de surveillant al het materiaal op. Daarna overhandigde hij de tweede set tekst en vragen, met dezelfde instructie als bij de eerste set. Doordat leerlingen in deze conditie niet werden aangespoord tot initiële tekstbestudering, verwachtten we dat ze de tekst vooraf enkel globaal of gedeeltelijk zouden bestuderen – of mogelijk zelfs direct zouden starten met de vragen – en zo vooraf een minder diep tekstbegrip zouden hebben opgebouwd dan de leerlingen in de twee voorgaande condities.

2.2 Voorbereiding op deelname aan experiment

Vooraf hadden deelnemende docenten Nederlands aangegeven dat hun 5-vwo-leerlingen niet vertrouwd waren met een vorm van tekstbegripstoetsing waarbij de tekst niet langer raadpleegbaar was bij de vraagbeantwoording. We wilden niet dat leerlingen uit onervarenheid met deze toetsingsvorm tijdens het experiment zouden worden verrast en mogelijk een inadequaat taakmodel zouden opstellen. Daarom bespraken deelnemende docenten enkele weken voorafgaande aan het experiment aan de hand van een PowerPoint met hun klassen welke implicaties (niet-)tekstraadpleegbaarheid heeft voor de initiële tekstbestudering en voor de vraagbeantwoording.

Daarbij ontvingen de leerlingen ook een korte introductie op het experiment. De docent gaf aan dat de toets onderdeel uitmaakte van een landelijke peiling van het tekstbegrip van vwo-leerlingen, naar aanleiding van recente berichten over een terugval in tekstbegrip in Nederland, onder anderen bij vwo-leerlingen (vgl. Gubbels et al., 2019).

2.3 Constructie van het materiaal

Omdat we nadrukkelijk het bestaande context- en taakmodel van 5-vwo-leerlingen bij tekstbegripstoetsen wilden aanspreken, werd bij het samenstellen van het materiaal zoveel mogelijk aangesloten op de reguliere toetsing van tekstbegrip in het voortgezet onderwijs. In havo- en vwo-bovenbouw wordt vaak geoefend met onderdelen uit de Centraal Schriftelijke Examens (CSE's) Nederlands havo en vwo (Canton et al., 2013). Als materiaal werden daarom twee zakelijke opiniërende 'teksten met vragen' geselecteerd uit CSE's havo en vwo.

Bij de selectie van de CSE-onderdelen golden drie criteria. 1) Er werd allereerst op gelet dat de teksten een heldere tekststructuur kenden: een vaardige lezer moest al bestudeerend in staat zijn een duidelijke, samenhangende gedachtegang in de tekst aan te wijzen. 2) Daarnaast moesten de CSE-vragen vooral een appel doen op het situatiemodel van een lezer. We selecteerden daarbij de vragen die als *elaboration questions* of *bridging inference questions* konden worden beschouwd, typerende situatiemodelvragen (Kintsch, 1998). *Elaboration questions* vragen om de verbinding van een tekstpassage met de achtergrondkennis van een lezer; *bridging inference questions* vereisen dat een lezer twee of meer tekstpassages inhoudelijk verbindt (vgl. Kintsch, 1998; Land, 2009; O'Reilly & McNamara, 2007). 3) Ten slotte zochten we CSE-onderdelen die een relatief hoge betrouwbaarheid kenden, waarbij we gebruik konden maken van interne Citorapportages. Op grond hiervan selecteerden we uiteindelijk één onderdeel uit het CSE vwo 2019-1 ('De Kunst en de regels') en één onderdeel uit het CSE havo 2017-2 ('De wilskrachtige is een gewoontedier'). In Appendix 1 is een selectie uit beide sets opgenomen.

Na selectie werden de teksten in deze CSE-onderdelen enigszins ingekort (max. \pm 900 woorden) om intensieve bestudering binnen een relatief kort tijdsbestek te vergemakkelijken en de noodzakelijke mentale inspanning van deelnemers niet te overvragen, vooral van de groep deelnemers die de vragen zonder de tekst zouden moeten maken.

Deze twee CSE-onderdelen kenden volgens Citogegevens een betrouwbaarheid van $GLB = .46$ en $GLB = .41$, relatief lage waarden, waarin we moeten verdisconteren dat elk van beide onderdelen slechts een beperkt aantal items bevat (resp. 9 en 10 items) en dergelijk weinig omvangrijke CSE-onderdelen doorgaans een (nog) lagere betrouwbaarheid kennen.³

Vragen werden zoveel mogelijk overgenomen uit de betreffende CSE-onderdelen. Vragen die vooral een appel deden op de tekstbasis werden daarbij verwijderd, evenals analytische vragen over alineafuncties, drogredenen en argumentatieschema's. Slechts vragen op het situatiemodelniveau, vragen die (onderdelen van) de centrale gedachtegang van de tekst bevroegen, werden gehandhaafd. Het betrof zowel open als gesloten vragen. Op basis van inhoudelijk commentaar van de vakvereniging van docenten Nederlands op een van beide CSE-onderdelen (LTN, 2019a), werd een klein aantal onvolkomenheden in de originele vraagstelling aangescherpt. Daarnaast werden in beide onderdelen enkele vragen zo aangepast dat leerlingen ze konden beantwoorden zónder dat ze de tekst nog konden raadplegen. Een tweetal vragen ten slotte werd door de onderzoekers nieuw geconstrueerd. Een screeningspanel, bestaande uit drie gepromoveerde onderzoekers in taalbeheersing en drie geroutineerde bovenbouwdocenten Nederlands, screende daarop al het materiaal uitvoerig; daarnaast stelde het panel het correctiemodel vast.

De deelnemende docenten Nederlands corrigeerden de toets op basis van het vastgestelde correctiemodel. Een van de onderzoekers corrigeerde daarnaast afzonderlijk al het materiaal; afwijkingen in score werden in overleg in overeenstemming gebracht.

2.4 Analyse

Controle manipulatie. Om te achterhalen in hoeverre leerlingen vooraf de teksten inderdaad in wisselende intensiteit per conditie bestudeerden, ontwikkelden we een aantal controle-instrumenten.

Ten eerste vroegen we leerlingen in alle condities – behalve de leerlingen in conditie IIC ('Vragen'), die de tekst niet hoefden te bestuderen – voorafgaand aan het maken van de vragen de tijdsduur aan te geven van hun tekstbestudering. We gaan ervan uit dat, als leerlingen in de ene conditie gemiddeld langer een tekst bestuderen dan de leerlingen in een andere conditie, ze daarmee blijkgewen van meer cognitieve verwerkingstijd en dus intensiever de tekst hebben bestudeerd.

Ten tweede scoorden we in de tekstboekjes van alle condities of a) leerlingen tekstelementen onderstreepten of *highlighten*, en daarnaast b) of leerlingen bij de tekst aantekeningen maakten, al dan niet op een apart blad. Daarbij gaan we ervan uit dat leerlingen intensiever de tekst hebben bestudeerd naarmate ze vaker onderstrepen en/of aantekeningen maken. Voor het scoren van de onderstrepingen en aantekeningen op de leesteksten maakten we gebruik van een tweepuntsschaal. Hierbij gold '0': leerlingen onderstrepen niet / maken geen aantekeningen; '1': leerlingen onderstrepen wel / leerlingen maken geen aantekeningen.

Analyse. Voor elke leerling is allereerst de totaalscore berekend voor de eerste set tekst en vragen én voor de tweede set tekst en vragen afzonderlijk. Voor de analyse van de verschillen in leestijd per set én de gemiddelde antwoordscore tussen de condities wordt vervolgens gebruikgemaakt van multiniveau-modellen, waarbij we als *fixed effects* de conditie en set 'tekst met vragen' 1 of 2 meenamen, als hoofdeffect en interactie-effect. Daarbij zijn scores steeds genest binnen leerlingen en klassen.

Om te bezien of er verschillen bestaan in de hoeveelheid onderstrepingen en aantekeningen tussen de condities, stelden we hierbij vooraf een binomiaal model op, waarbij we condities en 'eerste/tweede set' als hoofdeffecten en interactie-effect meenemen en we een variantie schatten tussen klassen.

3. Resultaten

3.1 Algemeen

Experiment I. Hierin verliep de afname grotendeels zoals beoogd. De leerlingen ($N = 174$) zijn redelijk evenwichtig verspreid over de twee condities (conditie IA ‘Raadpleegbaar’: $n = 77$; conditie IB ‘Niet-raadpleegbaar’: $n = 97$), waarbij jongens en meisjes niet afwijkend verspreid waren over de condities ($\chi^2(1) = 0.53, p = .47$). Deelnemende docenten Nederlands gaven aan dat leerlingen doorgaans voldoende geconcentreerd en gemotiveerd waren.

Op twee van de vier scholen verliep de afname zonder problemen, al was er door de Covid-epidemie wat meer lesabsentie dan gebruikelijk. Op de twee andere scholen werd afname van het experiment lange tijd verhinderd door de tweede lockdown van het Nederlandse voortgezet onderwijs, vanaf december 2020. Pas na de hervatting van het reguliere onderwijs in juni 2021, kon het experiment worden afgenomen, waarbij de ene school het materiaal enkel in gehalveerde klassen kon afnemen en de andere school met veel absentie te kampen had.

De twee testen tezamen bleken uiteindelijk in hoge mate betrouwbaar te differentiëren tussen de twee condities ($GLB = .86$).

Experiment II. Hierin verliep de afname zoals beoogd. De leerlingen ($N = 398$) zijn redelijk evenwichtig verspreid over de drie condities (conditie IIA ‘Intensief-lezen’: $n = 139$; conditie IIB ‘Lezen’: $n = 113$; conditie IIC ‘Vragen’: $n = 146$), waarbij jongens en meisjes niet afwijkend verspreid waren over de condities ($\chi^2(2) = 2.94, p = .23$). Surveillanten en deelnemende docenten Nederlands gaven aan dat leerlingen gemotiveerd en geconcentreerd deelnamen aan de toets. Wat conditie IIC (‘Vragen’) betreft, merkten surveillanten op dat het merendeel van de leerlingen na het uitdelen van de toetsboekjes niet startte met de vragen, maar met het lezen van de tekst.

De twee testen tezamen bleken voldoende betrouwbaar te differentiëren tussen de drie condities ($GLB = .64$).

3.2 Controle manipulatie

Vooraf-leestijden: experiment I. Hier vinden we voor de gemiddelde vooraf-leestijd allereerst een hoofdeffect van conditie ($F(1, 144.23) = 6.21, p = .01$). Leerlingen in conditie IA (‘Raadpleegbaar’) besteden gemiddeld 12.20 minuut ($se = .52$) aan een tekst terwijl de leerlingen in conditie IB (‘Niet-raadpleegbaar’) gemiddeld 12.88 minuut ($se = .50$) aan het vooraf lezen van een tekst besteden.

We kunnen niet aantonen dat de ene tekst vooraf langer of korter wordt gelezen dan de andere ($F(1, 154.89) = .28, p = .60$), noch vinden we een interactie-effect tussen tekst en conditie ($F(1, 237.97) = .88, p = .35$): we kunnen dus niet aantonen dat de ene tekst korter of langer wordt gelezen dan de andere tekst, afhankelijk van de conditie.

Vooraf-leestijden: experiment II. Hier vinden we dat leerlingen in conditie IIA (‘Intensief-lezen’) beide teksten vooraf duidelijk anders lezen dan leerlingen in conditie IIB (‘Lezen’) ($F(1, 242.95) = 50.92, p < .001$). De leerlingen in conditie IIA (‘Intensief-lezen’) besteden vooraf gemiddeld 14.31 minuut ($se = .28$, voor tekst 2) aan het vooraf lezen terwijl leerlingen in conditie IIB (‘Lezen’) gemiddeld 11.86 ($se = .31$) eraan besteden.

Er blijkt ook sprake van een verschil tussen teksten ($F(1, 241.95) = 5.91, p = .02$). Leerlingen blijken aan tekst 1, de vwo-tekst, gemiddeld .49 minuut ($se = .19$) meer tijd te besteden dan aan tekst 2, de havotekst.

Er is evenwel geen sprake van een interactie-effect tussen tekst en conditie ($F(1, 241.95) = .27, p = .61$).

Onderstrepingen en aantekeningen: experiment I. In 40 procent van de teksten zijn elementen *onderstreept*. Dit onderstrepen verschilt niet aantoonbaar tussen condities ($F(1, 313) = 3.12, p = .08$), noch tussen tekst 1 en tekst 2 ($F(1, 313) = 2.17, p = .14$). We kunnen evenmin een interactie-effect tussen tekst en conditie aantonen ($F(1, 313) = 3.68, p = .06$).

In 13 procent van de teksten zijn *aantekeningen gemaakt*. Dit aantekeningen maken verschilt niet aantoonbaar tussen condities (F

(1, 313) < .001, $p > .99$), noch tussen tekst 1 of tekst 2 ($F(1, 313) = .68, p = .41$). We kunnen ook geen interactie-effect tussen tekst en conditie aantonen ($F(1, 313) = .82, p = .37$).

In dit experiment vinden we dus, zoals verwacht, geen verschillen tussen condities op de controlevariabelen.

Onderstrepingen en aantekeningen: experiment II. Per conditie blijken leerlingen af te wijken in de gemiddelde hoeveelheid *onderstrepingen* ($F(2, 789) = 23.48, p < .001$). In 79.4 procent ($se = .04$, dus 4 procent) van de teksten van leerlingen in conditie IIA ('Intensief-lezen') zijn zaken onderstrept en in de teksten van leerlingen in conditie IIB ('Lezen') is dat 84.4 procent ($se = .04$). In de teksten van leerlingen in conditie IIC ('Vragen') is dat 57.8 procent ($se = .06$). De leerlingen in conditie IIA en IIB onderstrepen daarmee significant vaker ($p < .001$) dan de leerlingen in conditie IIC. Leerlingen in conditie IIA onderstrepen echter niet aantoonbaar meer of minder dan de leerlingen in conditie IIB ($p = .88$).

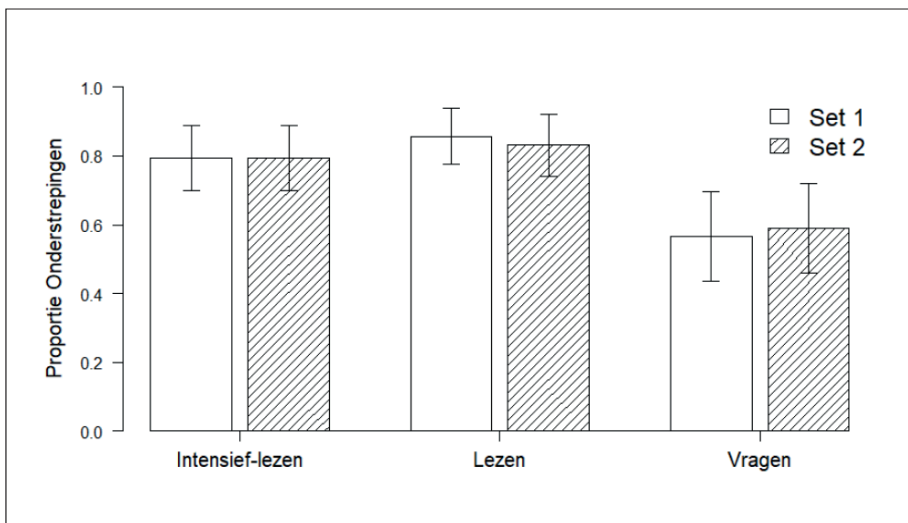
We kunnen niet aantonen dat leerlingen in tekst 1 minder of meer onderstrepen dan in tekst 2 ($F(1, 789) = .03, p = .85$), noch kunnen we een interactie-effect tussen tekst en conditie aantonen ($F(2, 789) = .22, p = .80$).

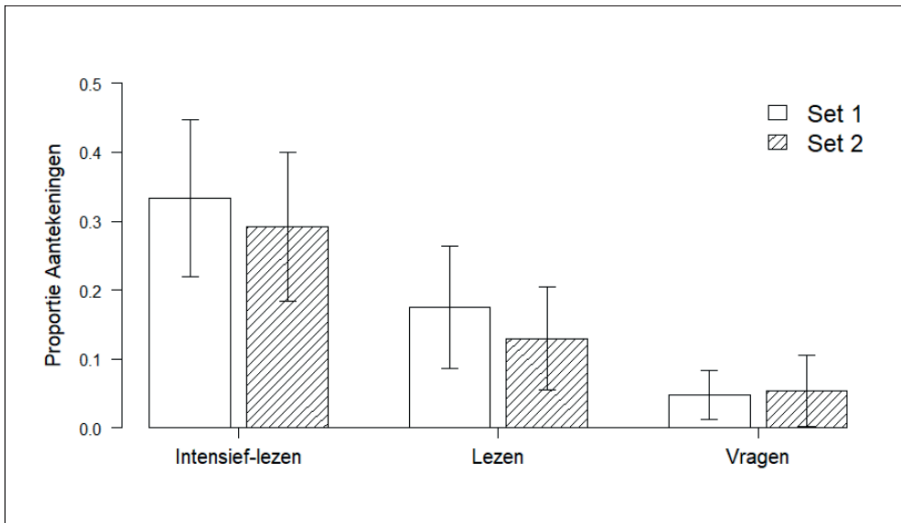
In Figuur 1 zijn de geschatte gemiddelden voor de onderstrepingen per conditie per tekst weergegeven.

Wat *aantekeningen maken* betreft, verschillen condities significant ($F(2, 789) = 30.76, p < .001$). Van de teksten van leerlingen in conditie IIA ('Intensief-lezen') bevat 31.2 procent aantekeningen ($se = .05$). Voor de teksten van leerlingen in conditie IIB ('Lezen') is dat 15.1 procent ($se = .03$), en voor de teksten van leerlingen in conditie IIC ('Vragen') 5.1 procent ($se = .02$). Daarmee onderstrepen de leerlingen in conditie IIA beduidend vaker dan de leerlingen in conditie IIB ($p < .01$) en vaker dan de leerlingen in conditie IIC ($p < .001$). Leerlingen in conditie IIB onderstrepen daarmee ook vaker dan de leerlingen in conditie IIC ($p = .03$). We kunnen dan ook concluderen dat de manipulatie in experiment II geleid heeft tot een andere taakaanpak tussen de condities.

Leerlingen blijken in tekst 1 niet aantoonbaar meer of minder aantekeningen te maken dan in tekst 2 ($F(1, 789) = .40, p = .53$). Er is evenmin sprake van een interactie-effect tussen tekst en conditie ($F(2, 789) = .29, p = .75$). In Figuur 2 zijn de geschatte gemiddelden voor het aantekeningen maken per conditie per tekst weergegeven.

Figuur 1
Proportie onderstrepingen bij een tekst, voor set 1 en 2, per conditie, voor experiment II ('Intensiteit')





Figuur 2
 Proportie aantekeningen maken bij een tekst, voor set 1 en 2, per conditie, voor experiment II ('Intensiteit')

Tabel 1
 Parameterschattingen voor de antwoordscore bij experiment I ('Raadpleegbaarheid')

Parameter	coëfficiënt	SE
<i>Conditie</i>		
conditie IA 'Raadpleegbaar': set tekst en vragen 2	9.01	.35
conditie IB 'Niet- raadpleegbaar': set tekst en vragen 2	6.54	.33
<i>Set</i>		
set tekst en vragen 1	-.10 ^a	.42
<i>Conditie * Set</i>		
Δconditie IA 'Raadpleegbaar': set tekst en vragen 1	-1.41	.65
<i>Varianties</i>		
S ² klas	.37	.33
<i>(Co)varianties tussen leerlingen</i>		
	S ² set1	S ² set2
S ² set1	13.55 (1.56)	.37 ^b
S ² set2	5.43 (.62)	2.93 (.76)

^aniet-significant;

^bcorrelatie tussen set1 en set2

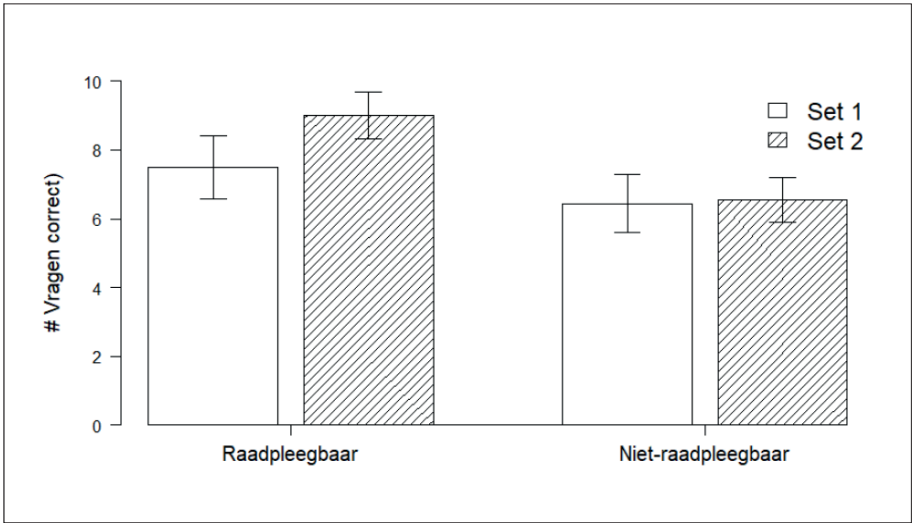
allereerst een duidelijk verschil tussen condities ($F(1, 110.62) = 22.44, p < .001$). In Tabel 1 zijn de schattingen voor de antwoordscore weergegeven voor experiment 1. We zien dan dat de leerlingen in conditie IB ('Niet-raadpleegbaar') per set aanmerkelijk lager scoren dan de leerlingen in conditie IA ('Raadpleegbaar'): leerlingen in conditie IA ('Raadpleegbaar') behalen een gemiddelde score van 9.01 ($se = .35$, voor de tweede set tekst en vragen), terwijl leerlingen in conditie IB ('Niet-raadpleegbaar') een gemiddelde score van 6.54 ($se = .32$) behalen.

We vinden eveneens een verschil in score tussen de eerste set tekst en vragen en de tweede set tekst en vragen ($F(1, 164.88) = 7.60, p < .01$), zoals er ook sprake blijkt van een interactie-effect tussen de sets en condities ($F(1, 238.82) = 4.75, p = .03$). Leerlingen in conditie IA ('Raadpleegbaar') blijken bij de eerste set tekst en vragen gemiddeld -1.41 ($se = .65$) lager te scoren dan bij de tweede set tekst en vragen. In Figuur 3 zijn de gemiddelde scores voor de beide sets in de twee condities weergegeven.

Leerlingen in conditie IB ('Niet-raadpleegbaar') scoren dus gemiddeld steeds lager dan de leerlingen in conditie IA ('Raadpleegbaar'), al is dit onderscheid in score tussen de condities groter voor de tweede set (de havo-

3.3 Antwoordscore

Experiment I. Kijken we vervolgens naar de gemiddelde antwoordscore, dan vinden we



Figuur 3
Gemiddelde antwoordscore voor set 1 en 2, per conditie, voor experiment I ('Raadpleegbaarheid')

Tabel 2
Parameterschattingen voor de antwoordscore bij experiment II ('Intensiteit'). (NB: het interactie-effect *Conditie*Set* is achterwege gelaten bij de hier gerapporteerde schattingen.)

Parameter	coëfficiënt	SE
<i>Conditie</i>		
conditie IIA ('Intensief-lezen': set tekst en vragen 2)	9.86	.20
conditie IIB ('Lezen': set tekst en vragen 2)	10.15	.21
conditie IIC ('Vragen': set tekst en vragen 2)	10.29	.19
<i>Set</i>		
set tekst en vragen 1	.39	.17
<i>Varianties</i>		
S ² klas	.15	.11
<i>(Co)varianties tussen leerlingen</i>		
	S ² set1	S ² set2
S ² set1	9.91 (.71)	.24 ^a
S ² set2	4.36 (.32)	1.57 (.34)

^a correlatie tussen set1 en set2

het vwo-examen, zeer hoog is (13.55, *se* = 1.56). De variantie bij de tweede set, het onderdeel uit het havo-examen, is daarmee vergeleken duidelijk lager (5.43, *se* = .62). Dit duidt erop dat de vwo-set meer verschillen tussen leerlingen laat zien dan de havo-set (en dus waarschijnlijk aanmerkelijk pittiger is).

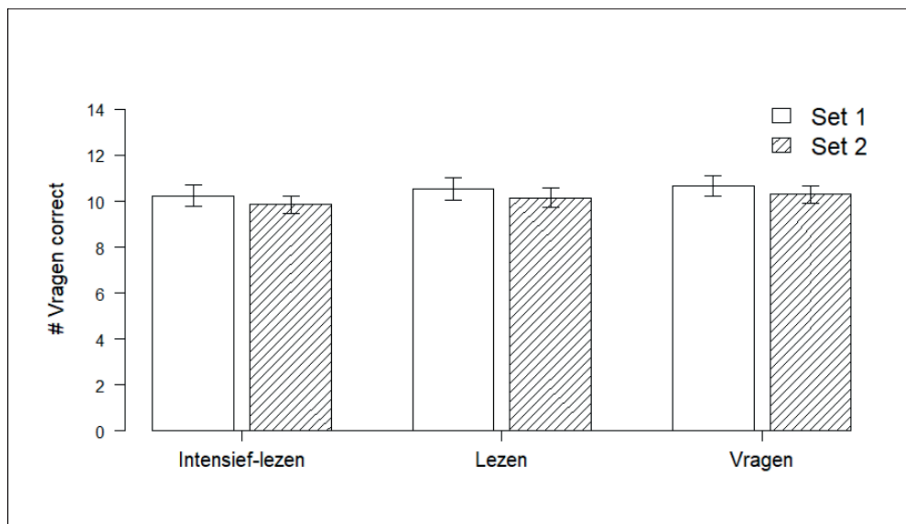
De covariantie tussen beide sets is 2.93 (*se* = .76), waarbij de correlatie tussen beide .34 is.

Experiment II. Vonden we eerder duidelijke verschillen in gemiddelde antwoordscore tussen de condities, bij het tweede experiment kunnen we hierin geen verschillen aantonen ($F(2, 395.76) = .55, p = .58$). In Tabel 2 zijn de schattingen voor de gemiddelde antwoordscore weergegeven: ook niet-significante parameterschattingen zijn hierin opgenomen. Daarin is te zien dat de leerlingen in conditie IIC ('Vragen') in absolute zin hoger scoren dan de leerlingen in de twee andere condities – zie Tabel 2 –, evenwel, dit onderscheid tussen condities is niet significant.

Wel vinden we een verschil in score tussen de eerste en de tweede set tekst en vragen ($F(1, 397.47) = 5.11, p = .02$): leerlingen scoren bij de eerste set over het algemeen .39 (*se* = .16) hoger dan bij de tweede set.

Een interactie-effect tussen de sets en con-

set) dan voor de eerste set (de vwo-set). Onderaan in Tabel 1 zijn de varianties weergegeven. We zien dan dat de variantie tussen klassen zeer gering is (.37, *se* = .33) en dat de variantie bij de eerste set, het onderdeel uit



Figuur 4
Gemiddelde antwoordscore voor set 1 en 2, per conditie, voor experiment II ('Intensiteit')

dities is evenwel niet aantoonbaar ($F(2, 397.48) = 2.05, p = .13$); in Tabel 2 is deze voor het berekenen van de schattingen daarom achterwege gelaten. In Figuur 4 zijn de gemiddelde scores voor de beide sets voor de drie condities weergegeven.

Onderaan in Tabel 2 zijn weer de varianties weergegeven. We zien dat de variantie tussen klassen erg beperkt is ($.15, se = .11$). De variantie in leeringscores bij de eerste set tekst en vragen (het vwo-examen) is wederom erg hoog ($9.85, se = .71$), terwijl de variantie bij de tweede set (het havo-examen) aanzienlijk lager is ($4.35, se = .31$): de pittige vwo-set lijkt weer tot meer verschillen tussen leerlingen te leiden dan de wat eenvoudiger havo-set. De covariantie tussen beide sets is relatief beperkt ($1.59, se = .34$), waardoor de correlatie tussen beide tests $.24$ bedraagt.

3.4 Samenvatting resultaten

Experiment I: raadpleegbaarheid. In de eerste variant, waarin het experiment als diagnostische toets werd afgenomen, kunnen we weinig tot geen verschillen in lees- en antwoordgedrag aantonen tussen de twee condities. Leerlingen in conditie IA ('Raadpleegbaar') onderstrepen niet meer of minder dan de leerlingen in conditie IB ('Niet-raadpleegbaar'), noch maken ze meer of minder aantekeningen.

De leerlingen in conditie IA besteden wat minder tijd aan het lezen van de tekst dan de leerlingen in conditie IB.

Wat de antwoordscore betreft, vinden we een aanmerkelijk verschil tussen de condities. De leerlingen in conditie IA ('Raadpleegbaar') scoren beduidend hoger dan de leerlingen in conditie IB ('Niet-raadpleegbaar').

Experiment II: intensiteit. In het tweede experiment, dat als summatieve toets werd afgenomen, zien we allereerst duidelijke verschillen in leesgedrag tussen de drie condities. Voor de leerlingen in conditie IIA ('Intensief-lezen') geldt dat ze een tekst langer lezen en vaker aantekeningen maken bij een tekst dan de leerlingen in conditie IIB ('Lezen'); bovendien onderstrepen de leerlingen in conditie IIA ('Intensief-lezen') vaker tekstelementen dan de leerlingen in conditie IIC ('Vragen'), zij het niet vaker dan de leerlingen in conditie IIB ('Lezen').

De leerlingen in conditie IIB ('Lezen') maken weer vaker aantekeningen en onderstrepen vaker tekstelementen dan de leerlingen in conditie IIC ('Vragen'). De leerlingen lijken dus per conditie, zoals beoogd, in wisselende mate van intensiteit de tekst hebben bestudeerd. *Grosso modo* kunnen we dus

stellen: naarmate in de conditie intensiever vooraf-leesgedrag werd afgedwongen, gingen leerlingen meer onderstrepen en aantekeningen maken.

We konden in dit experiment echter – tegen de verwachting in – geen verschil in antwoordscore aantonen tussen de drie condities.

Beide experimenten. Conditie IA ('Raadpleegbaar') en conditie IIA ('Intensief-lezen') wijken inhoudelijk niet af in materiaal en instructie, maar wel in hun taakcontext: diagnostische vs summatieve toets. Leerlingen in conditie IIA ('Intensief lezen'), die de summatieve toets maakten, scoren aanmerkelijk hoger op de vragen dan de leerlingen in conditie IA ('Raadpleegbaar'), die de diagnostische toets maakten, waarschijnlijk doordat deze eerste groep meer geconcentreerd en gemotiveerd deelnam.

Daarnaast: de variantie bij de eerste set (het vwo-examenonderdeel) is, zeker in vergelijking met de variantie bij de tweede set (het havo-examenonderdeel), bij beide experimenten zeer hoog. Deze eerste set lijkt al met al beduidend moeilijker te zijn geweest voor leerlingen. Deze moeilijkheid lijkt ook een mogelijke verklaring voor het feit dat in experiment I ('Raadpleegbaarheid') leerlingen bij set 1 lager scoren dan bij set 2, terwijl ze in experiment II ('Intensiteit') op set 1 juist hoger scoren dan op set 2: door de cognitieve inspanning die de vwo-set vergt, zetten leerlingen zich in de diagnostische context specifiek voor deze set waarschijnlijk minder in dan in de summatieve context.

4. Discussie

In deze studie wilden we de relatie tussen het vooraf bestuderen en de antwoordscore bij teksten met vragen nader blootleggen. Dat deden we door in twee experimenten twee variabelen te manipuleren: enerzijds de intensiteit van tekstbestudering wanneer vaardige lezers vooraf een tekst lezen, en anderzijds de mogelijkheid om de tekst te raadplegen wanneer vaardige lezers vragen beantwoorden.

Onze eerste hypothese luidde: 'Leerlingen

die de tekst vooraf intensief bestuderen én beschikbaar hebben bij de vragen, zullen aanmerkelijk hoger scoren dan leerlingen die de tekst intensief bestuderen en niet langer beschikbaar hebben bij de vragen.' Dit konden we zeer duidelijk aantonen. De verwachte, doorslaggevende rol die tekstraadpleging tijdens het antwoordproces heeft voor de antwoordscore, vinden we hier bevestigd. Doordat leerlingen via tekstraadpleging hun voor de vraag inadequate of ontbrekende tekstbegrip kunnen repareren en hun conceptantwoorden kunnen controleren, wordt de antwoordscore in sterke mate beïnvloed door de tekstraadpleegbaarheid tijdens de vraagbeantwoording.

Onze tweede hypothese luidde: 'Naarmate leerlingen een tekst *intensiever* bestuderen, zullen ze hoger scoren bij de vragen, indien de tekst beschikbaar blijft tijdens de antwoordfase'. Ondanks duidelijke, verwachte verschillen in leesgedrag tussen de drie condities in experiment II, konden we dit niet aantonen. Als een intensieve tekstbestudering de antwoordscore beïnvloedt, dan ontbreekt daarvoor in deze studie elke aanwijzing.

Dat we deze tweede hypothese niet hebben kunnen aantonen, valt des te meer op omdat we in experiment II daartoe een optimale context creëerden. De deelnemers waren vaardige, op de leestaak voorbereide lezers die, naar surveillantanten aangaven, geconcentreerd werkten aan een toets met een relatief zware weging voor het vak Nederlands. Deze lezers zullen, zogezegd, vaak gewerkt hebben vanuit een krachtig, adequaat context- en taakmodel. Daar komt nog bij dat ze uitsluitend situatiemodelvragen kregen voorgelegd, verbonden aan de centrale gedachtegang van de tekst. Als leerlingen vooraf zich deze centrale gedachtegang hebben eigengemaakt, dan zouden ze daarna juist deze vragen gemakkelijker (en dus vaker goed) moeten kunnen beantwoorden. Echter, zelfs onder deze vrijwel optimale omstandigheden konden we niet aantonen dat een intensieve tekstbestudering tot een hogere antwoordscore leidt. Kennelijk voegt een *intensieve* tekstbestudering weinig tot niets aantoonbaar toe aan de uiteindelijke kwaliteit van het antwoord. Daarmee lijkt de eerdere conclusie uit experi-

ment I andermaal te worden bevestigd: de tekstraadpleging tijdens de vragen geeft vooral de doorslag voor de antwoordscore – en niet zozeer een diepgaande tekstbestudering vooraf.

Het hypothetische taakmodel van een vaardiger lezer bij de taak tekst met vragen, zoals we dat in paragraaf 1.3 onderbouwden, kunnen we dus met deze bevindingen grotendeels bevestigen. Wanneer de tekst bij de vragen beschikbaar blijft, beïnvloedt vooral het antwoordproces de antwoordscore. Het vooraf-leesproces is daarbij instrumenteel: het staat lezers toe bij de beantwoording de vraagstelling beter te duiden en beter en sneller antwoord-relevante tekstdelen te lokaliseren (vgl. Bax, 2013; Brunfaut & McCray, 2015; Rooijackers et al., 2020). Aan dat hypothetische taakmodel kunnen we nu toevoegen dat een intensief vooraf-leesproces bij complexe situatiemodelvragen hoogstwaarschijnlijk geen positief effect heeft op het uiteindelijke antwoord. Wat leerlingen bij het vooraf-lezen nog hebben nagelaten te doorgronden, kunnen ze afdoende repareren tijdens het antwoordproces als de tekst dan beschikbaar blijft. Het volstaat om de tekst vooraf met enige aandacht door te nemen, want begripsconstructie tijdens het antwoordproces is het enige waar het bij deze taak om draait.

De impact van de taakcontext op de prestaties van leerlingen blijkt daarbij opvallend groot. Terwijl de onderzoekscondities IA ('Raadpleegbaar') en IIA ('Intensief-lezen') zich enkel onderscheiden in toetsingsvorm, bestuderen de leerlingen bij de diagnostische toets (experiment I) de tekst duidelijk korter dan de leerlingen bij de summatieve toets (experiment II), en hebben ze bovendien een aanmerkelijk lagere antwoordscore, met name bij de eerste, pittige vwo-set. Dit sluit aan bij de *parsing the context hypothesis* (Britt, Rouet, & Durik, 2018): naarmate het belang van de taak minder groot is, zal een lezer zich minder snel en minder diepgaand tot een complexe leestaak zetten. Een diagnostisch gebruikte toets zet een lezer dan – het ligt voor de hand – minder aan tot een diepe, nauwkeurige, gelaagde werkwijze dan een toets met een summatief doel.

4.1 Kanttekeningen

Natuurlijk zijn er bij deze studie de nodige voorbehouden te plaatsen. Vooropgesteld: dit onderzoek had uitsluitend betrekking op (tamelijk pittige) situatiemodelvragen, bij relatief lange, zakelijke opiniërende teksten, binnen de taakcontext van een (al dan niet diagnostische) leestoets, onder vaardiger (adolescent-)lezers. Andersoortige vragen (eenvoudigere of globalere vragen, analytische vragen, meer tekstbasis-gebonden vragen, etc.), bij andere teksten (kortere zakelijke teksten, andere genres zakelijke teksten, fictionele teksten, etc.), in een andere taakcontext (formatieve toetsing, 'gewone' leestaak in de les, etc.), onder andere lezers (vmbo-leerlingen, volwassen expert-lezers, etc.) zouden mogelijk andere resultaten hebben opgeleverd.

Daarnaast onderzochten we in deze studie de intensiteit van tekstbestudering, niet zozeer de kwaliteit ervan: het valt niet uit te sluiten dat een onderscheid binnen de steekproef tussen vaardiger en minder vaardige lezers andere accenten in de onderzoeksresultaten zou hebben opgeleverd. Verder dient opgemerkt dat leerlingen in onze studie de vragen vooraf niet konden inzien: wanneer ze deze kort vooraf hadden kunnen inzien voordat ze de tekst bestudeerden, hadden we mogelijk andere resultaten gevonden. Wel is het de vraag of vaardiger lezers bij een taak als deze vanuit zichzelf eerst de vragen kort doornemen: in conditie IIC 'Vragen' hadden de 5-vwo-leerlingen hiertoe de mogelijkheid maar startten ze desondanks doorgaans met het doornemen van de tekst.

Bovenal: de complexiteit, lokaliteit en aard van de vraagstelling én de complexiteit van de teksten zelf zijn waarschijnlijk zeer bepalend voor de resultaten in deze studie. Doordat we bij twee tamelijk complexe teksten situatiemodelvragen over de centrale gedachtegang stelden, werden hoge eisen aan het tekstbegrip van leerlingen gesteld. Daardoor zullen leerlingen vaak de behoefte hebben gehad de tekst nogmaals te raadplegen. Eenvoudigere, globalere vragen bij eenvoudiger teksten zullen lezers waarschijnlijk vaker vanuit hun vooraf-lezing van de tekst kunnen beantwoorden.

4.2 Implicaties voor het onderwijs

Deze studie heeft interessante implicaties voor het onderwijs in tekstbegrip. In Nederland staat de tekst met vragen als leestaak al sinds lange tijd ter discussie, omdat ze van een leerling een ‘weinig kritische lezer’ zou maken die ‘niet zelfstandig complexe teksten kan aanpakken zonder dat er vragen worden gesteld’ (Van de Wetering & Groenendijk, 2015, p. 5). Al bestaat hierover een brede consensus onder onderwijskundigen, docenten, taalwetenschappers en leesonderzoekers (bijv. Hoogeveen, 2018; LTN, 2019b; SLO, 2017), empirische onderbouwing bij deze kritiek ontbrak. Onderhavige studie impliceert dat een diepgaande, zelfstandige tekstbestudering vooraf voor leerlingen bij deze taak hoogstwaarschijnlijk geen meerwaarde heeft. Daarmee wordt de kritiek aannemelijk dat leerlingen niet zozeer met deze taak leren zelfstandig een tekst te bestuderen maar dat ze vooral ermee leren vragen over een tekst correct te beantwoorden.

Binnen de context van de gemiddelde leesles Nederlands wordt dat laatste zelfs nog aannemelijker. In onze studie wordt weliswaar materiaal uit de CSE’s Nederlands havo/vwo gebruikt, we handhaafden daarin alleen de situatiemodelvragen over de centrale tekstuele gedachtegang. In veel huidige leergangen Nederlands voor het Nederlandse secundaire onderwijs worden in leestaken echter sterk uiteenlopende vraagtypen gebruikt, in een nogal onvoorspelbare afwisseling: situatiemodelvragen staan tussen functioneel-analytische vragen, vragen over vocabulaire en argumentatie-analytische en -beoordelende vragen (Rooijackers, Van Silfhout, & Van den Bergh, 2021a). Een dergelijke taakinrichting bevordert nog sterker taakafhankelijk leesgedrag. Leerlingen hebben dan immers vooraf geen idee wat hun gevraagd zal gaan worden.

Dat leerlingen via deze reguliere invulling van de tekst met vragen vooral vragen over een tekst leren te beantwoorden, spreekt vanuit wetenschappelijk oogpunt overigens zo goed als vanzelf. Zonder duidelijke taakcontext, zonder concreet doel wordt de vooraftekstbestudering vooral een middel om het concrete, meetbare doel in de taak te bereiken: het correct beantwoorden van de vragen.

Dat heeft twee logische gevolgen voor het lezersgedrag bij dit soort taken. Ten eerste, omdat vooraf-lezen vooral een hulpmiddel is om de vragen beter te begrijpen en te beantwoorden, lezen leerlingen vooraf nauwelijks met diep begrip (Wolf, 2018; Onderwijsraad & Raad voor Cultuur, 2019) en lezen ze vooral ‘oppervlakkig’: leerlingen bouwen met name een tekstbasis op en enkel een zeer globaal situatiemodel. Dergelijk oppervlakkig leesgedrag werd inderdaad bij vwo-leerlingen aangetroffen (Breukink, Janssen, & Van den Bergh, (in voorbereiding); Rooijackers et al., 2020; Stevens, 2017). Ten tweede, doordat leerlingen zich vooral richten op de vragen, zullen ze vaak een hoge mate van taakvertrouwdheid en -afhankelijkheid opbouwen. Leerlingen worden zowel in primair als in secundair onderwijs vaak aan deze leestaak blootgesteld, en zo bouwen ze gedetailleerde kennis op over vraagtypen, en eveneens over vaardigheden en strategieën die per vraagtype wenselijk én adequaat zijn. Sporen van deze taakvertrouwdheid vonden we in een eerdere studie, deels van onze hand (Rooijackers et al., 2020).

Betekent dat nu dat een tekst met vragen geen tekstbegrip toetst? Voor de goede orde: wanneer leerlingen een vraag over een tekst goed beantwoorden, geven ze natuurlijk wél blijk van een vorm van tekstbegrip. In het leven van alledag hebben leerlingen immers vaak te maken met schoolse situaties waarin ze met een gerichte zoekvraag een tekst tot zich nemen, bijvoorbeeld bij werkstukken – zoals er ook tal van schoolse situaties zijn waarin ze vragen over een tekst moeten beantwoorden zónder dat de brontekst nog raadpleegbaar is, bijvoorbeeld bij proefwerken voor zaakvakken (Artelt, Schiefele, & Schneider, 2001; Ferrer et al., 2017; Schroeder, 2011). Deze beide vormen van tekstbegrip (vragen beantwoorden met én zonder de tekst erbij) lijken redelijk te correleren (Schroeder, 2011). Daarom kunnen de resultaten in deze studie in onze ogen ook niet worden gezien als ondersteuning bij de stelling dat een tekst met vragen slechts in beperkte mate begrip op het situatiemodel toetst (vgl. Kamalski, 2007), al sluiten de resultaten wel aan bij recente bedenkingen

over de beperkte cognitieve validiteit van de tekst met vragen als leestoets: de begripsprocessen bij het beantwoorden van vragen bij een tekst overlappen maar ten dele met de begripsprocessen van leerlingen bij 'echte' leestaken (zie met name O'Reilly et al., 2018). Implicaties van deze studie liggen, denken we, veeleer op inhoudelijk-didactische inrichting van de leesles Nederlands, met name ten aanzien van de tekst met vragen als trainingsinstrument. We werken er vier uit.

De eerste implicatie: vermijd een taakinrichting met vragen die inhoudelijk sterk uiteenlopen en elkaar onvoorspelbaar opvolgen en die niet aan een duidelijk doel of context zijn gekoppeld. Een andere werkwijze is aan te bevelen. Door bij deze taak te werken met een duidelijk, concreet leesdoel vooraf en met vragen die logisch aansluiten op dit concrete leesdoel, wordt de vooraf-bestudering van de tekst voor de lezer in een duidelijke leessituatie ingebed (McCrudden, Schraw, & Magliano, 2010; O'Reilly et al., 2018). Zo'n leesdoel kan bijvoorbeeld een voorbereidende leestaak voor een schrijfo opdracht zijn, maar kan ook worden aangeboden binnen een fictief, voor leerlingen herkenbaar scenario, van waaruit ze vooraf de tekst moeten lezen en daarna begripsvragen moeten beantwoorden (OECD, 2019; Sabatini, O'Reilly, Wang, & Dreier, 2018). Eventueel kan dan, zoals in onze studie, de tekst worden weggehaald bij het overhandigen van de vervolgvragen, waardoor de vragen voor leerlingen een check achteraf worden of ze de tekst vooraf voldoende hebben begrepen (Schroeder, 2011).

Een tweede implicatie: spreid in taaksoorten. Door in de leesles een veelvoud aan soorten taken aan te bieden (sorteertaken, argumentatietaken, synthesetaken, etc.), kunnen leerlingen minder taakafhankelijk en taakvertrouwd gedrag opbouwen. Als ze daarvoor de tekst vooraf zelfstandig moeten bestuderen, dan moet voor leerlingen wel vooraf duidelijk zijn aan welke dimensie van de tekst ze bij het lezen van de tekst speciale aandacht moet schenken: bijv. centrale gedachtegang, aanvaardbaarheid argumentatie, inhoudelijke consistentie, etc.

Een derde implicatie: stel bij leestaken vaker het leesproces zelf centraal. De aandacht gaat in de les zakelijk lezen nu vaak uit naar het leesproduct, doordat vragen bij een tekst al dan niet klassikaal worden beantwoord en besproken. Deze studie impliceert dat leerlingen voor een goede prestatie bij een tekst met vragen zich het beste kunnen richten op het antwoordproces, en niet zozeer op het vooraf-leesproces; daardoor raakt het zelfstandig bestuderen van een complexe tekst al snel buiten beeld. En weliswaar *modellen* docenten vaak, doen docenten voor hoe ze zelf zich teksten eigen maken (bijv. Linthorst & De Glopper, 2015), ze lijken in de les leerlingen zelden ertoe aan te zetten op hun eigen leesgedrag te reflecteren. In de leesles zou daarom bijvoorbeeld vaker het leesproces van leerlingen zelf centraal kunnen staan: hoe doorgronden leerlingen zelfstandig complexe teksten, welke leesproblemen ervaren ze dan en vanuit welke vaardigheden en strategieën kunnen ze nog beter dit soort teksten begrijpen (vgl. Couzijn, 1995; Rooijackers, Van Silfhout, & Van den Bergh, 2021b)?

Een vierde, meer indirecte implicatie is: verruim de blik op tekstbegrip in de les zakelijk lezen. In deze studie ontwierpen we twee experimenten waarbij leerlingen één samenhangende, kwalitatief hoogwaardige tekst via vragen gaandeweg moesten doorgronden, en daarmee sloten we aan bij de courante taakinrichting in huidige leidende leergangen Nederlands (Rooijackers et al., 2021a). Vakdidactische rationale achter deze taakinrichting is de gedachte dat leerlingen via vragen de diepere gedachtegang in een complexe betrouwbare tekst blootleggen en door oefening gaandeweg zichzelf deze benadering eigen maken (Hendrix & Hulshof, 1994). Daarmee blijven echter theoretische inzichten in tekstbegrip buiten beschouwing van met name ná de millenniumwisseling, inzichten die in het theoretisch kader van onze studie zijn meegenomen. Met name relevant lijkt het inzicht dat veel tekstbegrip in het leven van alledag wordt opgebouwd vanuit een specifiek leesdoel (Britt, Rouet, & Durik, 2018; vgl. McCrudden, Magliano, & Schraw, 2011). Door deze nadruk op de context rond

de tekst en de lezer, komen andere vaardigheden bij tekstbegrip in beeld: het kunnen bepalen van de kwaliteit, betrouwbaarheid en bruikbaarheid van een tekst, en het kunnen opbouwen van adequate taak- en contextmodellen bij leestaken (Britt & Rouet, 2012; Gubbels, 2020). Deze aspecten worden nu vaak niet of nauwelijks in de gemiddelde leestaak van het schoolvak Nederlands meegenomen (Gubbels, 2020; Rooijackers et al., 2021a), wat een van de verklaringen kan zijn waarom in internationaal onderzoek Nederlandse 15-jarige leerlingen relatief laag scoren op het domein reflecteren en evalueren (Gubbels et al., 2019).

Praktisch zorgt zo'n blikverruiming voor andere aandachtsgebieden in de leesles Nederlands (zie Rooijackers et al., 2021a). Leerlingen oefenen met het opbouwen en toepassen van adequate context- en taakmodellen in complexe leessituaties en reflecteren hierop: denk aan taken zoals werkstukken waarbij leerlingen informatie uit meerdere bronnen synthetiseren (Gil, Bråten, Vidal-Abarca, & Strømsø, 2010). Of leerlingen werken met leestaken waarin de teksten sterk variëren in kwaliteit, betrouwbaarheid en bruikbaarheid; daarbij ontwikkelen ze dan criteria en passen deze vervolgens toe op de set teksten (Anmarkrud, Bråten, & Strømsø, 2014). Of de aandacht gaat in de les uit naar de selectiecriteria van informatiebronnen op internet (Bråten, Stadtler, & Salmerón, 2018).

4.3 Conclusie

In deze studie onderzochten we via twee experimenten de relatie tussen het vooraf bestuderen van een tekst en de antwoordscore, bij twee teksten met vragen. Blijkens het eerste experiment komen de uiteindelijke antwoorden grotendeels tot stand doordat deelnemers tijdens de vragen de tekst kunnen raadplegen. De niet-betekenisvolle uitkomsten uit het tweede experiment lijken erop te wijzen dat leerlingen door vooraf intensief de tekst te bestuderen nauwelijks tot niet beter presteren bij deze taak.

Echter, er is nog de nodige ruimte voor vervolgonderzoek. Vervolgonderzoek zou bijvoorbeeld kunnen verkennen hoezeer de uitkomsten van dit onderzoek afhangen van

de steekproef (vaardige lezers), het soort teksten (complexe opiniërende zakelijke teksten) en het vraagtype (situatiemodelvragen). Daarnaast zou vervolgonderzoek ook de relatie tussen de vooraf-tekstbestudering en de antwoordscore bij andere taken (sorteertaken, etc.) kunnen onderzoeken.

De twee in deze studie beschreven experimenten vonden plaats met toestemming van de Facultaire Ethische Toetsingscommissie Geesteswetenschappen (FETC-GW) van de Universiteit Utrecht.

Graag danken we daarnaast de coördinerende docenten Nederlands van de aan de experimenten deelnemende scholen: mevr. M. van Houwelingen-Hoebink (Katholieke Scholengemeenschap Etten-Leur), mevr. I. van Luijteleaer (Sint-Janslyceum, 's-Hertogenbosch), dhr. T. Robben (Kandinsky College, Nijmegen), dhr. Y. Schotanus (Rhedens Dieren & Rhedens Rozendaal), mevr. H. Stinstra (Ds. Pierson College, 's-Hertogenbosch), mevr. J. Stuulen (Amstelveen College), en dhr. C. van 't Veer (Scholengemeenschap Spieringshoek, Schiedam).

Noten

- 1 In eerste instantie waren beide manipulaties in één experimentopzet opgenomen, als onderdeel van de summatieve tekstbegripstoetsing in het schoolvak Nederlands. Echter, deelnemende docenten Nederlands stonden zeer terughoudend tegenover een vorm van summatieve toetsing waarbij de tekst niet meer bij de vragen beschikbaar zou zijn: ze achtten de kans hoog dat onverantwoord grote verschillen in score zouden ontstaan tussen condities. Daarom besloten we tekstbegrip in experiment I ('Raadpleegbaarheid') diagnostisch te toetsen en dit in experiment II ('Intensiteit') summatief te toetsen.
- 2 Bij onderzoekscondities IA ('Raadpleegbaar') en IIA ('Intensief-lezen') werden de leerlingen misleid: de leesteksten zouden bij de vragen worden ingenomen, wat uiteindelijk niet gebeurde. Het ligt voor de hand aan te nemen dat leerlingen bij de tweede test deze misleiding met een korreltje zout zullen hebben genomen.

In de analyse is daarom erop gecontroleerd of een mogelijk verminderde effectiviteit van deze misleiding haar sporen heeft nagelaten in de resultaten in het experiment. Dit bleek echter niet aantoonbaar het geval.

3 Ter vergelijk, gebaseerd op de gegevens van de laatste drie jaar: CSE-onderdelen met een beperkt aantal vragen zoals in deze twee sets kennen doorgaans een nogal lage betrouwbaarheid ($GLB < .40$), terwijl een CSE havo/vwo in zijn geheel, met gemiddeld drie tot vijf teksten, met gemiddeld dertig tot veertig vragen, een GLB van circa $.70$ kent. In Nederland is het gebruik van GLB (*Greatest Lower Bound*) als maat voor betrouwbaarheid overigens nog relatief onbekend, terwijl deze maat een accuratere schatting hiervan geeft dan de traditionele maat, Cronbach's α , die altijd een ondergrensschatting is van de werkelijke betrouwbaarheid. Zie hiervoor Sijsma (2009) en Trizano-Hermosilla & Alvarado (2016).

Literatuur

- Agarwal, P. K., & Roediger III, H. L. (2011). Expectancy of an open-book test decreases performance on a delayed closed-book test. *Memory*, 19(8), 836-852.
- Anmarkrud, Ø., Bråten, I., & Strømsø, H. I. (2014). Multiple-documents literacy: Strategic processing, source awareness, and argumentation when reading multiple conflicting documents. *Learning and Individual Differences*, 30, 64-76.
- Ardoin, S. P., Binder, K. S., Zawoyski, A. M., Nimocks, E., & Foster, T. E. (2019). Measuring the Behavior of Reading Comprehension Test Takers: What Do They Do, and Should They Do It?. *Reading Research Quarterly*, 54(4), 507-529.
- Artelt, C., Schiefele, U., & Schneider, W. (2001). Predictors of reading literacy. *European Journal of Psychology of Education*, 16(3), 363-383.
- Bax, S. (2013). The cognitive processing of candidates during reading tests: Evidence from eye-tracking. *Language Testing*, 30(4), 441-465.
- Bogaerds-Hazenbergh, S., Evers-Vermeul, J., & Bergh, H. van den. (Ingediend.) What Textbooks Offer and What Teachers Teach: An Analysis of the Dutch Reading Comprehension Curriculum.
- Branden, K. van den. (2003). Leesonderwijs in Vlaanderen: van hoera! naar aha! *Vonk*, 32(3), 12-29.
- Bråten, I., Stadler, M., & Salmerón, L. (2017). The role of sourcing in discourse comprehension. In M.F. Schober, D.N. Rapp, & M.A. Britt (eds.), *The Routledge handbook of discourse processes* (p. 141-166). Routledge.
- Breukink, C., Janssen, T., & Bergh, H. van den. (In voorbereiding). Adolescents' text comprehension of poetry and prose in grade 8 to 12; a comparative study with reading tasks and stimulated recall group interviews.
- Britt, M. A., & Rouet, J.-F. (2012). Learning with multiple documents: Component skills and their acquisition. In J. R. Kirby & M. J. Lawson (Red.), *Enhancing the quality of learning: Dispositions, instruction, and learning processes* (p. 276-314). Cambridge University Press.
- Britt, M. A., Rouet, J.-F., & Durik, A. M. (2018). *Literacy beyond text comprehension: A theory of purposeful reading*. Taylor & Francis.
- Broek, P. van den, & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes*, 54(5-6), 360-372.
- Broek, P. van den, Ridsen, K., & Husebye-Hartmann, E. (1995). The role of readers' standard for coherence in the generation of inferences during reading. In R. F. Lorch & E. J. O'Brien (Eds.), *Sources of coherence in text comprehension* (p. 353-373). Erlbaum.
- Broek, P. van den, Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading: Inferences and the on-line construction of a memory representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (p. 71-98). Erlbaum.
- Brunfaut, T., & McCray, G. (2015). *Looking into test-takers' cognitive processes whilst completing reading tasks: a mixed-method eye-tracking and stimulated recall study*. (ARAGS Research Reports). British Council: s.l..
- Cain, K., & Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. *British Journal of Educational Psychology*, 76(4), 683-696.
- Canton, J., Aler, T., Heemskerk, K., van der Westen, W., & Willemsen, K. (red.). (2013). *Effecten*

van sturing op discrepantie tussen de cijfers van het centraal examen: Onderzoek naar de sturing door schoolleiders en de effecten daarvan op het taalonderwijs op havo en vwo en het schoolexamen bij de talen. VLLT.

- Cataldo, M.G., & Oakhill, J. (2000). Why are poor comprehenders inefficient searchers? An investigation into the effects of text representation and spatial memory on the ability to locate information in text. *Journal of Educational Psychology*, 92(4), 791-799.
- Cerdán, R., Pérez, A., Vidal-Abarca, E., & Rouet, J.F. (2019). To answer questions from text, one has to understand what the question is asking: differential effects of question aids as a function of comprehension skill. *Reading and Writing* 32, 2111-2124.
- Cerdán, R., Vidal-Abarca, E., Martínez, T., Gilbert, R., & Gil, L. (2009). Impact of question-answering tasks on search processes and reading comprehension. *Learning and Instruction*, 19(1), 13-27.
- Cohen, A. D., & Upton, T. A. (2006). Strategies in responding to the new TOEFL reading tasks. *ETS Research Report Series*, 2006(1), i-162.
- Couzijn, M.J. (1995). *Observation of writing and reading activities: Effects on learning and transfer*. Dissertatie UvA. Amsterdam: UvA.
- Drop, W. (1983). *Instrumentele tekstanalyse : ten dienste van samenvatten, opstellen van begripsvragen, tekstverbeteren*. Groningen: Wolters-Noordhoff.
- Dunlosky, J., & Rawson, K. A. (2005). Why does rereading improve metacomprehension accuracy? Evaluating the levels-of-disruption hypothesis for the reading effect. *Discourse Processes*, 40, 37-55.
- ExamenOverzicht. (2020). *Oefenboek Nederlands Examenstof 2020-2021*. Z.p.: ExamenOverzicht.
- Expertgroep Doorlopende Leerlijnen Taal. (2008). *Over de drempels met taal*. Enschede: SLO.
- Ferrer, A., Vidal-Abarca, E., Serrano, M. Á., & Gilbert, R. (2017). Impact of text availability and question format on reading comprehension processes. *Contemporary Educational Psychology*, 51, 404-415.
- Frank, H., Grezel, J.E., Kooiman, M., Steenbergen, W., Vos, M., & Zetstra, H. *Nieuw Nederlands: 4-5-6 vwo*. Noordhoff.
- Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H.I. (2010). Understanding and integrating multiple science texts: Summary tasks are sometimes better than argument tasks. *Reading Psychology*, 31, 30-68.
- Gil, L., Martínez, T., & Vidal-Abarca, E. (2015). Online assessment of strategic reading literacy skills. *Computers & Education*, 82, 50-59.
- Grezel, J. E. (2014). Vaardigheid of kunstje? Eindexamen Nederlands onder vuur. *Onze Taal*, 83(5), 109-115.
- Gubbels, J., Langen, A. van, Maassen, N., & Meelissen, M. (2019). *Resultaten PISA-2018 in vogelvlucht*. Universiteit Twente.
- Gubbels, J. (2020). Moeite met evalueren en reflecteren: wat betekent dat? Didactiek Nederlands – Handboek. Geraadpleegd op 5 juli 2021 via <https://didactieknederlands.nl/handboek/2020/12/moeite-met-evalueren-en-reflecteren-wat-betekent-dat/>.
- Hendrix, T., & Hulshof, H. (1994). *Leesvaardigheid Nederlands: Omgaan met zakelijke teksten*. Coutinho.
- Higgs, K., Magliano, J. P., Vidal-Abarca, E., Martínez, T., & McNamara, D. S. (2017). Bridging skill and task-oriented reading. *Discourse Processes*, 54(1), 19-39.
- Hoogeveen, M. (2018). *Het leesonderwijs in primair en voortgezet onderwijs: een stand van zaken en curriculaire aanbevelingen*. SLO.
- Hoogeveen, M., & Bonset, H. (1998). *Het schoolvak Nederlands onderzocht*. Leuven/Apeldoorn: Garant.
- Kamalski, J. (2007). *Coherence marking, comprehension and persuasion: On the processing and representation of discourse*. Dissertatie UU. LOT.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge university press.
- Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363-394.
- Linthorst, T.R., & de Gloppe, K. (2015). De didactiek van begrijpend lezen in het voortgezet onderwijs: lesobservaties bij Nederlands en zaakvakken. *Pedagogische Studiën*, 92, 150-166.
- Liu, Z. (2005). Reading behavior in the digital environment. *Journal of Documentation*, 61(6), 700-712.
- LTN [Levende Talen Nederlands]. (2019a). *Evaluatie Examens Nederlands 2019*. Z.p.: Levende Talen Nederlands. Geraadpleegd op 10 juli 2021

via <https://lerarennederlands.nl/wp-content/uploads/2020/03/Evaluatie-Examens-Nederlands-2019-met-voorblad.pdf>

- LTN [Levende Talen Nederlands]. (2019b). *Advies curriculum Nederlands in de bovenbouw van het voortgezet onderwijs*. Z.p.: Levende Talen Nederlands. Geraadpleegd op 10 juli 2021 via <https://levendetalen.nl/wp-content/uploads/2019/09/LT-advies-curriculum-Nederlands-2019.pdf>
- McCrudden, M. T., Magliano, J. P., & Schraw, G. (2010). Exploring how relevance instructions affect personal reading intentions, reading goals and text processing: A mixed methods study. *Contemporary Educational Psychology*, 35(4), 229-241.
- McCrudden, M. T., Magliano, J. P., & Schraw, G. (2011). Toward an integrated view of relevance in text comprehension. In M. McCrudden, J. Magliano, & G. Schraw (eds.), *Text relevance and learning from text* (p. 395-414). Information Age Publishing.
- Maki, R. H. (1998). Test predictions over text material. In D. J. Hacker, A. C. Graesser, & J. Dunlosky (Eds.), *Metacognition in educational theory and practice* (p. 117-144). Erlbaum.
- Maki, R. H., Shields, M., Wheeler, A. E., & Zachilli, T. L. (2005). Individual differences in absolute and relative metacomprehension accuracy. *Journal of Educational Psychology*, 97, 723-731.
- Milliano, I. de, & Gelderen, A. van. (2020). Taakgericht lezen en schrijven van vmbo'ers; Verschillen in de aanpak van leerlingen en kwaliteit van hun taakuitvoering. *Levende Talen Tijdschrift*, 21(4), 3-15.
- O'Reilly, T., Feng, D. G., Sabatini, D. J., Wang, D. Z., & Gorin, D. J. (2018). How do people read the passages during a reading comprehension test? The effect of reading purpose on text processing behavior. *Educational Assessment*, 23(4), 277-295.
- OECD. (2019). PISA 2018: Assessment and Analytical Framework. Paris: PISA, OECD. <https://doi.org/10.1787/b25efab8-en>
- Onderwijsraad & Raad voor Cultuur. (2019). *Lees! Een oproep tot een leesoffensief*. Raad voor Cultuur/Onderwijsraad.
- Ozuru, Y., Best, R., Bell, C., Witherspoon, A., & McNamara, D. S. (2007). Influence of question format and text availability on the assessment of expository text comprehension. *Cognition and Instruction*, 25(4), 399-438.
- Ozuru, Y., Rowe, M., O'Reilly, T., & McNamara, D. S. (2008). Where's the difficulty in standardized reading tests: The passage or the question? *Behavior Research Methods*, 40(4), 1001-1015.
- Payne, S. J., & Reader, W. R. (2006). Constructing structure maps of multiple on-line texts. *International Journal of Human-Computer Studies*, 64(5), 461-474.
- Pearson, P. D., & Hamm, D. N. (2005). The assessment of reading comprehension: A review of practices—past, present, and future. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp 13-69). Lawrence Erlbaum Associates.
- Rooijackers, P., Silfhout, G. van, & Bergh, H. van den. (2021a). Met zijwieltjes leren fietsen in de Dorpsstraat: Waarom de leestaken in de les Nederlands vaak inhoudelijk tekortschieten. *Levende Talen Magazine*, 108(2), 4-9.
- Rooijackers, P., Silfhout, G. van, & Bergh, H. van den. (2021b). Observerend leren lezen op het vwo; reflecteren op tekstbegrip via oogbewegingsfilmpjes. *Levende Talen Tijdschrift*, 22(3), 3-16.
- Rooijackers, P., Silfhout, G. van, Schuurs, U., Mulders, I., & Bergh, H. van den. (2020). Lezen en antwoorden bij teksten met vragen: een cross-sectionele eye-trackstudie onder 52 vwo-leerlingen. *Pedagogische Studiën*, 97(4), 281-308.
- Rouet, J.-F., Britt, M.A., & Durik, A.M. (2017). RESOLV: Readers' Representation of Reading Contexts and Tasks. *Educational Psychologist*, 52(3), 200-215. DOI: 10.1080/00461520.2017.1329015
- Rouet, J.-F., Vidal-Abarca, E., Bert-Erboul, A., & Millogo, V. (2001). Effects of information search tasks on the comprehension of instructional text. *Discourse Processes*, 31(2), 163-186.
- Rupp, A. A., Ferne, T., & Choi, H. (2006). How assessing reading comprehension with multiple-choice questions shapes the construct: A cognitive processing perspective. *Language Testing*, 23(4), 441-474.
- Sabatini, J., O'Reilly, T., Wang, Z., & Dreier, K. (2018). Scenario-Based Assessment of Multiple Source Use. In J. L. Braasch, I. Bråten, & M.T. McCrudden (Eds.), *Handbook of multiple source use* (p. 447-465). Routledge.

- Salmerón, L., Vidal-Abarca, E., Martínez, T., Mañá, A., Gil, L., & Naumann, J. (2015). Strategic decisions in task-oriented reading. *The Spanish Journal of Psychology*, 18, 1-10.
- Schaffner, E., & Schiefele, U. (2013). The prediction of reading comprehension by cognitive and motivational factors: Does text accessibility during comprehension testing make a difference?. *Learning and Individual Differences*, 26, 42-54.
- Schroeder, S. (2011). What readers have and do: Effects of students' verbal ability and reading time components on comprehension with and without text availability. *Journal of Educational Psychology*, 103(4), 877-896.
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika* 74, 107-120. doi: 10.1007/s11336-008-9101-0
- SLO. (2017). *Leerplankundige analyse PISA 2015*. SLO.
- Traub, R. E. (1993). On the equivalence of the traits assessed by multiple-choice and constructed-response tests. In R. E. Bennet & W. C. Ward (Eds.), *Construction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment* (p. 29-44). Lawrence Erlbaum Associates.
- Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best alternatives to Cronbach's alpha reliability in realistic conditions: congeneric and asymmetrical measurements. *Frontiers in Psychology*, 7, 769.
- Vidal-Abarca, E., Mañá, A., & Gil, L. (2010). Individual differences for self-regulating task-oriented reading activities. *Journal of Educational Psychology*, 102(4), 1-10. doi: 10.1037/a0020062.
- Wesdorp, H. (1981). *Evaluatietechnieken voor het moedertaalonderwijs: een inventarisatie van beoordelingsmethoden voor de stelvaardigheid, het begrijpend lezen, de spreek-, luisteren discussievaardigheid*. Staatsuitgeverij.
- Wetering, T. van de, & Groenendijk, R. (2015). Van toetsgerichte lezers naar zelfstandige, analytische lezers: Geen trucje maar een vaardigheid. *Levende Talen Magazine*, 102(2), 4-8.
- Wolf, M. (2018). *Reader, come home: The reading brain in a digital world*. Harper.
- Yeari, M., Broek, P. van den, & Oudega, M. (2015). Processing and memory of central versus pe-

ripheral information as a function of reading goals: Evidence from eye-movements. *Reading and Writing*, 28(8), 1071-1097.

Auteurs

Patrick Rooijackers is promovendus aan de Universiteit Utrecht (via een Dudoc Alfa-beurs). Daarnaast is hij werkzaam als docent Nederlands aan het Sint-Janslyceum in 's-Hertogenbosch en als toetsdeskundige Nederlands bij Cito. **Gerdineke van Silfhout** is werkzaam bij SLO als curriculumexpert en is daarnaast als programmamanager Actualisatie examenprogramma's verantwoordelijk voor het coördineren van de bijstelling van de examenprogramma's bovenbouw vo. **Uriël Schuurs** was tot voor kort toetsdeskundige Nederlands bij Cito. Momenteel werkt hij bij een bureau voor communicatietraining en -advies. **Huub van den Bergh** is als hoogleraar Didactiek en toetsing van het taalvaardigheidsonderwijs werkzaam aan de Universiteit Utrecht. *Correspondentieadres:* Patrick Rooijackers, Universiteit Utrecht, Trans 10, 3512 JK Utrecht; E-mail: p.j.h.rooijackers@uu.nl

Abstract

The relationship between reading texts in advance and answering comprehension questions
Two experiments among 11th-grade pre-university students

In Dutch education 'texts with questions' are often used to train reading skills. One of the tacit assumptions inherent to this task is that students who read the text well will answer more questions correctly.

The relationship between reading and answering process in this type of task is investigated, among 11th-grade pre-university students, in two experiments. In experiment I ($N = 174$) it was investigated to what extent *text availability* during answering process influences comprehension scores. In experiment II ($N = 398$) it was investigated to what extent *intensity of initial text reading* influences comprehension scores, if

texts remain available during answering process. Results. Experiment I: students who still had texts available during answering process had markedly higher comprehension scores than students who had not. Experiment II: students who initially had read texts intensively, do not have higher answering scores than students who had not. The question is therefore raised to what extent this task fosters better text comprehension.

Keywords: text comprehension, pre-university students, text with questions, school subject Dutch, expository texts

Appendix 1. Een selectie uit de twee sets ‘teksten met vragen’

Set ‘tekst met vragen’ 1, afkomstig uit het CSE Nederlands vwo 2019-1

De Kunst en de Regels (selectie)

[1] Vorige week woonde ik een opvoering bij van *Mariken van Nimwegen*, want mijn zoon was er figurant (een heel goede, concludeerden we achteraf). Zo leerde ik in de huise-lijke Stadsschouwburg van Groningen het verhaal van Mariken kennen. Een lacune in mijn academische opleiding die tegenwoordig niet meer mogelijk is: het internet wemelt van de leesdossiers waarin scholieren je een samenvatting plus exegese geven. Het is maar dat u het weet, mocht u ook iets in te halen hebben.

[2] De kernregel uit Marikens gedicht ‘Door d’onkonstige gaat die konste verloren’² is een noodkreet van begin zestiende eeuw die vreemd eigentijds aandoet. Mariken heeft het over de ‘seven vrie consten: rethorijcke, musijcke, logica, gramatica ende geometrie, arismetica ende alkenie’.³ Ik zou daar vijf eeuwen later de kunst van goed onderwijs, onderzoek, zorg en zo meer aan willen toevoegen. Wie daarvan de kunst niet verstaat, kan volgens Mariken slechts schade aanrichten. Het is dus een oude, maar helaas brandende kwestie.

[3] Een paar voorbeelden uit de zorg. Een huisarts meldt dat hij vijf A4’tjes moest invullen om dieetvoeding voor een terminale patiënt te bemachtigen. De minutenregistratie van zorg ‘achter de voordeur geleverd’ is berucht. Hoe moet de tijd van de ingang van een flat naar de patiënt op de tiende etage geregistreerd worden – wat geldt hier als voordeur? Of het consult met de huisarts op diens kantoor, wat dus geen achter de voordeur geleverde zorg is? De regels van hen die de kunst niet verstaan, staan hier de kunst van goede zorg in de weg.

[4] Het nare is dat de regels het snelst groeien waar ze de meeste schade aanrichten. Waarom? Noem het de Wet van Mariken. Ze draait om het begrip *tacit knowledge*,⁴ een idee van Michael Polanyi. Het gaat om ervaringskennis die niet in een boekje is op te schrijven en niet overdraagbaar is door een paar uur instructie of inwerken. Vakken vullen, bollen pellen en boompjes stekken vergen vrijwel geen ervaringskennis (ik kom uit de buurt van Boskoop, dus spreek uit ervaring). Je kunt vanaf dag één aan de slag. Voor meesters en juffen, wetenschappers en verzorgenden is dat wel anders – en de lijst is natuurlijk veel langer. Want wat is een goede docent, verzorgende, of onderzoeker? Dat is lastig te zeggen – en dus is er volop ruimte voor regels die dat specificeren. (Denkt u dat zulke regels er ook zijn voor onkruidwieders en vakkenvullers? Welnee. Daar zie je direct of het werk goed gedaan is.)

[5] Tegelijk zijn regels het schadelijkst juist daar waar veel ervaringskennis is. Want de regels van onwetende buitenstaanders dwingen de ervaringsdeskundigen hun werk naar de verkeerde maatstaven te beoordelen en uit te voeren, en daar bovendien veel tijd aan te besteden. Denk aan het construeren van ranglijstcriteria, het tellen van stappen en minuten, en het aanvinken van vakjes in grote tabellen (ik spreek opnieuw uit ervaring). Meestal zonde van de tijd, en geen kunst. Maar er is meer (...)

Naar: Dirk Bezemer

Vragen bij De Kunst en de Regels (selectie)

- 2p **1** In alinea 4 van de tekst wordt het begrip *tacit knowledge* geïntroduceerd. Welke vier van onderstaande groepen zouden vooral beschikken over *tacit knowledge*, gelet op de strekking van de tekst? Kies uit: bollenpellers, buitenstaanders, ‘konstigen’, managers, onderwijzers, ‘onkonstigen’, patiënten, scholieren, verzekeraars, verzorgenden en wetenschappers.
- 2p **2** De titel van de tekst luidt ‘De Kunst en de Regels’. Leg uit hoe de Kunst zich verhoudt tot Regels, gelet op de strekking van de tekst. Geef antwoord in één zin.
- 1p **3** Wat is het tekstdoel van de tekst, gelet op de strekking van de tekst? ‘De Kunst en de Regels’ is een
- A beschouwing, want de kernregel uit Mariken van Nimwegen wordt vanuit verschillende invalshoeken in een moderne context geplaatst.
 - B beschouwing, want van diverse beroepen in onderwijs, onderzoek en zorg wordt bekeken of de Wet van Mariken erop van toepassing is.
 - C betoog, want aan de hand van de Wet van Mariken wordt gepropageerd dat het snel groeiende aantal regels bij ervaringsberoepen een kwalijke zaak is.
 - D betoog, want door het geven van schrijnende voorbeelden wordt bepleit dat we werk niet vooral als bron van inkomen moeten zien.

In de tekst komen praktijkgerichte beroepen en beroepen met veel *tacit knowledge* ter sprake. Tussen beide **beroepsoorten** staan enkele **inhoudelijke tegenstellingen** centraal.

- 3p **4** Noem drie van deze inhoudelijke tegenstellingen. Vul dit in in onderstaand schema.

praktijkgerichte beroepen	beroepen met veel <i>tacit knowledge</i>
1a)	1b)
2a)	2b)
3a)	3b)

Tekst De Kunst en de regels: antwoorden

1 maximumscore 2

'konstigen', onderwijzers, verzorgenden en wetenschappers

indien vier goed 2

indien drie goed 1

indien twee of minder goed 0

2 maximumscore 2

De kern van een goed antwoord is:

De Kunst wordt gehinderd door de Regels / Vaak hebben regels een negatief effect op werkzaamheden waar ervaringskennis voor nodig is / Hoe meer regels, des te eerder raakt kunst in het gedrang

3 C 1

4 maximumscore 3

drie van onderstaande inhoudelijke tegenstellingen:

praktijkgerichte beroepen	ervaringsgerichte beroepen
a) geen/nauwelijks regels	b) veel regels
a) vrijwel geen ervaringskennis / vrijwel geen <i>tacit knowledge</i>	b) veel ervaringskennis / veel <i>tacit knowledge</i>
a) gemakkelijk aan te leren	b) lang leertraject noodzakelijk
a) eenvoudig beoordeelbaar	b) lastig beoordeelbaar
a) geen schade als gevolg van regels	b) negatieve effecten als gevolg van regels
a) (kennis is) gemakkelijk overdraagbaar	b) (kennis is) lastig overdraagbaar

per juist benoemde tegenstelling 1

De wilskrachtige is een gewoontedier (selectie)

[1] Wat eet je liever: een appel of een chocoladereep? Worteltjes of patat? De appel en de worteltjes zijn natuurlijk de gezondste keuze: caloriearm, vol vezels en vitamines. Zoete troep en vette hap zijn nergens goed voor, maar wél lekker. Dus wat kies je? Utrechtse psychologen legden hun proefpersonen onlangs keuzes als deze voor. En wat bleek?

(2) Mensen die van nature wilskrachtig zijn, kiezen vaker voor de appel en de worteltjes dan de minder wilskrachtigen, de uitlappers en de flapuits. Ze kiezen niet alleen vaker, maar ook sneller voor het verstandige alternatief. Ze twijfelen minder, voeren slechts een korte of helemaal geen innerlijke strijd.

[3] Dit nieuwe onderzoek sluit aan bij eerdere bevindingen van het *Selfregulation Lab*, zoals een groep psychologen aan de Universiteit Utrecht zichzelf noemt. Sinds 2002 bestuderen zij de wilskracht: het vermogen om impulsen te onderdrukken met het oog op een grotere beloning in de toekomst. Wie slank wil blijven, bezwijkt niet voor patat, gaat niet (te vaak) de kroeg in. Naast intelligentie is zelfcontrole een belangrijke sleutel tot zowel persoonlijk als maatschappelijk succes.

[4] Onder leiding van hoogleraar gezondheidspsychologie Denise de Ridder deed deze groep enkele jaren geleden al een interessante ontdekking: het kost wilskrachtige mensen vaak heel weinig moeite verleiden het hoofd te bieden. Ze zijn geen mentale

krachtpatsers die hun verlangens naar junkfood, alcohol en overspelige seks steeds met geweld weten te onderdrukken: ze vóelen die verlangens eenvoudigweg niet of veel minder dan hun zwakkere broeders en zusters. Innerlijke conflicten weten ze te omzeilen met routines.

[5] Daar hebben wilskrachtige mensen er namelijk relatief veel van, blijkt uit studies van het Utrechtse team. Ze hebben meer gezonde en minder ongezonde gewoontes dan de patateters en de kroegtijgers. Voor het ontbijt nemen ze yoghurt met muesli, rond koffietijd op kantoor een appel en geen gevulde koek. Dankzij de automatische piloot hoeven ze hun zelfbeheersing niet in stelling te brengen, zegt De Ridder. Hun tankje wilskracht blijft gevuld voor nood-gevallen.

[6] Uit onderzoek blijkt dat dit tankje bij iedereen leeg kan raken. Vergelijk wilskracht met een spier: als je die flink belast, wordt hij moe. Ook goedgepaste bodybuilders zijn weleens uitgeput, en moeten dan rusten om weer op krachten te komen.

[7] Het beeld van de spier is van Roy Baumeister, hoogleraar sociale psychologie aan de Vrije Universiteit in Amsterdam. Hij is ook de uitvinder van de term 'egodepletie': een verminderd vermogen om je gedrag en gevoelens te controleren. Bij egodepletie ben je (tijdelijk) door je voorraad wilskracht heen en voel je intensere verlangens naar verboden vruchten. (...)

Naar: Iris Pronk

Vragen bij De wilskrachtige is een gewoontedier (selectie)

- 2p **1** In de openingsalinea's is sprake van **recent** Utrechts onderzoek. Welke nieuwe inzichten over wilskrachtige mensen gaf dit recente Utrechtse onderzoek? Geef antwoord in een of meer volledige zinnen.

In de eerste helft van de tekst wordt gesproken over recent én ouder onderzoek onder leiding van hoogleraar gezondheidspsychologie Denise de Ridder.

- 1p **2** Welke van de onderstaande omschrijvingen geeft de beste samenvatting weer van de resultaten van dat onderzoek?

Wilskrachtige mensen

- A** hebben zoveel gezonde gewoontes dat ze de lokroep van ongezonde verleidingen minder horen en daardoor hun wilskracht nauwelijks hoeven aan te spreken.
- B** onderdrukken hun verlangens naar ongezonde verleidingen door snel over te stappen op een automatische piloot die hen ervoor behoedt toe te geven aan die verleidingen.
- C** voelen wel degelijk een bepaalde honger naar junkfood en alcohol, maar door een beroep te doen op hun voorraad wilskracht weten ze die drang om te zetten in gezonde gewoontes.
- D** zijn vatbaarder voor verlangens dan mensen die bepaalde routines opgebouwd hebben omdat die laatste groep meer succeservaringen heeft geboekt op het gebied van zelfcontrole.

In de tekst valt de term 'egodepletie'.

- 1p **3** Welk van de onderstaande omschrijvingen geeft het best weer wat egodepletie te maken heeft met wilskracht?
- A** Als je wilskracht niet voldoende gebruikt, krijg je een tekort aan egodepletie.
- B** Als je je voorraad wilskracht hebt opgebruikt, kun je spreken van egodepletie.
- C** Als je je wilskracht op de proef stelt, voorkom je een toestand van egodepletie.
- D** Als je wilskracht regelmatig traint, kom je sneller in aanmerking voor egodepletie.

In de eerste helft van de tekst wordt uitgelegd dat er voordelen zitten aan het worden van een gewoontedier.

- 3p **4** Noem de voordelen uit tekst 2.
Geef antwoord in een of meer volledige zinnen.

Tekst De wilskrachtige is een gewoontedier: antwoorden

1 maximumscore 2

De kern van een goed antwoord is:

- Wilskrachtige mensen kiezen vaker het verstandige alternatief / beheersen zich vaker 1
- (en) doen dat sneller / twijfelen veel korter over hun keuze / nemen sneller een besluit 1

2 A 1

3 B 1

4 maximumscore 3

De kern van een goed antwoord is:

- a) Het kost weinig moeite om verleidingen te weerstaan 1
- b) Ze kunnen innerlijke conflicten omzeilen (met routines) 1
- c) Hun wilskracht blijft beschikbaar voor noodgevallen / Hun tankje wilskracht blijft gevuld voor noodgevallen 1