

Discussievaardigheid in de basisvorming

TRUDY HEUVES

Inleiding

De invoering van de basisvorming in 1993 heeft voor scholen en docenten veel veranderingen met zich meegebracht. Bij het vak Nederlands is de nadruk komen te liggen op communicatief onderwijs en op het aanleren van vaardigheden. Zo moeten docenten Nederlands in de basisvorming meer aandacht besteden aan spreekvaardigheid. Voor het domein spreekvaardigheid zijn in de basisvorming drie kerndoelen vastgesteld: het houden van een monoloog, het voeren van een dialoog en het voeren van een polyloog (discussie).

Docenten Nederlands hebben doorgaans nog weinig ervaring met het onderwijzen en beoordelen van discussievaardigheid. Dit wekt weinig verbazing als we bedenken dat bruikbaar lesmateriaal schaars is, het onderwijzen van het discussiëren in de klas op organisatorische problemen stuit en geschikte evaluatie-instrumenten vooralsnog ontbreken. Bovendien is discussievaardigheid, als onderdeel van spreekvaardigheid of communicatievaardigheid, een nog nauwelijks ontgonnen onderzoeksgebied. Dit zal ongetwijfeld samenhangen met het arbeidsintensieve en kostbare karakter van het meten van productieve taalvaardigheid. Zo vermeldt Carroll (1993) in zijn review van factor-analytische studies naar taalvaardigheid slechts één onderzoek waarin aandacht is besteed aan het participeren in groepsdiscussies.

In de kerndoelen voor de basisvorming staat de polyloog als volgt omschreven: 'Leerlingen kunnen in klas- of schoolverband aan een polyloog deelnemen met het oog op het geven of verkrijgen van de nodige informatie, het weergeven van een eigen mening en het bereiken van een voor hen aanvaardbaar resultaat.' Aansluitend bij deze onderwijsdoelstelling is in dit onderzoek gekozen voor de meningsvormende discussie. In zo'n discussie geven leerlingen hun mening over een bepaald onderwerp en anderen reageren daarop. Men is het eens of oneens en geeft daar argumenten voor. Om mee te kunnen doen aan een discussie moeten leerlingen onder meer hun mond open durven te doen, eventueel hun gevoelens durven uiten en goed kunnen luisteren naar anderen (Bonset, De Boer & Ekens, 1995).

In een groepsdiscussie spelen velerlei vaardigheden een rol; bijgevolg kan de discussievaardigheid van de leerling op velerlei aspecten beoordeeld worden (Van Gelderen, 1987; Van Gelderen & Oostdam, 1992; Rijlaarsdam & Westorp, 1984; Schreuder, 1983; 1989; Leidse werkgroep moedertaal didactiek, 1981). De vaardigheden waarop leerlingen zoal beoordeeld zouden kunnen worden, zijn hun communicatieve houding, de inhoud van het gesprokene, vaardigheid in het taalgebruik (bijvoorbeeld de woordkeus, zinsbouw, wel of geen standaardtaal) en de techniek van het spreken (bijvoorbeeld de uitspraak of de spreeknelheid) (Leidse werkgroep moedertaal didactiek, 1981). Niet al deze gedragsaspecten hoeven relevant te zijn voor docenten en leerlingen in de onderbouw van het voortgezet onderwijs.

Een speciaal probleem vormt de beoordeling van discussievaardigheid. Anders dan bij meerkeuzetoetsen, die machinaal gescoord kunnen worden, is de beoordeling van discussievaardigheid een min of meer subjectieve aangelegenheid. Het gaat tenslotte om een 'vluchtig' product, dat ter plekke beoordeeld moet worden. Natuurlijk kan een discussie opgenomen worden op video, met als voordeel dat de docent de opname nog eens rustig kan bekijken en de discussiedeelnemers kan beoordelen. Het opnemen op video heeft echter ook een keerzijde. Zo hebben niet alle docenten Nederlands de beschikking over een videocamera en kan de aanwezigheid van een camera in de klas een versturende invloed hebben op het gedrag van de

leerlingen. Bovendien is bij dit soort toetsen het oordeel zeker niet feilloos (Meuffels, 1993). In zijn algemeenheid geldt dat de objectiviteit en betrouwbaarheid van een oordeel groter wordt naarmate het op meer beoordelaars gebaseerd is. De docent kan natuurlijk een collega-docent vragen om de discussiedeelnemers te beoordelen, maar dit brengt veel organisatorische problemen met zich mee. Een andere mogelijke oplossing is het betrekken van leerlingen bij de beoordeling. Dit stelt de docent in staat om meer aandacht te besteden aan de organisatorische aspecten van de discussie. De veronderstelde meerwaarde van het inschakelen van leerlingen bij de beoordeling zal evenwel alleen optreden als de oordelen van de leerlingen serieus genomen worden en de leerlingen de discussievaardigheid op dezelfde wijze beoordelen als de docent zelf.

De validiteit van een nieuw meetinstrument kan onder meer tot uiting komen in de mate waarin het instrument een onderscheid weet te maken tussen groepen van een verschillend prestatieniveau (Thorndike, 1982). Onderzoek heeft aangetoond dat scholen in het voortgezet onderwijs sterk verschillen in het niveau van taalvaardigheid Nederlands; bovendien bestaan er (binnen scholengemeenschappen) grote prestatieverschillen tussen leerlingen van verschillende opleidingstypen (De Glopper, 1988; Kuhlemeier & Van den Bergh, 1989, Kuhlemeier, Kremers & Kleintjes, 1996). Zo bleek in de peiling Nederlands maar liefst 28% van de totale variantie in spreekprestaties verklaard te worden door de school die de leerling bezocht (Kuhlemeier, 1996). Om van een valide meting te mogen spreken, zouden de discussietoetsen deze goed gedocumenteerde prestatieverschillen tussen scholen en opleidingstypen moeten kunnen aantonen.

Tijdens de lezing heb ik verslag gedaan van de belangrijkste resultaten van een onderzoek dat is uitgevoerd met als doel te komen tot een meetinstrument dat valide en betrouwbaar de discussievaardigheid van leerlingen kan meten. Daarnaast was het van belang te onderzoeken in hoeverre deze nieuwe toetsvorm acceptabel was voor docenten en in hoeverre ze ermee konden werken in de klas (organisatie).

De eerste twee onderzoeken zijn uitgevoerd in 1994 en 1995 en hadden als voornaamste doel het ontwikkelen van een hanteerbare checklist en onderzoeken of het mogelijk was leerlingen te betrekken bij de beoordeling. Hoofdstuk 1 van dit verslag gaat nader in op de gebruikte instrumenten en in hoofdstuk 2 staan de belangrijkste resultaten van het grootschalige onderzoek dat in 1996 is uitgevoerd.

Voor meer en uitgebreide informatie over met name dit onderzoek zij verwezen naar *Tijdschrift voor Taalbeheersing* waar begin 1998 een artikel over dit onderzoek is verschenen.

1

De discussie-opdrachten. Voor alle drie onderzoeken zijn vier verschillende discussie-opdrachten ontwikkeld. Bij het eerste onderzoek handelden alle vier de opdrachten over 'Straf', bij het tweede onderzoek over 'Geld' en bij het derde onderzoek over 'Lijf en leden'. Het gaat hierbij om gestructureerde opdrachten, dat wil zeggen dat in de opdrachten een aantal situaties worden beschreven die betrekking hebben op het betreffende onderwerp met daarbij apart vermeld de mogelijke standpunten en argumenten die leerlingen tijdens de discussie mogen inbrengen. Uiteraard mogen leerlingen ook eigen standpunten en argumenten inbrengen. Voor deze gestructureerde opdrachten is gekozen opdat alle leerlingen een gelijk uitgangspunt hebben als ze beginnen met de discussie. Voorkomen moet worden dat leerlingen geen inbreng in de discussie hebben omdat ze bijvoorbeeld weinig of niets van het onderwerp afweten. Het doel van de discussie is de leerlingen hun eigen mening omtrent het onderwerp te laten weergeven.

Instructie voor de docent. In deze instructie staan onder andere aanwijzingen en suggesties voor de docent. Deze aanwijzingen hebben betrekking op het organiseren van de discussie in de klas en hoe de leerlingen zich kunnen voorbereiden op de discussie. In de instructie staat verder

dat elke discussiedeelnemer tijdens de discussie beoordeeld wordt door de docent en door medeleerlingen.

Het afnemen. Een discussie-opdracht wordt uitgevoerd door zes à acht leerlingen per klas. Het afnemen duurt ongeveer veertig tot vijftig minuten, waarvan tien à vijftien minuten voor voorbereiding, twintig tot dertig minuten voor discussie en de resterende tijd voor evaluatie. Ter voorbereiding krijgen de discussiedeelnemers de beschikking over een tekst over het onderwerp van de discussie en de bijbehorende standpunten en argumenten.

Opdracht voorzitter. De discussie wordt geleid door een voorzitter. Voor deze voorzitter is een opdracht ontwikkeld waarin aanwijzingen staan hoe de discussie te leiden. Deze aanwijzingen hebben betrekking op hoe de discussie te openen, erop te letten dat alle discussiedeelnemers de gelegenheid krijgen hun mening over het onderwerp te zeggen, erop te letten dat leerlingen bij het onderwerp blijven, te voorkomen dat er onderonsjes worden gehouden, bij gelegenheid een korte samenvatting te geven en hoe de discussie te sluiten.

De beoordeling. Elke discussiedeelnemer wordt beoordeeld door de docent en door twee of drie medeleerlingen die niet aan de discussie deelnemen. Door hoeveel leerlingen een discussiedeelnemer uiteindelijk wordt beoordeeld, is onder andere afhankelijk van de klasgrootte. Bij de beoordeling vindt geen onderling overleg plaats.

Gedragsaspecten. De beoordeling vindt plaats aan de hand van een checklist met gedragsaspecten.

2

Het derde onderzoek is uitgevoerd in het voorjaar van 1996. Uit de voorgaande twee onderzoeken bleek dat er een hanteerbare checklist was voor zowel de docent als voor de beoordelende leerlingen. De checklist bevatte relevante gedragsaspecten en de beoordelaars-overeenstemming tussen docent en leerlingen was hoog. Ook stonden er in de handleiding voldoende aanwijzingen om de discussie te organiseren in de klas. De checklist die in 1996 is gebruikt, bevatte de volgende punten:

De discussiant

- 1 *geeft een of meer standpunten.* Beoordeeld wordt in hoeverre de discussiant één of meer standpunten correct weergeeft en helder verwoordt (waarbij de discussiant eigen standpunten mag toevoegen).
- 2 *geeft argumenten die het standpunt ondersteunen.* Beoordeeld wordt in hoeverre de discussiant het standpunt met argumenten weet te ondersteunen en deze helder verwoordt (waarbij de discussiant eigen argumenten mag toevoegen).
- 3 *laat anderen uitspreken.* Beoordeeld wordt in hoeverre de discussiant anderen (storend) in de rede valt, niet laat uitspreken.
- 4 *is voor iedereen te volgen.* Beoordeeld wordt in hoeverre de discussiant voor iedereen verstaanbaar is en in een voor iedereen te volgen tempo spreekt (d.w.z. het spreektempo aanpast aan de groep).
- 5 *is betrokken bij de discussie.* Beoordeeld wordt in hoeverre de discussiant door zijn houding kenbaar maakt dat hij luistert, de spreker aankijkt, eventueel instemmend knikt, in zijn bijdragen aansluit bij wat de vorige spreker gezegd heeft, adequate vragen stelt en geen 'onderonsjes' houdt.

Beoordeling op meer dan vijf gedragsaspecten tegelijk werd niet zinvol geacht omdat dit de hanteerbaarheid voor zowel de docent als voor de leerlingen bemoeilijkt (Wesdorp, 1981). Er kan niet worden volstaan met het toekennen van 1 of 0 punten per gedragsaspect. Het kan immers voorkomen dat alle discussianten een standpunt met daarbij argumenten geven. De ene discussiant kan dit net iets beter, overtuigender doen dan een andere leerling. De beoordelaars

moeten dus bij het toekennen van punten hierin kunnen variëren. Er is gekozen voor de volgende puntenverdeling per gedragsaspect:

0 punten = gedrag wordt niet vertoond

1 punt = gedrag wordt enigszins vertoond

2 punten = gedrag wordt duidelijk vertoond

Drie jaar na invoering van de basisvorming waarin de docenten en de leerlingen uitgebreid hebben kunnen oefenen met deze nieuwe toetsvorm, mocht deze toets geen problemen meer opleveren. Besloten is dan ook om een grootschalige proefafname te organiseren waarbij scholen a-select werden uitgekozen. Aan dit onderzoek is deelgenomen door 317 leerlingen van in totaal 44 scholen (45 klassen). Er zijn twee discussietoetsen uitgezet en elke klas maakte een toets. De eerste discussietoets is afgenomen op 18 scholen bij 125 leerlingen; de tweede toets op 27 scholen bij 192 leerlingen. De verdeling van leerlingen naar opleidingstype was 30% (i)vbo, 35% mavo en 36% havo/vwo.

Bij dit derde onderzoek stonden onder andere de volgende onderzoeksvragen centraal:

- 1 In welke mate is de discussietoets acceptabel en hanteerbaar voor docenten?
- 2 In hoeverre verschilt het oordeel van de beoordelende leerlingen van dat van de docent?
Deze derde vraag bestaat uit drie deelvragen:
 - a) Maken leerlingen een even goed onderscheid tussen de discussianten als de docent?
 - b) Zijn leerlingen milder of strenger in hun oordeel dan de docent?
 - c) Zijn de oordelen van de docenten betrouwbaarder dan die van de leerlingen?
- 3 Hoe groot zijn de prestatieverschillen tussen scholen (vraag 3a) en tussen leerlingen van verschillende opleidingstypen (vraag 3b)?

3. Resultaten

3.1 Acceptabiliteit en hanteerbaarheid.

In welke mate is de discussietoets acceptabel en hanteerbaar voor docenten (onderzoeksvraag 1)? Door middel van een vragenlijst is docenten gevraagd naar hun mening over deze toetsvorm en de bijbehorende checklisten. De vragenlijst is door 51 docenten ingevuld. Enkele resultaten:

- Het onderwijzen en beoordelen van discussievaardigheid wordt nog nauwelijks in praktijk gebracht. Van de 51 docenten gaven er slechts drie aan dat hun leerlingen bekend waren met deze toetsvorm.
- De aanwijzingen voor het invullen van de checklist waren voor 84% van de docenten voldoende duidelijk; voor de overeenkomstige aanwijzingen voor de leerling bedroeg dit percentage eveneens 84%. Ook bleken de meeste docenten van mening dat de checklist hanteerbaar was voor zowel docenten (55%) als voor leerlingen (76%). Wel gaven docenten aan het beoordelen van zes of acht leerlingen tegelijkertijd moeilijk te vinden.
- Ook de gedragsaspecten werden door de meeste docenten (78%) voldoende duidelijk gevonden. Wel hadden sommige docenten problemen met het aspect 'Uitspreken'. Dit werd als moeilijk te beoordelen ervaren. Geen enkele docent vond dat er overbodige gedragsaspecten waren opgenomen, en 18% miste juist essentiële gedragsaspecten. Genoemd werden onder andere: ingaan op argumenten, taalgebruik en vragen stellen.

Al met al kan geconcludeerd worden dat de acceptabiliteit en hanteerbaarheid van de discussietoetsen redelijk was, ook al hadden zowel docenten als leerlingen nog weinig ervaring met deze nieuwe toetsvorm.

3.2 *Verschillen in beoordelingen tussen docenten en leerlingen*

In hoeverre maken de beoordelende leerlingen een even scherp onderscheid tussen de discussiedeelnemers als de docent zelf (onderzoeksvraag 2a)? Bij de eerste discussietoets maken leerlingen een even scherp onderscheid als de docent. Bij de tweede toets is er sprake van een minder scherp onderscheid.

In hoeverre zijn leerlingen milder of juist strenger in hun oordeel dan de docent zelf (onderzoeksvraag 2b)? Zowel bij toets 1 als bij toets 2 oordelen leerlingen iets milder dan de docent.

Zijn de oordelen van de docenten meer of minder betrouwbaar dan die van leerlingen (onderzoeksvraag 2c)? Betrouwbaarheid toets 1: leerlingen en docenten: .83. Betrouwbaarheid toets 2: leerlingen .73 en docenten .80

3.3 *Prestatieverschillen tussen scholen en opleidingstypen*

Een belangrijke vraag betreft de verschillen tussen scholen (onderzoeksvraag 3a). In hoeverre is het - gegeven de instrumentatie en het onderzoeksontwerp - mogelijk te differentiëren tussen scholen met hoge en lage gemiddelde prestaties voor discussievaardigheid? Van de totale prestatieverschillen in discussievaardigheid blijkt slechts een klein deel toe te schrijven aan de school die de leerling bezoekt. Er is weinig onderscheid tussen de gemiddelde prestaties van (i)vbo-mavo-havo-vwo-leerlingen.

In hoeverre wordt de discussievaardigheid van leerlingen uit de verschillende opleidingstypen verschillend beoordeeld (onderzoeksvraag 4b)? Op grond van de oordelen van leerlingen en docenten kan maar zelden een duidelijk onderscheid worden gemaakt tussen de gemiddelde prestaties van vbo-, mavo- en havo/vwo-leerlingen. Bij toets 1 beoordelen mavo-leerlingen de discussievaardigheid van hun klasgenoten wat lager dan vbo-leerlingen.

4. **Conclusie**

Al met al kunnen we concluderen dat de acceptabiliteit en hanteerbaarheid van deze nieuwe toetsvorm akkoord is. Docenten geven wel aan problemen te hebben met het tegelijkertijd beoordelen van 6 à 8 leerlingen. Uit dit onderzoek blijkt dat het oordeel van de docent sterk samenhangt met dat van de leerlingen. Op basis hiervan zouden de docenten kunnen beslissen de beoordeling aan de leerlingen over te laten en zelf de organisatorische aspecten en/of de voorzittersrol op zich te nemen.

Te verwachten was dat VBO-leerlingen lager zouden scoren dan VWO-leerlingen, maar dit bleek niet zo te zijn. Er is dus weinig verschil in prestaties tussen verschillende opleidingstypen. Tussen leerlingen binnen scholen zijn wel duidelijke verschillen aan te wijzen.

Bibliografie

- Bonset, H., M.de Boer & T.Ekens, *Nederlands in de basisvorming. Een praktische didactiek*. Bussum: Coutinho, 1995
- Carroll, J.B., *Human cognitive abilities. A survey of factor-analytic studies*. Cambridge: University Press, 1993
- Gelderens, A van, *Taalmaten, constructie van gedetailleerde beoordelingsprocedures ten behoeve van peilingsonderzoek, deel 1: het beoordelen van spreekprestaties*. Amsterdam, SCO-reeks 128, 1987

- Glopper, K. de**, *Schrijven beschreven. Inhoud, opbrengsten en achtergronden van het schrijfonderwijs in de eerste vier leerjaren van het VO*. Den Haag: SVO, 1988
- Kuhlemeier, J.B.**, *Taalvaardigheid, taalactiviteiten en taalattitudes: een validatiestudie*. Proefschrift. Arnhem: Instituut voor Toetsontwikkeling, 1996
- Kuhlemeier, J.B. & Bergh, H. van den**, *De Proefpeiling Nederlands: een onderzoek naar de haalbaarheid van peilingsonderzoek in het voortgezet onderwijs*. Specialistisch Bulletin Nr. 74. Arnhem: Instituut voor Toetsontwikkeling, 1989
- Kuhlemeier, J.B., E.J.J Kremers & F.G.M. Kleintjes**, *Gebruik en moeilijkheid van de eerste generatie afsluitingstoetsen basisvorming in het schooljaar 1994/95*. Arnhem: Instituut voor Toetsontwikkeling, 1996
- Leidse werkgroep moedertaal didactiek, Moedertaal didactiek. Een handleiding voor het voortgezet onderwijs. Muiderberg: Coutinho, 1991**
- Meuffels, B.**, *De verguisde beoordelaar: opstellen over opstelbeoordeling*. Amsterdam: Thesis Publisher, 1993
- Rijlaarsdam, G. & H. Wesdorp**, *Het beoordelen van taalvaardigheid in het onderwijs*. Amsterdam: SCO-rapport 28, 1984
- Schreuder, K.**, *Spreekvaardigheid Nederlands. Voorstellen voor toetsconstructie*. Documentatierreeks nr. 112. Arnhem: Instituut voor Toetsontwikkeling, 1983
- Schreuder, K.**, *Spreektoetsen Nederlands voor onderbouw lbo/mavo*. Bulletinreeks nr. 73. Arnhem: Instituut voor toetsontwikkeling, 1989
- Wesdorp, H.**, *Evaluatietechnieken voor het moedertaalonderwijs*. Den Haag: SVO, 1981