

Beoordeling van de samenvattingsopdracht in het nieuwe examen Nederlands

Eind 1995 verschenen de voorstellen van de Vakontwikkelgroep Nederlands (VOG-N) voor het Examen Nederlands in havo en vwo. In die voorstellen is een belangrijke plaats ingeruimd voor een samenvattingsopdracht als onderdeel van de (centrale) toetsing van leesvaardigheid.

Aan het SCO-Kohnstamm Instituut van de Universiteit van Amsterdam is een eerste onderzoek verricht naar de nieuwe samenvattingsopdracht en mogelijke beoordelingsprocedures. Daarnaast is aan docenten gevraagd welke beoordelingsprocedure hun voorkeur heeft bij de beoordeling van samenvattingen in het ce.

De VOG-N kent de samenvattingsopdracht een belangrijke plaats toe in de centrale toetsing van 'leesvaardigheid', waarbij de algemene doelstelling voor leesvaardigheid is dat 'de leerling [...] verschillende soorten niet-fictionele teksten [kan] analyseren, interpreteren, beoordelen en samenvatten' (Stuurgroep Profiel tweede fase 1995, 10). Voor het 'subdomein' samenvatten geldt als eindterm dat 'de leerling [...] in staat [is] teksten en tekstgedeelten beknopt samen te vatten, d.w.z. te reduceren tot de hoofduitspraak (hoofduitspraken) met bijbehorende ondersteuning of (belangrijke) ondergeschikte uitspraken' (o.c.: 11). Met deze voorstellen krijgt de samenvattingsopdracht niet alleen een centrale plaats in de toetsing van de leesvaardigheid in zowel het vwo als het havo, maar verandert er ook het een en ander aan de vorm van de opdracht. Met die verandering wordt tegemoet gekomen aan kritiek die er al enige tijd bestaat op de huidige samenvattingsopdracht.

De huidige opdracht

Dat het kunnen samenvatten van een gelezen tekst een belangrijke vaardigheid is voor het voortgezet en hoger onderwijs en soms de beroepsuitoefening, is in divers onderzoek aangetoond en staat over het algemeen buiten kijf (zie De Glopper en Van Schooten 1990). Het samenvatten is tot nu toe het herkenbaarst in het vwo waar het deel uitmaakt van het centraal examen (ce). In de samenvattingsopdracht in het ce van het vwo wordt van de examenkandidaat verlangd dat hij de inhoud van een gemengd infor-

matiefbetogende tekst van ongeveer 2300 woorden terugbrengt tot een eigen tekst van maximaal 500 woorden. De samenvatting moet dusdanig zijn dat 'iemand die de uitgangstekst niet kent, de samenvatting goed kan volgen'.

De docent scoort vervolgens de inhoud van de samenvatting (sinds 1988 verplicht) met behulp van een analytisch schema, waarin de noodzakelijke informatie-elementen zijn opgenomen met bijbehorende normering. Voor overschrijding van de maximum-lengte en gemaakte taalfouten (zinsbouw, spelling en interpunctie) bestaan aparte regelingen.

De bezwaren die van verschillende kanten tegen deze toetsvorm geuit zijn, hebben deels betrekking op de opdracht en deels op de beoordelingsprocedure.

De gewenste opdracht

Een belangrijk probleem bij de huidige samenvattingsopdracht is de definiëring van de opdracht: 'Welk type samenvatting moet worden geschreven?' De gebruikte opdrachtvorm lijkt niet inhoudsvalide. De leerlingen wordt gevraagd een 'representatieve samenvatting' te schrijven, terwijl met de opdracht veeleer een 'informatieve samenvatting' beoogd zou moeten worden (Van Eemeren e.a. 1975, Hendrix en Sanders 1985, Braet 1988). Braet (1988) stelt dat bij de samenvattingsopdracht feitelijk geen (korte) samenvatting wordt gevraagd, maar een (lang) uittreksel. Drop en De Vries (1977, 143) noemen een samenvatting die één derde van de oorspronkelijke tekst bevat zelfs 'een vrij dwaze opdracht'.

Verskillende opvattingen over de doelstelling van het samenvattingsonderwijs hebben ook aanleiding gegeven tot discussies over de opdracht en de beoordeling van de samenvatting (Hendrix 1985, Van Aken-van de Leur 1988, Doesborgh 1989, Sectiebestuur Nederlands 1989). De leraren op de bijeenkomst georganiseerd door het Sectiebestuur Nederlands in 1989, waar deze doelstelling centraal stond, gaven unaniem te kennen dat de huidige samenvattingsexamenpraktijk moet veranderen. Een grote meerderheid sprak zich uit voor een kortere samenvatting volgens de ideeën van Braet (1988, Sectiebestuur Nederlands 1989). Ook de respondenten uit het behoefteonderzoek van De Glopper en Van Schooten (1990) geven de voorkeur aan een korte samenvatting ('volledig kort samenvatten'). De samenvattingsopdracht in het ce is in 1990 als compromis slechts op

kleine punten gewijzigd als een overgangsregeling naar een nieuw eindexamen, dat nu in de nieuwe voorstellen vormgegeven is.

De beoordelingsbetrouwbaarheid

Los van de opdrachtvorm is de betrouwbaarheid van de beoordeling steeds een punt van zorg geweest. In een vooronderzoek ter voorbereiding van de invoering van de analytische beoordeling schatten Sanders, Hendrix en Luijten op basis van de beoordelingen van zes docenten dat de gemiddelde individuele beoordelaarsbetrouwbaarheid .50 bedraagt (1984), hetgeen onvoldoende is. In een groter vervolgonderzoek vinden ze een hogere betrouwbaarheid, namelijk .84, maar helaas lijkt de betrouwbaarheidsschatting uit deze studie geflatteerd door een evident gebrek aan onafhankelijkheid van de (tweede) beoordelingen (Hendrix en Sanders 1987). De voor de cijferbepaling zeer belangrijke inhoudsscores van de eerste beoordelaar lijken vaak overgenomen door de tweede beoordelaar en de werkelijke betrouwbaarheid ligt zeer waarschijnlijk lager dan de gerapporteerde .84. Geconcludeerd moet worden dat niet duidelijk is wat de betrouwbaarheid is van de beoordeling van de samenvatting, maar echt geruststellend zijn de genoemde uitkomsten niet.

De taakvariabiliteit

Tenslotte willen we nog de aandacht vestigen op een aspect van de samenvattingsopdracht waarover we nagenoeg niets weten, namelijk de zogenoemde taakvariabiliteit. Hiermee bedoelen we dat niet duidelijk is wat de prestaties op de ene samenvattingsopdracht zeggen over prestaties op een volgende. Met andere woorden: in hoeverre is het cijfer van een leerling afhankelijk van die ene toevallige samenvattingsopdracht? In schrijfvaardigheidsonderzoek is gebleken dat de prestaties van leerlingen zeer sterk verschillen van schrijfopdracht tot schrijfopdracht (Van den Bergh, De Glopper en Schoonen 1988). Grote taakvariabiliteit impliceert dat men een groot aantal taken aan een leerling moet afnemen om een volledig beeld van zijn 'samenvatvaardigheid' te krijgen. Voor de huidige (en nieuw voorgestelde) samenvattingsopdracht zijn er, voor zover ons bekend, geen gegevens gerapporteerd over de taakvariabiliteit. In de adviezen geformuleerd door de Commissie Vernieuwing Eindexamenprogramma's Nederlandse taal- en letterkunde vwo en havo (CVEN) wordt terecht gewezen op dit probleem. Een probleem dat bij de nieuw voorgestelde (kortere) samenvattingsopdracht nog weleens groter zou kunnen zijn dan bij de huidige (CVEN 1991).

Een nieuwe samenvattingsopdracht, een nieuwe beoordelingsmethode?

Voorstellen voor een fundamentele verandering van het ce Nederlands worden al gedaan in het rapport van de CVEN (1991). Voorgesteld wordt dat vwo- én havo-abituriënten een qua taken vergelijkbaar eindexamen 'leesvaardigheid' doen; er zal alleen sprake zijn van een verschil in moeilijkheidsgraad, een niveaoverschil. Dit leesvaardigheidsexamen behelst

het volgende. Zowel de havo- als de vwo-abituriënt krijgt minimaal twee teksten voorgelegd. Bij één of enkele hiervan moeten vragen worden beantwoord en één tekst van 1000 tot 1500 woorden moet worden gecomprimeerd tot een 'beknopte' samenvatting. Het samenvattingsonderdeel zou dan één derde tot de helft van het leesvaardigheidscijfer bepalen (CVEN 1991, 118). Overigens bestaat over de weging van de verschillende onderdelen binnen het examen Nederlands nog onzekerheid (Stuurgroep Profiel tweede fase 1995).

Aan de samenvatting worden in het CVEN-rapport de volgende eisen gesteld:

- De samenvatting moet beperkt blijven tot de hoofduitspraak met bijbehorende ondersteuning of (belangrijke) ondergeschikte uitspraken, die inhoudelijk correct in hun onderlinge samenhang moeten worden gegeven;
 - De samenvatting moet een goed geformuleerde tekst zijn die los van de uitgangstekst te begrijpen valt, waarbij de ordening mag afwijken van de oorspronkelijke tekst en woorden en zinswendingen mogen worden overgenomen;
 - De stijl dient direct te zijn.
- (CVEN 1991, 119)

De opdracht is met deze criteria afgestemd op de 'informatieve samenvatting' van een betogende tekst zoals omschreven door Van Eemeren e.a. (1975; zie ook Braet 1988).

Over de beoordeling van de samenvattingen merkt de CVEN op:

Voor de (vwo- en havo-)samenvatting zal nog nader onderzoek verricht moeten [worden] naar de beste beoordelingsmethode. De gedachten van de commissie gaan uit naar schaalbeoordeling waarbij de gemaakte samenvattingen vergeleken worden met een schaal die bestaat uit voorbeelden van slechte, matige en goede samenvattingen.

(CVEN 1991, 171)

De aanbeveling voor een samenvattingsopdracht nieuwe stijl in zowel havo als vwo is door de Vakontwikkelgroep Nederlands overgenomen (Stuurgroep Profiel tweede fase 1995). Over de te prefereren beoordelingsmethode rept de VOG-N niet.

De nieuwe samenvattingsopdracht wordt dus zeer waarschijnlijk de toekomstige examenpraktijk. Blijft echter de vraag van de te verkiezen beoordelingsmethode en de betrouwbaarheid ervan, want bij de korte samenvattingsopdracht zou de beoordelingsbetrouwbaarheid weleens kunnen afnemen. Bovendien rest ook nog de vraag van de taakvariabiliteit van de nieuwe samenvattingsopdracht. Bij kortere uitgangsteksten en kortere samenvattingen zouden toevallige factoren weleens een grotere (ongewenste) rol kunnen spelen.

Een eerste onderzoek

Tegen deze achtergrond van aanstaande veranderingen en nog te maken keuzes is een eerste onderzoek uitgevoerd aan het SCO-Kohnstamm Instituut van de Universiteit van Amsterdam in opdracht van het ministerie van OCenW. Met het onderzoek beoogden we niet alleen meer te weten te komen over de te verwachten betrouwbaarheid van een drietal mogelijke beoordelingsmethoden en over de te verwachten taakvariabiliteit van de nieuwe opdrachtvorm, maar

ook wilden we inzicht verwerven in de bruikbaarheid en de te verwachten acceptatie 'in het veld' van een voor te stellen beoordelingsmethode. De drie beoordelingsmethoden die onderzocht zijn, zijn de meest voor de hand liggende, namelijk: globale, analytische en schaalbeoordeling. De eerste twee methoden zullen bekend zijn, de derde misschien niet. Bij schaalbeoordeling gebruikt de beoordelaar een reeks samenvattingen als ankers voor de cijfergeving. De (meestal) vijf samenvattingen hebben al een cijfer gekregen en zij markeren het bereik van de te verwachten scores. De reeks bestaat uit een zeer zwakke samenvatting, een zwakke samenvatting, een gemiddelde samenvatting, een goede en een zeer goede samenvatting. Door de te beoordelen samenvattingen te vergelijken met de reeks van vijf voorbeelden die als een soort liniaal fungeert, komt de beoordelaar tot zijn cijfers (zie verder voor deze aan schrijfvaardigheidsonderzoek ontleende beoordelingsmethode Van den Bergh en Rijlaarsdam 1986). Zo'n eerste onderzoek naar een opdrachtvorm die nog geen ingang gevonden heeft, kent uiteraard een aantal beperkingen. Die beperkingen vloeien voort uit de paradoxale situatie dat men iets wil onderzoeken dat nog niet ingevoerd is en tegelijkertijd wil men iets pas invoeren als er onderzoek naar gedaan is. In ons geval moesten we leerlingen vragen om samenvattingsopdrachten te maken volgens richtlijnen waarmee zij (waarschijnlijk) nog geen ervaring hebben en waarmee ze niet hebben kunnen oefenen. Bovendien moesten we docenten vragen samenvattingen te beoordelen die voor hen nieuw zijn en daarbij moesten zij (een) beoordelingsmethode(n) gebruiken waarmee ze niet vertrouwd zijn. Deze beperkingen ten spijt hebben we toch getracht een zo goed mogelijk inzicht te krijgen van wat er verwacht kan worden bij de nieuwe samenvattingsopdracht en een eventueel nieuwe beoordelingsprocedure.

De beoordelingen

Ruim 170 leerlingen uit vwo-5 maakten twee samenvattingsopdrachten (tekst A en B) waarbij ze een tekst van ca. 1400 woorden tot ongeveer 10% moesten reduceren. De leerlingen waren zoals gezegd enigszins gehandicapt omdat de opdracht nieuw voor hen was. Bovendien kregen de leerlingen relatief weinig tijd voor hun samenvatting; om praktische redenen hebben we de beperking van één lesuur aangehouden. Voor sommige leerlingen was dit te krap en zij hebben kladversies ingeleverd en andere leerlingen zullen vanwege tijdsdruk misschien niet optimaal gepresteerd hebben. Dit betekent dat het niveau van de samenvattingen lager zal liggen dan men bij een eindexamen mag verwachten. Overigens ging het er ons niet om uitspraken te doen over de vaardigheid van de leerlingen, maar het ging om de betrouwbaarheid waarmee verschillende beoordelingsmethoden toegepast kunnen worden en om de taakvariabiliteit. Twaalf docenten Nederlands beoordeelden (elkaar overlappende) subsets van de samenvattingen. De onderzoeksopzet was er op gericht dat van elke docent de beoordelingsbetrouwbaarheid bepaald kon worden voor elk van de drie beoordelingsmethoden.

Voor de precieze en technische details van de onderzoeksopzet verwijzen we naar het onderzoeksrapport (Schoonen 1996).

Bij de toepassing van de beoordelingsmethoden is zoveel mogelijk rekening gehouden met de procedures zoals die door het Cito bij het ce gebruikt worden, zij het dat de aftrekregels voor lengte-overschrijding en taalgebruik niet toegepast zijn. Deze aftrekregels hebben voor alle drie de beoordelingsmethoden dezelfde uitwerking en staan hier niet ter discussie. In het vervolg staan we eerst kort stil bij de uitkomsten van het beoordelingsonderzoek: de beoordelingsbetrouwbaarheid en de taakvariabiliteit. Vervolgens gaan we in op de appreciatie van de beoordelingsmethoden door de docenten.

De resultaten: beoordelingsbetrouwbaarheid en de taakvariabiliteit

Over de vraag naar de beoordelingsbetrouwbaarheid die met de drie methoden bereikt wordt, zullen we hier kort zijn. We richten ons op de vraag: 'Rangordenen de beoordelaars de leerlingen op dezelfde wijze?'

• Rangordening

We drukken de overeenstemming in de rangordeningen van beoordelaars uit in een correlatiecoëfficiënt tussen 0 en 1. Voor elke beoordelaar kon bepaald worden wat zijn of haar correlatie met een willekeurig andere beoordelaar is. Deze individuele beoordelaarsbetrouwbaarheid hebben we van de twaalf beoordelaars gemiddeld. In tabel 1 wordt voor elk van de drie methoden en de twee onderwerpen de gemiddelde beoordelaarsbetrouwbaarheid gerapporteerd.

	Tekst A	Tekst B
Schaalbeoordeling	.355	.445
Analytische beoordeling	.532	.506
Globale beoordeling	.611	.430

tabel 1: Gemiddelde beoordelaarsbetrouwbarheden voor elk van de drie beoordelingsmethoden en de twee teksten.

Op basis van de betrouwbaarheden in tabel 1 moet men toch concluderen dat de schaalbeoordeling geen gewenste beoordelingsmethode is, m.n. voor tekst A is de betrouwbaarheid laag (.355). Gemiddeld genomen ontlopen de analytische en globale beoordeling ontlopen elkaar niet zoveel. De globale beoordeling lijkt gevoeliger voor de verschillende teksten (.611 versus .430) dan de analytische beoordeling die wat dat betreft een stabielere indruk maakt (.532 versus .506).

Op basis van deze gegevens zou men een (lichte) voorkeur voor de analytische beoordeling kunnen uitspreken. De globale beoordeling lijkt een 'goede' tweede.¹ Verderop zullen we nagaan in hoeverre deze voorkeur gedragen wordt door de docenten die straks de beoordeling in praktijk moeten brengen. Eerst komen terug op het mogelijke probleem van de taakvariabiliteit.

• Taakvariabiliteit

Als we voorlopig uitgaan van de analytische beoordelingen dan kunnen we nagaan in hoeverre de prestaties van de leerlingen op de twee samenvattingsop-

drachten met elkaar samenhangen. In hoeverre lijkt iemands 'samenvatvaardigheid' afhankelijk van de toevallige samenvattingsopdracht of brontekst. Uit onze gegevens blijkt dat (gecorrigeerd voor onbetrouwbaarheid van de beoordelingen) de prestaties op de twee samenvattingen (tekst A en B) betrekkelijk weinig samenhangen (zgn. 'ware' correlatie van .352). Ook als we van globale beoordeling uit zouden gaan, komen we tot dezelfde conclusie. Met andere woorden: of je goed of slecht scoort, lijkt voor een groot deel afhankelijk van de tekst die je moet samenvatten en dat is natuurlijk een uiterst ongewenste situatie; een probleem dat losstaat van de beoordelingsproblematiek.

Docenten over de beoordelingsmethoden

Op basis van het beoordelingsonderzoek zou men een (lichte) voorkeur voor de analytische beoordeling kunnen uitspreken. De verschillen met de globale beoordeling zijn echter zeer klein en dan kan men zich afvragen of de voordelen die een analytische beoordeling geeft, opwegen tegen de extra inspanningen die ze van de beoordelaars vraagt. Zo kost een analytische beoordeling naar opgave de twaalf docenten ruim anderhalf keer zoveel tijd als een globale beoordeling. Bovendien kan men zich voorstellen dat er andere verschillen, voor- of nadelen, in het gebruikersgemak van de beoordelingsmethoden zijn die bij de keuze een rol moeten spelen. Om hier meer inzicht in te krijgen zijn leerkrachten benaderd en gevraagd naar hun oordeel.

Er is een enquête uitgevoerd onder ruim 100 havo- en vwo-docenten van wie er 59 reageerden met een ingevuld formulier. Zij hebben op basis van voorbeeldsets van de beoordelingsvoorschriften de drie methoden een schoolcijfer (0-10) gegeven voor bruikbaarheid in het ce. De geënquêteerde docenten wisten overigens niet van de hier gerapporteerde betrouwbaarheden, wél gaven we hun de geschatte tijd die voor een beoordeling volgens de betreffende methode nodig is. Behalve drie schoolcijfers konden leerkrachten commentaar geven bij de drie beoordelingsmethoden of hun cijfers. Om de interpretatie van de schoolcijfers enigszins vergelijkbaar te houden moest het cijfer 5 als twijfelachtig opgevat worden. Tabel 2 geeft de waardering van de docenten voor de drie methoden.

	gem.	s
Schaalbeoordeling	4.06	1.90
Analytische beoordeling	7.44	1.76
Globale beoordeling	5.04	1.95

tabel 2: Waardering voor de drie beoordelingsmethoden: gemiddelde (gem.) en spreiding (s) van de 'schoolcijfers'.

In tabel 2 zien we dat er maar één voldoende valt en dat is voor de analytische beoordelingsprocedure (7.44). Globale beoordeling vindt men van twijfelachtige kwaliteit (5.04) en de schaalbeoordeling krijgt een duidelijke onvoldoende (4.06). Overigens illustreren de indices voor de spreiding in de schoolcijfers (standaarddeviatie, s) dat men niet unaniem is: docenten verschillen nogal in hun appreciatie van de beoordelingsmethoden. Maar ondanks deze onderlin-

ge verschillen is er grosso modo een duidelijke voorkeur voor de analytische beoordeling. De oordelen over de methoden worden niet of nauwelijks beïnvloed door ervaring van de docenten; het beeld is in dat opzicht redelijk consistent. Het enige opmerkelijke is dat de nieuw voorgestelde schaalbeoordeling bij minder ervaren docenten (<10 jaar) iets meer genade vindt dan bij ervarener docenten, maar ook bij de minder ervaren docenten geniet de analytische beoordeling duidelijk de voorkeur.

Interessant is ook de toelichting die de respondenten soms bij de schoolcijfers gaven. Hoewel het om betrekkelijk weinig commentaar gaat van slechts 18 docenten – dat bovendien sterk varieerde van adhesie-betuigingen voor de analytische beoordeling en de samenvattingsopdracht-nieuwe-stijl tot een kritisch exposé over de samenvattingsopdracht in het ce – is er één opmerkelijk punt dat regelmatig terugkeerde. Een aantal opmerkingen concentreerde zich vooral op de transparantie van het analytische oordeel. De transparantie die volgens sommigen objectiverend werkt, werd niet alleen als plezierig voor de docent zelf ervaren, maar vooral ook voor het overleg met de tweede beoordelaar. Discrepancies in oordelen kunnen beter opgespoord en uitgesproken worden.

Conclusies

Hoewel de algemene beoordelaarsbetrouwbaarheid niet hoog te noemen is, verdient de analytische beoordeling toch de voorkeur van de drie onderzochte methoden: de betrouwbaarheid is beter dan bij de schaalbeoordeling en vergelijkbaar met die van de globale beoordeling; daarbij lijkt die betrouwbaarheid minder afhankelijk van de tekst. Bovendien hebben de toekomstige gebruikers, de docenten, er een duidelijke voorkeur voor.

Had de CVEN misschien gehoopt dat de schaalbeoordeling de kwaliteit van de beoordelingen zou kunnen verbeteren ten opzichte van de globale of de huidige analytische beoordeling, dan moeten onze onderzoeksresultaten in dat opzicht teleurstellend genoemd worden. De schaalbeoordeling blijft achter bij de globale en analytische beoordeling.

Voor de gevonden, betrekkelijk lage beoordelaarsbetrouwbaarheden kunnen verschillende oorzaken aangevoerd worden. In de eerste plaats was het voor de beoordelaars een onderzoekssituatie en geen examensituatie, en mogelijk is in de laatste situatie de zorgvuldigheid iets groter. Overigens vonden Sanders e.a. (1984) ook dergelijk lage betrouwbaarheden en is de beoordeling van geschreven producten een notoir lastig probleem.

Voorts kan in ons onderzoek de beoordelingstaak extra bemoeilijkt zijn doordat de beoordelaars te maken kregen met samenvattingen van leerlingen die een dergelijke opdracht waarschijnlijk niet eerder gemaakt hebben. Deze mogelijke verklaringen suggereren dat de lage beoordelaarsbetrouwbaarheden misschien een te somber beeld schetsen, maar geheel gerust mag men daar toch zeker niet op zijn. Een tweede *onafhankelijke* beoordeling lijkt dan ook

een bittere noodzaak om tot een betrouwbare beoordeling van de voor de leerlingen zo belangrijke samenvattingen te komen.

Aanbevelingen

Gegeven onze onderzoeksresultaten verdient het aanbeveling zich vooral te richten op de verbetering van de analytische beoordeling. Daarbij zou nagegaan moeten worden of de beoordelingsbetrouwbaarheid omhoog gaat als het om betere samenvattingen gaat van leerlingen die beter voorbereid zijn op de nieuwe samenvattingsopdracht (en er meer tijd voor krijgen). Binnen de examencontext verdient het ook aanbeveling meer werk te maken van de tweede beoordeling. Gezien de reeds vaak aangehaalde zware taakbelasting van de docenten Nederlands is dit uiteraard een heikel punt. Extra vergoedingen of vrijstellingen voor de docent Nederlands lijken bij dergelijk arbeidsintensieve taken onontbeerlijk om te komen tot een

goede beoordeling van (onder meer) de samenvattingsopdracht.

Tenslotte willen we nog aanbevelen na te gaan wat nu de eerder genoemde taakvariabiliteit van de nieuwe samenvattingsopdracht bepaalt. Daarmee bedoelen dat het van belang is te weten hoe iemands 'samenvatvaardigheid' afhankelijk is van de toevallige samenvattingsopdracht of brontekst. De geconstateerde grote taakvariabiliteit impliceert dat er eigenlijk meer opdrachten nodig zijn om een volledig beeld van iemands vaardigheid te krijgen. Bij de samenvattingsopdrachten is het daarom noodzakelijk te weten welke kenmerken van de opdracht, d.w.z. brontekst en opdrachtformulering, van invloed zijn op de taakvariabiliteit. Inzicht in deze kenmerken lijkt ons essentieel voor de constructie van betrouwbare en valide examens Nederlands met vergelijkbare samenvattingsopdrachten die appelleren aan een en dezelfde set vaardigheden.

Noot

We beperken ons hier tot een relatieve vergelijking van de drie beoordelingsmethoden. In absolute zin zijn de uitkomsten (d.w.z. de betrouwbaarheden) zeker niet geruststellend, te meer als men bedenkt dat het om (toekomstige) examensituaties gaat. Daar staat tegenover dat het geschetste beeld misschien iets te somber is, zoals in de conclusies gesuggereerd wordt.

Literatuur

- Aken-van de Leur, A. van, Aan alle collega's neerlandici, belast met de correctie van de samenvatting vwo, in: *Levende Talen* 435 (1988), 618-619.
- Bergh, H. van den, K. de Glopper en R. Schoonen, Directe metingen van schrijfvaardigheid: validiteit en taakeffecten, in: F.H. van Eemeren en R. Grootendorst, *Taalbeheersing*, Dordrecht: Foris Publications, 1988, 370-378.
- Bergh, H. van den, en G. Rijlaarsdam, Problemen met opstelbeoordeling? Een recept, in: *Levende Talen* 413 (1986), 448-454.
- Braet, A., Van het eindexamen-uitreksel naar een samenvatting van het betoog, in: *Levende Talen* 427 (1988), 8-14.
- CVEN, *CVEN-rapport. Eindverslag van de commissie Vernieuwing eindexamenprogramma's Nederlandse taal en letterkunde vwo en havo*, Den Haag: Sdu Uitgeverij, 1991.

- Doesborgh, P., Het bindend correctievoorschrift cse-samenvatting Nederlands vwo; standpunt van het sectiebestuur Nederlands, in: *Levende Talen* 437 (1989), 57-58.
- Drop, W., en J.H.L. de Vries, *Taalbeheersing. Handboek voor taalhantering*, Groningen: Wolters-Noordhoff, 1977.
- Eemeren, F.H. van, e.a., *Samenvatten. Nederlands 2*, 1975, 41-58.
- Glopper, K. de, en E. van Schooten, *De inhoud van de examens Nederlands voor havo en vwo*, Amsterdam: SCO, 1990. SCO-rapport 243.
- Hendrix, T., De samenvatting beoordeeld. Tekstbegrip Nederlands op het cse van het vwo, in: *Levende Talen* 405 (1985), 557-563.
- Hendrix, T., en P. Sanders, Analytische beoordeling van de samenvattingsopdracht c.s.e. Nederlands voor het vwo, in: W.K.B. Koning (red.), *Taalbeheersing in theorie en praktijk*, Dordrecht: Foris Publications, 1985, 353-361.
- Hendrix, A.C., en P.F. Sanders, De beoordeling van de samenvatting Nederlands vwo: de analytische beoordelingsprocedure in de examenpraktijk, in: *Tijdschrift voor Taalbeheersing* 9 (1987), 151-165.
- Sanders, P.F., A.C. Hendrix en A.J.M. Luijten, De beoordeling van de samenvatting Nederlands. *Tijdschrift voor Taalbeheersing* 6 (1984), 241-251.
- Schoonen, R., *Een nieuwe samenvattingsop-*

dracht voor het centraal examen Nederlands: een onderzoek naar drie beoordelingsmethoden en naar de taakvariabiliteit, Amsterdam: SCO-Kohnstamm Instituut, 1996. SCO-rapport 443.

- Sectiebestuur Nederlands, Studiedag over de samenvatting Nederlands vwo, in: *Levende Talen* 446 (1989), 781-785.
- Stuurgroep Profiel tweede fase, *Advies Examenprogramma's havo en vwo Nederlands*, Den Haag: SLO, 1995.

Rob Schoonen



Studeerde Nederlands en Algemene Taalwetenschap aan de Universiteit van Amsterdam. Sinds 1985 werkzaam aan het SCO-Kohnstammstituut en daarnaast sinds 1988

verbonden aan de vakgroep Algemene Taalwetenschap (beide UvA). Promoveerde in 1991 op onderzoek naar de validiteit van diverse vormen van schrijfvaardigheidsmeting. Adres: SCO-Kohnstamm Instituut, Wibautstraat 4, 1091 GM Amsterdam.