

Toetsing in context vraagt doordenking van de rol van taal

Dr. Jacqueline Evers-Vermeul – universitair docent Nederlands Universiteit Utrecht

De concept-context-benadering wordt in de onderwijspraktijk vaak toegejuicht: leerlingen moeten kennis niet (alleen) in isolatie, maar ook in context aanleren en kunnen toepassen. Deze visie werkt door in de toetsing: contextafhankelijke toetsvragen zijn populair. Dit leidt echter ook tot klachten over de 'taligheid' van toetsen en examens. Dit spanningsveld vereist dat toetsmakers de rol van taal in toetsen beter gaan doordenken: hoe kun je de taligheid van vragen aanpassen aan het gewenste niveau, welke tekst- en beeldkenmerken maken een vraag makkelijker of juist moeilijker, en hoe bepalend mag iemands leesvaardigheid zijn? In deze bijdrage laat ik zien dat de constructie van toetsen idealiter een samenspel wordt tussen vakinhoudelijke experts, psychometristen en taalexperits.

Toetsen in context

Onze 21^{ste}-eeuwse samenleving is een informatiemaatschappij: burgers worden geacht zelf allerlei informatie op te kunnen zoeken en te selecteren (vooral via het internet), en de verzamelde informatie te kunnen vertalen naar hun eigen situatie. Het is dan ook noodzakelijk dat ons onderwijs goed voorbereidt op deelname aan deze maatschappij. Dit vereist dat kennis niet alleen in isolatie, maar ook in context wordt aangeleerd en toegepast. De concept-context-benadering heeft dus terecht een belangrijke plek in onze onderwijspraktijk.

Deze benadering werkt ook door in de toetsing. Zo schreef het toenmalige College voor Examens (2010, p. 16) in haar brochure *Examen en werkelijkheid – Van Cevo naar College voor Examens*:

"Het is een bewuste keuze om de werkelijkheid in alle examens aan de orde te laten komen. Kennis is immers zinloos als je die niet kunt toepassen. Daarom staan er echte krantenartikelen in het examen Duits en Engels en levensechte situaties bij wiskunde, natuurkunde en scheikunde. Een redelijke beheersing van het Nederlands is nodig om het diploma te halen – en niet alleen voor het examencijfer Nederlands. In het vervolgonderwijs, bij het werk en in de samenleving is die beheersing van het Nederlands immers óók nodig."

Contextafhankelijke items zijn dan ook niet meer weg te denken uit examens en methodegebonden toetsen.

Toetsen te talig?

Tegelijkertijd levert de toepassing van de concept-context-benadering in toetsen en examens allerlei klachten op. Met name bij Maatschappijleer wordt er geklaagd dat leerlingen teveel moeten lezen, en de rekentoetsen liggen regelmatig onder vuur vanwege de 'taligheid' en de hoeveelheid 'verhaaltjessommen'. Zo sprak de commissie-Bosker zich negatief uit over de kwaliteit van de rekentoetsen vo en examens mbo, "die te moeilijk en/of te gekunsteld en/of te talig zouden zijn" (Bosker & Van de Vorle 2014, p. 12).

Eerdere klachten over het taalgebruik in examens vormden al de aanleiding voor het College voor Toetsing en Examens om de Universiteit Utrecht hier onderzoek naar te laten doen (zie o.a. Evers-Vermeul & Land 2012a, b; Evers-Vermeul & Pardoel 2014). Opmerkelijk genoeg worden uitspraken over de grote 'taligheid' van examens of van bepaalde schoolvakken vaak namelijk nauwelijks onderbouwd met argumenten of resultaten uit onderzoek. Zo onderzochten Österholm en Bergqvist (2013) 50 wetenschappelijke artikelen met daarin 311 claims over het taalgebruik in wiskundeschoolboeken. Zij ontdekten dat 38% van deze claims totaal niet werd onderbouwd en dat het gros van de overige claims slechts gelardeerd werd met exemplarische voorbeelden (27%) of verwijzingen naar werk van collega-wetenschappers die eveneens slechts een claim poneerden (14%). Vergelijkende empirische studies bleken zeldzaam en spraken bovendien op meerdere punten tegen dat wiskunde relatief moeilijke taal zou bevatten. Hoe 'talig' toetsen en schoolboeken van verschillende vakken precies zijn, is dus grotendeels nog een kwestie die empirisch onderzoek vraagt.

Meer lezen een probleem?

Hoe dan ook doen contextafhankelijke opgaven een groter beroep op de leesvaardigheid van leerlingen dan contextloze opgaven. Meer leeswerk of lastiger taalgebruik bij de zaakvakken kan taalzwakke leerlingen op een achterstand zetten. Dit betekent dat er een *test bias* of construct-irrelevante variantie kan ontstaan (Abedi 2006; Haladyna & Downing 2004). Hoe werkt dit? Uit onderzoek blijkt dat vooral zwakkere lezers moeite hebben om veel informatie tegelijkertijd te verwerken. De *Cognitive Load Theory* (Sweller e.a. 1998) gaat ervan uit dat het werkgeheugen van zwakke lezers meer belast wordt met basale leesprocessen (zoals het decoderen van woorden), omdat deze leesprocessen minder geautomatiseerd verlopen dan bij goede lezers. Ook zijn zwakke lezers minder goed in staat om verbanden tussen zinnen te herleiden (Van Silfhout e.a. 2013). Lange stukken tekst leggen dus een extra druk op hun werkgeheugen, wat ten koste kan gaan van hun prestatie op de vraag.

Of dat een probleem is, hangt af van twee zaken. Ten eerste is er de afweging hoe zwaar leesvaardigheid moet meewegen bij andere vakken dan Nederlands. Hier moeten beleidsmakers per vak dus een uitspraak over doen, een uitspraak die gevolgen kan hebben voor de verhouding contextafhankelijke en contextloze opgaven in een toets, en voor de hoeveelheid leeswerk per item.

Ten tweede moet bepaald worden of de door contexten vereiste leesvaardigheid wel past bij het onderwijsniveau waarop getoetst wordt: wanneer is het taalgebruik te moeilijk voor een referentieniveau? Evers-Vermeul en Holtermann (2013) betogen dat dit geen eenvoudige opgave is en laten zien dat er nog weinig concrete richtlijnen voorhanden zijn om het taalgebruik af te stemmen op specifieke onderwijsniveaus. Samenwerking met en nader onderzoek door taalexperts is hier dan ook gewenst.

Minder leeswerk

Wat nu als toetsmakers de rol van leesvaardigheid willen verkleinen? Hieronder bespreek ik vier opties om de hoeveelheid leeswerk in toetsen te verminderen en geef ik aan welke beleidsmatige keuzes hierbij om de hoek komen kijken.

1. Verlaag het aantal contextafhankelijke opgaven. Hiervoor moeten toetsconstructeurs afwegen wat de verhouding contextafhankelijke en contextloze opgaven moet worden, en wanneer het toevoegen van een context gerechtvaardigd is. De concept-context-benadering impliceert niet dat we aan alle vragen een context moeten toevoegen. Onderzoek laat zien dat hier winst te behalen valt: examens bevatten regelmatig vragen – met name kennisvragen – die ook zonder context te beantwoorden zijn (Evers-Vermeul & Pardoel 2014).
2. Kort contexten in. Inkorten kan alleen als duidelijk is welke informatie noodzakelijk is voor het beantwoorden van de vraag. Regelmatig bevatten contexten ‘onnodige’ informatie, zodat ook de opzoekvaardigheden van leerlingen getoetst kunnen worden. Het is dan de vraag in hoeveel vragen toetsing van die opzoekvaardigheden centraal moet staan, en in hoeveel vragen de context zo kort mogelijk kan blijven. Bij het inkorten van contexten moeten toetsconstructeurs ook overwegen welke motieven zij hebben om echte bronnen als context te selecteren en of die motieven zó zwaar wegen dat ze het extra beroep op leesvaardigheid rechtvaardigen. Dit is vooral van belang bij een vak als Maatschappijleer, dat frequent gebruik maakt van (licht bewerkte) krantenberichten als context.
3. Werk met item-clusters in plaats van met losse items. Evers-Vermeul en Pardoel (2014) laten zien dat de frequentie van clusters in digitale vmbo-bb-examens nogal verschilt per vak. Wiskunde bevat fictieve, maar realistisch ogende contexten, wat het mogelijk maakt om vier of vijf vragen over eenzelfde context te stellen (zoals over Anil die zijn slaapkamer gaat opknappen), waarbij vaak gaandeweg meer details aan die context worden toegevoegd. Dat geeft een groot verschil in hoeveelheid leeswerk met Maatschappijleer-examens, waarin clusters van twee items al zeldzaam zijn doordat vrijwel elk item met een nieuw krantenbericht of videofragment begint.

Ook hier speelt een rol hoe brongetrouw contexten moeten zijn. Bij digitale examens is het bovendien raadzaam een manier te vinden die inzichtelijk maakt welke informatie reeds bekend en welke informatie nieuw is. Daarmee blijft reeds bekende informatie beschikbaar, maar wordt voorkomen dat leerlingen deze automatisch gaan herlezen als ze bij een vervolg-item komen.

4. Streef naar een optimale leesroute binnen items. In de huidige toetspraktijk worden contexten veelal zonder vooraankondiging gepresenteerd. Dit resulteert in onnodig herleeswerk, omdat leerlingen zonder zo'n vooruitblik niet weten met welk doel ze de bron lezen of welk type informatie ze hieruit zouden moeten filteren. Oogbewegingsonderzoek onder vmbo-leerlingen die vlak voor hun eindexamen zaten, laat zien dat examentrainingen om eerst de vraag te lezen geen effect sorteren: het gros van de leerlingen begon met het lezen van de context (Van den Bosch e.a. 2012). Het is dus raadzaam leerlingen wél vooraf duidelijk te maken met welke 'bril' ze de gegeven informatie moeten bekijken. Met name bij digitale toetsen vraagt dit om expertise over de manier waarop combinaties van tekst en beeld verwerkt worden, zoals bijvoorbeeld de *Cognitive Load Theory* (Sweller e.a. 1998) en de 'grammatica' van visueel ontwerp (Kress & Van Leeuwen 2006).

Conclusie

Onze informatiemaatschappij vereist dat we onze leerlingen zo goed mogelijk leren lezen en zoveel mogelijk kennis en vaardigheden aanleren die ze in context kunnen hanteren. Toetsing in context is hierbij een nuttig hulpmiddel om te achterhalen hoe de ontwikkeling van leerlingen ervoor staat. Hoe die toetsing in context precies vorm moet krijgen, vraagt een grondige doordinking van de rol van taal daarbinnen en een afweging van de verschillende doelen die met toetsen worden nagestreefd. Ongefundeerd klagen over de 'taligheid' van examens is daarbij niet constructief. Beleidsmakers kunnen zich beter bezighouden met vragen als 'hoe zwaar mogen leesvaardigheid en opzoekvaardigheden meewegen bij verschillende vakken?', 'welke mate van taalcomplexiteit past bij welk referentieniveau?', 'hoe belangrijk is het dat examens echte in plaats van fictieve bronnen bevatten?' en 'hoe optimaliseren we de leesroute binnen toetsvragen?' Om antwoord te vinden op deze vragen en daarmee de kwaliteit van onze toetsen te verhogen, is dan ook intensieve samenwerking nodig tussen mensen met kennis van de vakinhoud, van het ontwerp van toetsen en van taal.

Literatuur

- Abedi, J. (2006). Language issues in item development. In S. M. Downing & T. M. Haladyna (red.), *Handbook of test development* (pp. 377-398). Mahwah, NJ: Lawrence Erlbaum.
- Bosch, L.J. van den, Evers-Vermeul, J. & Phielix, C. (2012). Back on track: the role of structure in context-dependent exam questions. In *Student Research Conference 2012* (pp. 29-32). Utrecht: SRC/VSNU.
- Bosker, R. & Vorle, R. van de (red.) (2014). *Advies over de uitwerking van de referentieniveaus 2F en 3F voor rekenen in toetsen en examens*. Enschede: SLO.
- College voor Examens (26 januari 2010). *Examen en werkelijkheid – Van Cevo naar College voor Examens*. Zie <http://www.cve.nl/item/brochures>.
- Evers-Vermeul, J. & Holtermann, M. (2013). Doorlopende leerlijnen: implicaties voor *leveling* van leer- en examenteksten voor het middelbaar onderwijs. *Tijdschrift voor Taalbeheersing* 35(1), 1-24.
- Evers-Vermeul, J. & Land, J. (2012a). Slecht gelezen of slecht geleerd? – 1: Persoonlijk taalgebruik en woordkeuze in vmbo-examens. *Examens – tijdschrift voor de toetspraktijk* 9(1), 22-25.
- Evers-Vermeul, J. & Land, J. (2012b). Slecht gelezen of slecht geleerd? – 2: Sitatieschetsen en informatieopbouw in vmbo-examens. *Examens – tijdschrift voor de toetspraktijk*, 9(2), 25-29.
- Evers-Vermeul, J. & Pardoel, M. (2014). *Contexten in digitale examens: brug of barrière? Een onderzoek naar de leesbaarheid van vmbo-BB-examens Wiskunde en Maatschappijleer 2*. Ongepubliceerd onderzoeksverslag in opdracht van College voor Examens.

- Haladyna, T.M. & Downing, S.M. (2004). Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice* 23(1), 17-27.
- Kress, G. & Leeuwen, T. van (2006). *Reading images: The grammar of visual design*. 2^e ed. London & New York: Routledge.
- Österholm, M. & Bergqvist, E. (2013). What is so special about mathematical texts? Analyses of common claims in research literature and of properties of textbooks. *ZDM - The International Journal on Mathematics Education* 45(5), 751-763.
- Silfhout, G. van, Evers-Vermeul, J. & Sanders, T.J.M. (2013). *Omdat leerlingen moeite hebben met die teksten. Daarom! Tekstbegrip verbeteren in het voortgezet onderwijs*. In D. Schram (red.), *De aarzelende lezer over de streep: Recente wetenschappelijke inzichten* (pp. 85-104). Delft: Eburon Academic Publishers.
- Sweller, J., Merriënboer, J. van & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review* 10, 251-296.

Auteur

Dr. Jacqueline Evers-Vermeul werkt als docent-onderzoeker bij het Departement Talen, Literatuur en Communicatie aan de Universiteit Utrecht. Zij doet onderzoek naar taal- en geletterdheidontwikkeling en de rol die teksten daarin spelen: welke kenmerken maken een tekst lastiger te begrijpen, hoe komt dat en wat betekent dit voor de onderwijspraktijk? E-mail: j.evers@uu.nl.