

Tekster ontleed:

Over de waarde, bouwstenen, werkzame bestanddelen en basis van een effectief schrijfprogramma voor de bovenbouw van de basisschool

KEES DE GLOPPER

Op 2 september 2016 verdedigden Monika Koster en Renske Bouwer aan de Universiteit Utrecht met succes hun proefschrift *Bringing Writing Research into the Classroom*: zij kregen er beiden het *judicium cum laude* voor. In *Schrijffonderzoek naar het klaslokaal*, de Nederlandse titel van de dissertatie, doen zij verslag van een onderzoek naar de effecten van Tekster, een nieuw ontwikkeld lesprogramma voor leerlingen in het basisonderwijs.* Het proefschrift van Bouwer en Koster is de moeite van het lezen meer dan waard. Om te beginnen wordt het oog van de lezer gestreeld door het uiterlijk ervan: de omslag en het binnenwerk zijn fraai door de vele subtiele illustraties en accenten en de lay-out is aantrekkelijk. Het kloeke boek is verder uiterst leesbaar geschreven. Maar het belangrijkste is de inhoud: Bouwer en Koster doen verslag van twee belangwekkende experimenten en van een aantal omliggende studies. Redenen genoeg om het hier uitgebreid te bespreken en stil te staan bij de betekenis van hun studie voor het onderzoek naar en de praktijk van het schrijffonderwijs op de basisschool.

De kern van de dissertatie wordt gevormd door twee interventiestudies: een eerste studie waarin de interventie werd uitgevoerd

door leerkrachten (N=76) die getraind zijn door experts (lees: de onderzoekers) en een tweede studie waarin de interventie nog eens is uitgevoerd, maar nu met een ingreep in de trainingscomponent. De helft van de leerkrachten (N=31) werd getraind door experts. Een deel van hen trainde vervolgens collega-leerkrachten (N=37) die het programma daarna uitvoerden.

Het eerste experiment

In hoofdstuk 5 bespreken Bouwer en Koster de effecten van het programma Tekster op de schrijfprestaties van leerlingen uit groep 6, 7 en 8. Tekster biedt de leerlingen een algemene strategie voor de aanpak van schrijftaken waarin plannen, schrijven en reviseren centraal staan. Voor groep 6 is VOS ontwikkeld (Verzinnen, Ordenen, Schrijven), voor groep 7 DODO (Denken, Ordenen, Doen, Overlezen) en voor groep 8 EKSTER (Eerst nadenken, Kiezen & ordenen, Schrijven, Teruglezen, Evalueren, Reviseren). Deze strategieën komen in elke les aan bod. De leerlingen uit groep 6, 7 en 8 (1420 in getal, afkomstig uit 27 scholen) volgden 16 Teksterlessen in een tijdsbestek van acht weken.

Het experiment is opgezet volgens het *switching replications design* (Shadish, Cook, &

Campbell, 2002). De eerste groep leerkrachten voerde Tekster uit tussen het eerste en tweede meetmoment terwijl de tweede groep leerkrachten tijdens die periode het reguliere schrijfprogramma uit de eigen taalmethode volgde. Na het tweede meetmoment werd Tekster door de tweede groep leerkrachten uitgevoerd, terwijl groep een nu het reguliere schrijfprogramma weer oppakte.

Op meetmoment 1, 2 en 3 werd de schrijfvaardigheid van de leerlingen met drie verschillende taken getoetst: een verhaal, een beschrijving en een overtuigende brief. Zo schreef iedere leerling negen teksten. Alle teksten, een ware berg van bijna 13.000 stuks, werden beoordeeld door jury's van drie beoordelaars.

De resultaten van het eerste experiment laten effecten op de schijfprestaties van de leerlingen zien. Tussen het eerste en het tweede meetmoment nemen de prestaties van de leerlingen uit de eerste groep toe. Hetzelfde geldt tussen het tweede en derde meetmoment voor de prestaties van de tweede groep leerlingen. Generaliserend over klassen, leerlingen en taken is het effect 0,32, wat neerkomt op een derde standaarddeviatie. Na afloop van de interventie gaat het effect niet verloren, zo laten de gegevens uit de eerste groep zien. Wel is er tussen meetmoment 2 en 3 in groep 7 en 8 sprake van een neerwaartse trend, zoals te zien valt in figuur 5.2 op pagina 105 van het proefschrift.

Gegevens uit logboeken en klasobservaties laten zien dat de leerkrachten goed met Tekster overweg konden. Wel deden zich tussen leerkrachten grote verschillen voor in het toepassen van de belangrijkste componenten van het programma. Ook kwamen lang niet alle leerkrachten aan alle lessen toe: van de 16 lessen werden er gemiddeld 10 gegeven.

Het tweede experiment

In hoofdstuk 9 beschrijven Bouwer en Koster de tweede interventiestudie waarin, zoals

gezegd, sprake is van doelbewuste variatie in de scholing van de leerkrachten: door experts of door collega-leerkrachten die eerder door experts zijn getraind. De opzet van het tweede experiment komt verder nauw overeen met die van het eerste. Een belangrijk verschil is wel de looptijd: de 16 lessen werden nu niet in 2 maar in 4 maanden aangeboden. Deze meer gespreide aanbieding van Tekster ging gepaard met minder lesuitval: van de 16 lessen werden er nu gemiddeld 13 gegeven.

De resultaten van het tweede experiment laten opnieuw effecten zien op de schrijfprestaties van de leerlingen uit groep 6, 7 en 8 (1365 in getal, afkomstig uit 25 scholen). Generaliserend over klassen, leerlingen en taken bedraagt het effect van Tekster nu 0,55, iets meer dan een halve standaarddeviatie. Van een neerwaartse trend in de schrijfprestaties na afloop van de interventie is ditmaal geen sprake.

Van de aard van de scholing van de leerkrachten ging geen effect uit op de prestaties van hun leerlingen. Wel bleek dat de leerkrachten die door experts waren getraind wat positiever oordeelden over wat ze in de training hadden geleerd dan de leerkrachten die door collega's waren getraind. Uit logboeken, vragenlijsten, klasobservaties en interviews bleek verder dat de leerkrachten de inhoud van hun lessen en hun manier van lesgeven hebben aangepast in de door Tekster bedoelde richting. Ten slotte bleek dat beide groepen leerkrachten na afloop van het experiment positiever waren over het geven van schrijffonderwijs en zich bekwaamer voelden om zwakke en sterke leerlingen te leren schrijven.

De waarde van Tekster

De opzet, uitvoering en analyse van de beide experimenten zijn voorbeeldig te noemen. Het gekozen design vormt een goede waarborg voor de interne validiteit van de

experimenten: plausibele alternatieve verklaringen voor de geconstateerde effecten vallen niet goed te bedenken. Tekster werkt en doet dat bij uitvoering op een zeldzaam grote schaal, en ook wanneer de onderzoekers de training van de leerkrachten die het programma uitvoeren uit handen geven. Dit maakt dat vragen naar de waarde, bouwstenen, werkzame bestanddelen en basis van Tekster opportuun zijn. Hoe moeten we de effecten van het lesprogramma in termen van leerwinst waarderen? Hoe is Tekster eigenlijk opgebouwd? Wat zijn de werkzame bestanddelen van het programma? En welke visie op schrijfvaardigheid ligt aan de basis van Tekster?

De onderzoekers geven zelf een antwoord op de vraag naar de waarde van Tekster. Het effect van de eerste studie, waarin de leerlingen gedurende 8 weken gemiddeld 10 lessen volgen, stellen zij gelijk aan een leerwinst in schrijfvaardigheid van ruim een half leerjaar. De schaal en maat voor dit groeitempo ontlenen Bouwer en Koster aan de prestatieverschillen tussen de leerlingen van groep 6, 7 en 8 op het eerste meetmoment. Het effect van de tweede studie met zijn gemiddelde van 13 lessen gedurende 16 weken wordt op dezelfde wijze naar leerwinst vertaald en correspondeert dan met een groei van maar liefst anderhalf leerjaar. Een half jaar leerwinst of anderhalf jaar leerwinst? Dat maakt nogal uit. Op dit verschil wordt in het proefschrift opmerkelijk genoeg niet ingegaan. En in hoeverre het gerechtvaardigd is om verschillen op de schrijftaken uit het onderzoek te gebruiken om de effecten van de interventie te vertalen naar het 'gewone' schrijfonderwijs, is een vraag die in het proefschrift niet wordt gesteld. Toch is dit geen onbelangrijke vraag, want voor bepaling van de waarde van het programma is generalisatie nodig van de in het onderzoek geconstateerde effecten naar te verwachten leerwinst in de reguliere klaspraktijk. Daarbij dienen zich allerlei vragen aan; vragen die in het proefschrift niet

worden gesteld en waarop het begin van een antwoord ontbreekt. In hoeverre zijn de effecten afhankelijk van de gebruikte schrijftaken en van de mate waarmee leerlingen met zulke taken vertrouwd zijn, ook al voorafgaand aan de instructie? Nieuwe bezems vegen immers schoon en met nieuwe griffels en schone leien is het goed schrijven. En hoe zouden de effecten eruitzien wanneer leerlingen in groep 7 en groep 8 niet voor de eerste maal met het programma in aanraking zouden komen, maar wanneer het schrijfprogramma zou voortbouwen op wat er in eerdere groepen aan bod kwam en geleerd is? Worden, met andere woorden, de per leerjaar geconstateerde effecten bij longitudinale implementatie van Tekster op elkaar gestapeld, of is misschien de oogst van het eerste jaar de rijkste?

Het antwoord op de vraag naar de waarde van Tekster is, zo gezien, niet direct voorhanden. Dat laat onverlet dat het programma onmiskenbaar effectief is. En daarom zijn de vragen naar de bouwstenen, werkzame bestanddelen en basis ervan relevant.

De bouwstenen van Tekster

De bouwstenen voor Tekster hebben Bouwer en Koster ontleend aan een meta-analyse van interventiestudies die zich richten op de verbetering van schrijfvaardigheid van leerlingen van groep 6 tot en met 8 in het reguliere basisonderwijs. Hoofdstuk 2, dat verslag doet van deze analyse, is gepubliceerd als tijdschriftartikel (Koster, Tribushinina, De Jong, & Van den Bergh, 2015). De meta-analyse heeft betrekking op 32 verschillende experimentele studies die 59 vergelijkingen bieden van vormen van schrijfonderwijs waarvan de focus van instructie verschilt. Om greep te krijgen op mogelijke verschillen in effectiviteit zijn de studies in maar liefst 10 groepen geïdentificeerd, al naar gelang hun dominante focus van instructie: schrijfstrategieën, tekststructuur,

peer-interactie, tekstevaluatie, doelen stellen, feedback, grammaticale instructie, revisie, prewriting en procesaanpak. Dat classificeren is een hachelijke onderneming omdat interventieprogramma's zelden of nooit op een paard wedden, ook niet wanneer de voor een programma verantwoordelijke onderzoekers dat beweren te doen. Veel van de onderscheiden categorieën sluiten elkaar ook niet uit of gaan zelfs noodzakelijkerwijs gepaard met andere. Peer-interactie moet immers over het schrijfproduct of het schrijfproces gaan, feedback heeft ergens betrekking op en strategieën voor het plannen van een tekst roepen elementen van de tekststructuur op. Door het grote aantal onderscheidingen zijn de meeste categorieën met slechts enkele verschillende studies gevuld. Dat maakt dat de schattingen van de effecten weinig precies zijn en erg afhankelijk van de beschikbare voorbeelden.

Uit de meta-analyse komen niettemin 5 aanpakken als de meest effectieve naar voren: doelen stellen, strategie-instructie, feedback, tekststructuurinstructie en peer-interactie. Samen vormen deze de bouwstenen voor Tekster. Maar daarmee is niet alles gezegd. Aanvullende bouwstenen zijn er namelijk ook, in de vorm van observationeel leren, zelfregulatie en afname van hulp aan c.q. toename van autonomie van de schrijver. Al deze elementen zijn gecombineerd in een samenhangend programma dat gericht is op zowel het schrijfproces (strategieën), als het schrijfproduct (vooral tekststructuur en met doel en publiek samenhangende evaluatiecriteria) en de schrijver zelf (zelfregulatie). Modelleren, uitleggen en feedback en hulp bieden zijn de centrale activiteiten voor de leerkracht. Corresponderende leerlingactiviteiten zijn observeren, discussiëren, vergelijken, luisteren, aantekeningen maken, evalueren, feedback geven, strategieën toepassen en reguleren.

Hoe zou Tekster in een toekomstige meta-analyse moeten worden geïdentificeerd? Als strategieonderwijs? Dat zou waarschijnlijk het

antwoord van Bouwer en Koster zijn, want als kern van Tekster beschouwen zij de algemene strategie voor de aanpak van schrijftaken die correspondeert met de elementen van het schrijfproces: plannen, schrijven en reviseren. Maar andere bouwstenen lijken evenzeer onmisbaar: het leren van voorbeelden en het verkrijgen van inzicht in de samenhang tussen doel, publiek en inhoud en structuur van teksten. Tekster valt daarom vanuit het standpunt van een meta-analytisch onderzoek misschien het beste als een hybride aanpak karakteriseren. Van een praktijkstandpunt bezien is het verstandig breed en evenwichtig.

Wat Tekster leerlingen niet biedt, wordt bij lezing van het proefschrift en bestudering van het lesmateriaal ook duidelijk. Feedback van de leerkracht lijkt een ondergeschikte rol te spelen; peer feedback is wel een integraal onderdeel van de lessen. Ook schriftelijk formuleren staat niet centraal. Aan de talige middelen die leerlingen moeten inzetten om teksten te creëren die doel- en publiekgericht zijn, wordt in Tekster nauwelijks of geen expliciete aandacht besteed. Voor en na het schrijven staan de inhoud en de structuur van de tekst centraal. In termen van het promotieonderzoek van Hoogeveen (2012), een interventiestudie naar de gecombineerde effecten van genrekennis en peer response, gaat het in Tekster eerder om algemene aspecten van communicatief schrijven (AACS) dan om specifieke genrekenmerken (SGK). Hoogeveen toetste in haar onderzoek de veronderstelling dat instructie in het gebruik van talige middelen leerlingen uit groep 8 concrete handvatten biedt voor het plannen, schrijven, bespreken en reviseren van verhalende en instructieve teksten. Volgens haar studie is instructie in het gebruik van indicatoren voor plaats en tijd en in hun functie in verschillende genres effectief: de SGK-leerlingen schrijven betere teksten en reviseren meer dan de AACS-leerlingen. De exploratieve studie van Broekhuis en De Glopper (2016) levert in

vervolg op de studie van Hoogeveen (2012) aanwijzingen op dat instructie in specifieke genrekenmerken positief effecten heeft op de planningsactiviteiten van leerlingen uit groep 8 tijdens het schrijven van een verhaal.

Hoewel meesterschap zich in beperkingen toont, valt er veel voor te zeggen om naast (of in een volgende versie van) Tekster aandacht te besteden aan de talige middelen die nodig zijn bij het schrijven van teksten in verschillende genres. Niet alleen door het aanleren van strategieën, maar ook langs deze weg kan de complexiteit van schrijftaken immers worden teruggedrongen en kunnen leerlingen worden geholpen met het leren schrijven op een wijze die meer aansluit bij het kennis-transformatiemodel dan bij het kennisweergavemodel (Bereiter & Scardamalia, 1987).

Verrijking van Tekster lijkt ook mogelijk op het punt van structuurkennis. Tekster zet op dit punt namelijk erg breed en waarschijnlijk daardoor weinig diep in. Zo blijft de structuurkennis met betrekking tot verhalen in groep 8 beperkt tot het aanbieden van de onderscheiding tussen inleiding, kern en slot, waarbij wordt uitgelegd dat de inleiding het begin is van het verhaal waarin meestal kort verteld wordt over wie het verhaal gaat en waar en wanneer het gebeurt, dat de kern het middenstuk is waarin je de rest van het verhaal vertelt maar nog niet de afloop. Die vertel je in het slot. Het valt verder te overwegen om anders dan Tekster per leerjaar in te zoomen op een kleiner aantal genres dan de (ongeveer) 10 die er nu in het materiaal voor groep 6, 7 en 8 aan bod komen.

Ook op andere punten kan Tekster worden uitgebreid of verbonden met andere vormen en inhoud van schrijfonderwijs. Een snelle vergelijking met de aanpak die Van Norden (2014) schetst, maakt dat duidelijk. Te denken valt aan informatie verzamelen en verwerken, inhoudelijke startgesprekken voeren, groeps-gewijs hardop denkend een tekst schrijven en bruggen slaan naar de zaakvakken, om maar een paar punten te noemen.

De werkzame bestanddelen van Tekster

Wat zijn de werkzame bestanddelen van Tekster? Aan welke aspecten van de inhoud en de focus van het lesprogramma vallen de effecten toe te schrijven? Het proefschrift bevat een drietal hoofdstukken die op deze vraag een antwoord kunnen bieden.

In hoofdstuk 7 wordt ingegaan op de effectiviteit van de eerste stap in de schrijfstrategie: het genereren en ordenen van ideeën voor de inhoud van de tekst. Bouwer en Koster beschrijven een analyse van het gebruik dat de leerlingen maakten van het kladpapier bij hun schrijftaken. Het onderzoeksmateriaal is afkomstig uit de tweede interventiestudie en betreft een steekproef uit het schrijfwerk (N=3680). De analyse heeft betrekking op de schriftelijke neerslag van de prewriting-activiteiten die vooraf gingen aan het schrijven van narratieve teksten (groep 6), descriptieve teksten (groep 7) en persuasieve teksten (groep 8) op de drie verschillende meetmomenten. Het onderwijs in plannen heeft effect. Na de interventie schreven de leerlingen minder vaak (alleen maar) een kladversie van hun tekst en produceerden ze voorafgaand aan het schrijven van hun teksten vaker al dan niet geordende ideeën. Prewriting vertoont verder samenhangen met de kwaliteit van de teksten van de leerlingen. Dit geldt overigens niet alleen voor de strategie die is onderwezen (het genereren en organiseren van ideeën), maar ook voor het schrijven van een kladversie. De analyses van deze samenhangen betreffen, voor zover uit het proefschrift valt op te maken, alle materiaal, ongeacht het meetmoment of de conditie. Op de causaal geformuleerde conclusie dat het gebruik van de strategie leidt tot teksten met een hogere kwaliteit, valt daarom wel wat af te dingen.

Aan de effecten van reviseren, het sluitstuk van de schrijfstrategie, wordt door Bouwer en Koster geen aandacht besteed. Dat is jammer,

want daardoor blijft het inzicht in de werkzame bestanddelen van het lesprogramma beperkt. De onderzoekers merken deze beperking zelf op en geven, ter verantwoording, aan dat onderzoek bestaat uit het maken van keuzes. Dat is waar, en daarom is de vraag gerechtvaardigd of zij op dit punt niet beter een andere keuze hadden kunnen maken. Een deelstudie naar revisie had bijvoorbeeld de plaats kunnen innemen van het onderzoek dat wordt beschreven in hoofdstuk 8. Dit hoofdstuk doet verslag van een experiment waarin 14 basisschoolleerkrachten feedback gaven op 15 leerlingteksten in drie verschillende genres: een verhaal, een persuasieve brief en een argumentatief tekstje. De 5 teksten per genre waren verschillend in kwaliteit: heel goed, heel slecht, of iets er tussenin. De leerkrachten verschilden onderling sterk in de manier waarop zij feedback gaven. Hun feedback had vooral betrekking op tekortkomingen in de teksten. In hun feedback kwamen lagere orde- en hogere orde-aspecten gelijkmatig aan bod. Op goede en slechte teksten formuleerden de leerkrachten gemiddeld genomen evenveel lagere orde- en hogere orde-feedback. De goede teksten kregen wel vaker positief commentaar. Aan deze globale overeenkomsten en verschillen in de feedback op goede en slechte teksten verbinden de onderzoekers de conclusie dat leerkrachten hun feedback onvoldoende afstemmen op het niveau en de behoefte van leerlingen. Voor deze gevolgtrekking lijken echter andere analyses en andere data nodig. In plaats van de globale vergelijkingen tussen de aard van de feedback op goede en slechte teksten zouden specifieke analyses passen die nagaan welke sterke en welke zwakke punten de teksten bevatten en welke daarvan op welke wijze door de leerkrachten worden geadresseerd. Het zou verder beter zijn als de feedbackpraktijken van de leerkrachten in hun eigen praktijk onderzocht zouden worden, waarbij ze zich buigen over de teksten van hun eigen leerlingen. Zulke data zouden meer licht

kunnen werpen op de kwaliteit en effectiviteit van leerkrachtfeedback.

Naast de algemene schrijfstrategie vormt structuurkennis een centrale component van het lesprogramma. Het zou daarom interessant zijn om te weten welke kennis over schrijven en teksten de leerlingen onder invloed van Tekster verwerven. Wat leren zij over schrijfprocessen en schrijfproducten? En hoe hangt de ontwikkeling van hun kennis samen met de inhoud van het programma? In hoofdstuk 6 gaan Bouwer en Koster in op de kennis van leerlingen. Het hoofdstuk doet verslag van een experimentele deelstudie met een *posttest only design* onder 593 leerlingen uit groep 6, 7 en 8. Alle leerlingen schreven een brief aan een denkbeeldige klasgenoot waarin zij advies gaven over hoe je een goed cijfer voor schrijven kunt krijgen. De leerlingen in de experimentele conditie kregen drie Teksterlessen, terwijl de leerlingen in de controlegroep de adviesbrief schreven zonder voorbereidende instructie. Een analyse van de gegeven adviezen laat zien dat de lessen effect hebben op de adviezen die leerlingen geven. De Teksterleerlingen gaven meer adviezen en meer adviezen met betrekking tot hogere orde-aspecten van schrijven. Uit analyses blijkt verder dat de aard van de adviezen samenhangt met de kwaliteit van de teksten. De kwaliteit van de adviesbrieven is hoger naarmate de leerlingen meer adviezen geven over het gebruik van interpunctie en hoofdletters, spelling en grammatica, presentatie, organisatie van de tekst en het schrijfproces. De frequentie van adviezen over inhoud en stijl hangt verrassend genoeg niet samen met tekstkwaliteit, maar zulke adviezen worden ook maar heel spaarzaam gegeven. Aan het slot van het hoofdstuk formuleren de onderzoekers terecht voorzichtige conclusies. De drie lessen hebben effect op de kwaliteit van de teksten van de leerlingen en op hun indirect gemeten kennis over schrijven en teksten. Die voorzichtigheid laten ze los in de Nederlandstalige samenvatting die aan het

slot van hun dissertatie is opgenomen: “door middel van het aanleren van een schrijfstrategie krijgen leerlingen meer kennis over schrijven, wat er vervolgens toe leidt dat ze betere teksten gaan schrijven” (Bouwer & Koster 2016, p. 300). Een dergelijke causale interpretatie vindt geen steun in de data. De onderzoekers laten immers niet zien dat de leerlingen hun nieuw verworven kennis gebruiken bij het schrijven. Tegen de causale interpretatie pleit ook het feit dat tekstkwaliteit het sterkst samenhangt met kennis die niet in het programma wordt aangeboden (interpunctie en hoofdletters, spelling en grammatica, presentatie). Verder hebben de analyses van de samenhangen tussen adviezen en tekstkwaliteit betrekking op al het onderzoeksmateriaal, ongeacht de herkomst ervan uit de experimentele dan wel de controleconditie.

Hoofdstuk 6 en ook de rest van het proefschrift werpt geen licht op de vraag in hoeverre het aanbieden van structuurkennis heeft bijgedragen aan de effecten van Tekster. De analyse van de adviesbrieven is daarvoor niet specifiek genoeg. Hetzelfde geldt voor de beoordeling van de leerlingteksten, in dit deelonderzoek en in de beide interventiestudies. Bij het beoordelen van de vele duizenden leerlingteksten is namelijk gebruik gemaakt van schaalbeoordeling. Deze vorm van beoordeling staat centraal in de twee laatste deelstudies die in deze bijdrage besproken worden.

De basis van Tekster: begrip en meting van schrijfvaardigheid

In hoofdstuk 4 worden de betrouwbaarheid en validiteit van verschillende beoordelingsprocedures met elkaar vergeleken: holistische, analytische en schaalbeoordeling. 36 studenten beoordeelden persuasieve teksten van 34 leerlingen uit groep 8, afkomstig uit de studie van Pullens (2012). Per leerling werden twee teksten beoordeeld. De resultaten

laten zien dat de verschillen tussen de beoordelingsmethoden beperkt zijn. Bouwer en Koster geven de voorkeur aan schaalbeoordeling en wijden aan deze aanpak nog een vervolgstudie waarin 10 studenten driemaal 40 teksten beoordelen: twee teksten uit hetzelfde genre (een persuasief briefje) en een uit een afwijkend genre (een adviesbrief). De studie laat zien dat een en dezelfde beoordelingschaal gebruikt kan worden voor teksten van een zelfde genre, maar met een (iets) ander onderwerp. Deze uitkomst sterkt de onderzoekers in hun voorkeur voor het werken met schaalbeoordeling.

In hoofdstuk 3 doen de onderzoekers verslag van een onderzoek naar de vraag naar het aantal schrijftaken en beoordelaars dat nodig is om een betrouwbaar oordeel over de individuele schrijfvaardigheid van leerlingen te kunnen vellen. In het hoofdstuk, dat ook gepubliceerd is als tijdschriftartikel (Bouwer, Béguin, Sanders & Van den Bergh, 2015; Bouwer & Van den Bergh, 2015), wordt een analyse gepresenteerd van oordelen over een steekproef van teksten van 67 leerlingen uit groep 8, opnieuw afkomstig uit de studie van Pullens (2012). De steekproef bevat 12 teksten per leerling: 3 teksten in 4 verschillende genres (een persuasief briefje, een betogende tekst die dient als voorbereiding op een discussie, een avonturenverhaal voor de lezers van de schoolkrant en een persoonlijk verhaal). Iedere tekst is beoordeeld door een jury van drie beoordelaars die bij het beoordelen gebruik maakten van een ankertekst van gemiddelde kwaliteit. Analyses van de generaliseerbaarheid van de tekstbeoordelingen laten zien dat de variantie in de scores voor tekstkwaliteit aan verschillende zelfstandige of interacterende bronnen kan worden toegeschreven, bronnen die inhoudelijk interessant en potentieel valide zijn: voor 10% aan de leerling (ongeacht het genre en het onderwerp daarbinnen), voor 11% aan het genre (ongeacht de leerling en het onderwerp), voor 4% aan de interactie tussen

leerling en genre, voor 2% aan de interactie tussen genre en onderwerp en voor 19% aan de interactie tussen leerling, onderwerp en genre. Twee andere bronnen van variantie baren zorg omdat ze niet alleen met de leerling, het genre of het onderwerp verbonden zijn, maar ook met de beoordelaar: 18% van de variantie zit vast aan de interactie tussen beoordelaar, taak en genre en maar liefst 36% aan de combinatie van foutenvariantie en de interactie tussen beoordelaar, leerling, genre en onderwerp.

Vanuit de aanname dat schrijfvaardigheid een individuele eigenschap is die ongeacht genre en onderwerp gemeten kan en moet worden, zijn deze resultaten verontrustend. Dan luidt immers de conclusie, zoals die van de onderzoekers, dat slechts 10% van variantie in de scores met individuele schrijfvaardigheid is geassocieerd. Erg realistisch is deze aanname echter niet. De data uit deze en tal van andere studies wijzen op een andere aanname, en wel een die beter correspondeert met wat we weten over de aard, complexiteit en gesitueerdheid van schrijven. Schrijvers moeten werken met de kennis uit hun langetermijngeheugen en met informatie die in de schrijffomgeving beschikbaar is of beschikbaar gemaakt moet worden. Zij moeten verder leren schrijven conform bepaalde genres en ook leren hoe ze daar doelbewust van kunnen afwijken. Inhoud en genre tellen met andere woorden mee. Dat is een ongemakkelijke waarheid voor wie wil toetsen, maar goed schrijfonderwijs hoeft daar geen last van te hebben. Zoveel laten de studies van Bouwer en Koster ook zien.

* Achtergronden van Tekster, voorbeeldmateriaal en gebruikerservaringen zijn te vinden op <https://tekster.nl>.

LITERATUUR

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Erlbaum.
Bouwer, R., Béguin, A., Sanders, T., & Bergh,

H. van den (2015). Effect of genre on the generalizability of writing scores. *Language Testing*, 32, 83–100.

Bouwer, R., & Bergh, H. van den (2015). Toetsen van schrijfvaardigheid: hoeveel beoordelaars, hoeveel taken? *LevendeTalen Tijdschrift*, 16(3), 3–13.

Bouwer, R., & Koster, M. (2016). *Bringing writing research to the classroom. The effectiveness of Tekster, a newly developed writing program for elementary students*. Academisch proefschrift. Universiteit Utrecht.

Broekhuis, S., & Glopper, K. de (2016). Plaatsen in de tijd. De invloed van instructie in specifieke genrekenmerken op de planningsactiviteiten tijdens het schrijven van een verhaal. *LevendeTalen Tijdschrift*, 17(4), 3–14.

Hoogeveen, M. (2012). *Writing with peer response using genre knowledge. A classroom intervention study*. Academisch proefschrift. Universiteit Twente.

Koster, M., Tribushinina, E., De Jong, P. F., & Van den Bergh, H. (2015). Teaching children to write: A meta-analysis of writing intervention research. *Journal of Writing Research*, 7(2), 249–274.

Norden, S. van (2014). *Iedereen kan leren schrijven. Schrijfplezier en schrijfvaardigheid in het basisonderwijs*. Bussum: Coutinho.

Pullens, T. (2012). *Bij wijze van schrijven: Effecten van computerondersteund schrijven in het primair onderwijs*. Academisch proefschrift. Universiteit Utrecht.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin Company.

KEES DE GLOPPER (1956) is hoogleraar Taalbeheersing van het Nederlands aan de Rijksuniversiteit Groningen. Zijn onderzoek en onderwijs hebben betrekking op de ontwikkeling van, het onderwijs in en het gebruik en de consequenties van geleterdheid.
E-mail: c.m.de.glopper@rug.nl