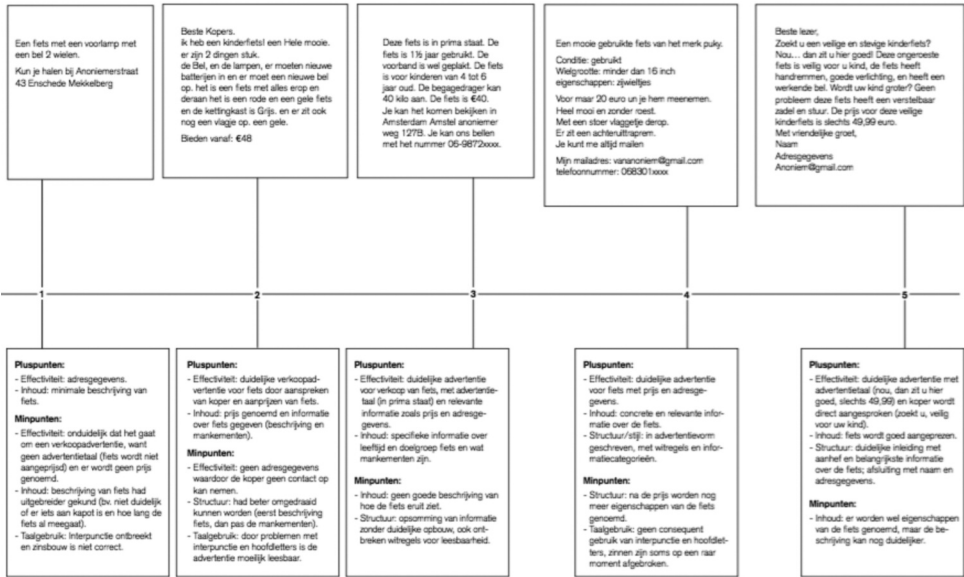


Het beoordelen van schrijfvaardigheid en het geven van feedback

In de huidige, digitale kenniseconomie verloopt het merendeel van onze communicatie schriftelijk, via e-mail, *social media*, blogs, websites en (digitale) formulieren. Goed kunnen schrijven is een essentiële vaardigheid om volwaardig deel te kunnen nemen aan de maatschappij. Het is daarom belangrijk dat leerlingen al op jonge leeftijd goed leren schrijven. In het *Referentiekader Taal en Rekenen* is vastgelegd dat een leerling aan het eind van de basisschool korte, eenvoudige teksten moet kunnen schrijven over alledaagse onderwerpen of over onderwerpen uit zijn leefwereld. Voorbeelden hiervan zijn: brieven, kaartjes, e-mails, werkstukken, samenvattingen, maar ook verhalen en gedichten (Expertgroep Doorlopende Leerlijnen Taal en Rekenen 2009). Maar hoe kun je bepalen of leerlingen deze doelstelling hebben behaald?

Of een tekst geslaagd is of niet, hangt van allerlei aspecten af, zoals de inhoud van de tekst, de structuur, de stijl en de woordkeuze. Maar ook aspecten zoals ‘spelling’, ‘interpunctie’ en ‘handschrift’ kunnen hierbij een rol spelen. Beoordelaars verschillen in welke aspecten ze belangrijk vinden bij het beoordelen van een tekst. Zo zal de ene beoordelaar meer waarde hechten aan creativiteit en fantasie, waar de ander zich meer zal richten op de correctheid van de tekst. Het is dan ook niet verwonderlijk dat uit onderzoek is gebleken dat beoordelaars enorm van elkaar kunnen verschillen in hun beoordeling als aan hen wordt gevraagd om een oordeel te geven over de tekst als geheel (Wesdorp 1981; Diederich, French & Carlton 1961). Dat onderzoek laat zien dat beoordelaars houvast nodig hebben om tot een betrouwbaar oordeel te kunnen komen. Een oplossing hiervoor is om analytische schema's te gebruiken waarin elk deelaspect gescoord moet worden. Dit soort beoordelingsschema's verhoogt de betrouwbaarheid van de oordelen, maar er kleven ook belangrijke nadelen aan vast. Zo is het lastig om op basis van analytische scores iets te zeggen over de kwaliteit van de tekst als geheel. Een ander groot nadeel van analytisch beoordelen is dat het tijdrovend is, omdat iedere tekst meerdere malen gelezen moet worden. De uitdaging is dus om een manier van beoordelen te vinden die niet al te tijdrovend is, recht doet aan de tekst als geheel, maar toch betrouwbaar is.

Om de bovengenoemde bezwaren te kunnen ondervangen, hebben wij in het onderzoek naar de effectiviteit van *Tekster*, een lesmethode voor schrijfvaardigheid voor groep 6 t/m 8, gebruikgemaakt van een beoordelingsprocedure op basis van beoordelingsschalen met ankerteksten (Bouwer & Koster 2016; Koster, Bouwer & Van den Bergh 2014a; 2014b; 2014c). Deze schalen bestaan uit vijf teksten, oplopend in tekstkwaliteit, die de spreiding van de verschillen in tekstkwaliteit van groep 6 tot en met groep 8 weerspiegelen. Figuur 1 laat zien hoe dat eruitziet. Het grote voordeel van zo'n beoordelingsschaal is dat het beoordelaars ondersteuning biedt bij het beoordelen van de tekst als geheel.



Figuur 1 – Voorbeeld van een beoordelingsschaal.

De middelste tekst op de schaal is ongeveer van gemiddelde kwaliteit (hij voldoet aan de opdracht, de boodschap komt over, maar ook niet meer dan dat). De tekst rechts hiervan is beter en de tekst uiterst rechts is zeer goed. De tekst links van het midden is minder goed dan de gemiddelde tekst en de tekst uiterst links is zwak. Onder elke tekst staat een korte toelichting waarom hij geplaatst is op deze plek op de schaal. Beoordelaars komen tot een oordeel over de kwaliteit van een tekst door deze te vergelijken met de teksten op de schaal. Deze procedure blijkt meer betrouwbaar dan holistisch beoordelen zonder ankerteksten. Bovendien gaat dit ook niet ten koste van de validiteit. De tekst wordt immers nog steeds als geheel beoordeeld.

Wat betreft de praktische toepassing van dit type beoordelingsschalen, is gebleken dat het niet noodzakelijk is om voor iedere schrijftaak een nieuwe schaal met ankerteksten te ontwikkelen. De schalen blijken namelijk ook bruikbaar voor het beoordelen van taken die anders zijn dan de teksten op de schaal, op voorwaarde dat de teksten van

hetzelfde genre zijn. De oordelen blijken bovendien even betrouwbaar als wanneer de leerlingteksten overeenkomen met de ankerteksten op de schaal (Bouwer & Koster 2016). De betrouwbaarheid van de oordelen neemt echter af bij het beoordelen van teksten uit een ander genre, wat aangeeft dat de ankerschalen hierbij minder ondersteuning bieden. Het is daarom raadzaam om voor verschillende genres aparte schalen te ontwikkelen.

Wat verder pleit voor het ontwikkelen van beoordelingsschalen in verschillende genres, is het feit dat de prestaties van individuele leerlingen enorm van elkaar kunnen verschillen, afhankelijk van het genre of van het onderwerp van de schrijftaak. Een leerling die veel fantasie heeft, kan veel beter zijn in het schrijven van verhalen dan in het schrijven van een informatieve tekst, bijvoorbeeld. Verder is het ook vaak makkelijker om een tekst te schrijven over een onderwerp waar je al veel van afweet dan over een onderwerp dat totaal nieuw voor je is. We doen leerlingen dan ook tekort als we een oordeel vellen over zijn of haar schrijfvaardigheid, gebaseerd op slechts één schrijftaak of één genre. Het is beter om een oordeel over de schrijfvaardigheid van een leerling te baseren op meerdere teksten, bij voorkeur van verschillende genres. Onderzoek heeft laten zien dat voor een echt betrouwbaar oordeel van de algemene schrijfvaardigheid leerlingen ten minste drie teksten in vier verschillende genres zouden moeten schrijven en dat deze teksten door ten minste twee beoordelaars zouden moeten worden beoordeeld (Bouwer & Van den Bergh 2015).

In de dagelijkse onderwijspraktijk is het niet haalbaar en niet wenselijk om iedere leerling 12 teksten te laten schrijven en deze ook nog eens door meerdere beoordelaars te laten beoordelen. Het onderzoek van Bouwer & Van den Bergh (2015) toont echter wel het belang aan van het gebruik van meerdere schrijftaken en van goede beoordelingsinstrumenten als je uitspraken wilt doen over de schrijfvaardigheid van een leerling. In het onderzoek naar *Tekster* hebben we leerlingen een verhaal, een beschrijving en een overtuigende tekst laten schrijven om een goede indruk te krijgen van hun schrijfvaardigheid. Deze genres sluiten aan bij de schrijfdoelen die worden genoemd in de kerndoelen Nederlands voor het primair onderwijs: ‘plezier verschaffen’, ‘informereren’ en ‘overtuigen’ (Grever & Letschert 2006). Voor deze drie genres hebben we beoordelingsschalen met ankerteksten ontwikkeld. Het feit dat de beoordelingsschalen in verschillende leerjaren (groep 6 t.e.m. groep 8) gebruikt kunnen worden, biedt de mogelijkheid om de schrijfontwikkeling van een leerling te volgen over de leerjaren heen.

Doordat de schaal met ankerteksten voorbeelden geeft van teksten van oplopende kwaliteit, geeft deze ook aanknopingspunten voor het geven van gerichte feedback om een leerling een stapje verder te helpen. Door de tekst van de leerling op de schaal te zetten en te vergelijken met de tekst van het volgende ankerpunt, wordt inzichtelijk op welke aspecten de leerling zijn tekst kan verbeteren, zodat hij één stapje verder op de schaal kan uitkomen. Effectieve feedback is afgestemd op de behoefte van de leer-

ling, geeft hem inzicht in zijn prestatie en aanwijzingen hoe hij verder kan komen. In de beperking toont zich de meester: kies de twee tot drie belangrijkste verbeterpunten en houd hierbij vooral de communicatieve effectiviteit van de tekst in het oog. De beoordelingsschaal kan vervolgens gebruikt worden als hulpmiddel bij de bespreking van een tekst met een leerling. De schalen bieden dus meerdere toepassingsmogelijkheden voor de onderwijspraktijk.

Referenties

- Bouwer, R. & M. Koster (2016). *Bringing writing research into the classroom. The effectiveness of Tekster, a newly developed writing program for elementary students*. [Doctoral thesis]. Utrecht: Universiteit Utrecht.
- Bouwer, R. & H. van den Bergh (2015). "Toetsen van schrijfvaardigheid: Hoeveel beoordelaars, hoeveel taken?". In: *Levende Talen Tijdschrift*, 16 (3), p. 3-12.
- Diederich, P., J. French & S. Carlton (1961). *Factors in judgment of writing ability*. Princeton, NJ: Educational Testing Service.
- Expertgroep Doorlopende Leerlijnen Taal en Rekenen (2009). *Referentiekader Taal en Rekenen: De referentieniveaus*. Den Haag: Ministerie van OCW.
- Grever, J. & J. Letschert (2006). 'Kerndoelen Primair Onderwijs'. Den Haag: Ministerie van OCW.
- Koster, M., R. Bouwer & H. van den Bergh (2014a). *VOS: werkboek en docentenhandleiding voor groep 6*. Utrecht: Universiteit Utrecht.
- Koster, M., R. Bouwer & H. van den Bergh (2014b). *DODO: werkboek en docentenhandleiding voor groep 7*. Utrecht: Universiteit Utrecht.
- Koster, M., R. Bouwer & H. van den Bergh (2014a). *EKSTER: werkboek en docentenhandleiding voor groep 8*. Utrecht: Universiteit Utrecht.
- Wesdorp, H. (1981). *Evaluatietechnieken voor het moedertaalonderwijs*. Den Haag: SVO.