

Breder evalueren

Bij toetsontwikkeling wordt vaak ingezet op de ontwikkeling van op grote schaal inzetbare toetsen. Vaak draait het daarbij in de eerste plaats om de eenmalige, globale beoordeling van 'einde-van-de-rit' prestaties (summatieve evaluatie).

Wie A zegt,

Om verschillende redenen is er echter behoefte aan bredere manieren van evalueren. Manieren die vertrekken vanuit de behoeftes en daaraan gekoppelde doelstellingen van de leerder. Manieren waarbij je net zo goed validiteit en betrouwbaarheid kunt bereiken als met eenmalige toetsen die in het teken staan van vergelijkbaarheid, efficiëntie, standaardisering, ... Daarbij kan de invalshoek van waaruit we evalueren met betrekking tot een aantal aspecten verruimd worden. Hoe doe je dat?

moet ook B zeggen

Hieronder een Vlaamse proeve van uitwerking van de hand van Katrien Bultynck.

In de laatste decennia is onze kennis over en onze visie op tweedetaalverwerving sterk geëvolueerd.

Vanuit Vlaanderen bekeken kan je die veranderende visie op tweedetaalverwerving als volgt omschrijven:

- 1 De aandacht is verschoven van 'taalkennis' naar 'taalvaardigheid': centraal komt te staan wat een leerder in het Nederlands moet kunnen in functionele, voor hem relevante situaties.
- 2 Taalleren is wel cumulatief maar niet lineair. Een leerder bouwt weliswaar verder op wat hij al kan. Maar hoe en wat hij daarop verder bouwt, dat kan voor elk individu verschillend zijn.
- 3 Tweedetaalleren gaat meteen ook over kennis van de wereld opbouwen. Cognitieve vaardigheden en leerstrategieën spelen een rol in het leerproces evenals affectieve factoren als motivatie, betrokkenheid en attitudes.
- 4 Zicht hebben op je eigen leerproces en er mee de verantwoordelijkheid voor dragen, blijkt één van de belangrijke factoren te zijn die het succes van een leerproces beïnvloeden.
- 5 In een gestuurd taalproces zijn instructie en evaluatie onlosmakelijk met elkaar verbonden. Permanente feedback is dus onontbeerlijk om de leerder zicht op zijn eigen leerproces te bieden (cf. punt 4).

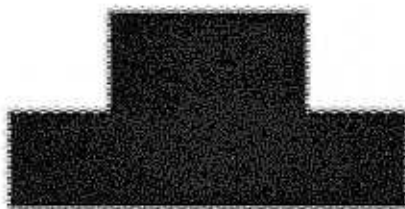
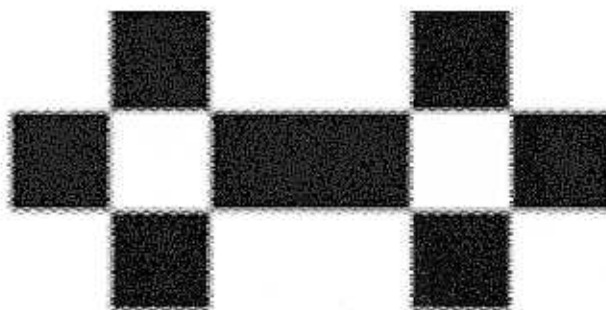
Deze veranderingen vertalen zich langzaam maar zeker in de dagelijkse Nt2-lespraktijk. Zo wordt bijvoorbeeld door meer en meer taallessgevers in Vlaanderen gebruik gemaakt van functioneel en taakgericht lesmateriaal.⁽¹⁾

Taalevaluatie maakt ook deel uit van de lespraktijk. Op dit vlak zou dus een soortgelijke evolutie aan de gang moeten zijn.

En inderdaad: er wordt hard gewerkt aan vormen van assessment die op een geïntegreerde manier de taalvaardigheid van de kandidaat meten, eerder dan de taalkennis. Het draait daarbij echter meestal in de eerste plaats om de eenmalige, globale beoordeling van 'einde-van-de-rit' prestaties (summatieve evaluatie). De toets wordt daarbij meestal ontwikkeld met een welomschreven algemeen 'niveau' als uitgangspunt. Bovendien worden de vier vaardigheden meestal geïsoleerd getoetst.⁽²⁾

Door deze ontwikkelingen komen bovendien lesgever en taallearder steeds verder van het evaluatieproces af te staan.

De rol van de docent beperkt zich tot het afnemen van een elders ontwikkelde toets bij zijn cursistengroep. De resultaten worden vaak centraal op



een andere plaats verwerkt. Lesgevers hebben dus nog weinig zicht op: 'Wat wordt er eigenlijk getoetst?', 'Waarom?' en 'Welke informatie kan ik uit de resultaten halen om het leerproces bij een cursist beter of anders te ondersteunen?'

Hetzelfde geldt voor de taalleerder, van wie we in de inleiding toch stelden dat hij een actieve rol zou moeten spelen. Zijn rol wordt beperkt tot de rol van 'object'. De informatie die wordt verzameld is niet voor hem maar gaat enkel over hem.

Zonder twiifel is het belangrijk dat aandacht aan de ontwikkeling en de verspreiding van de hierboven gescherste instrumenten wordt besteed.

In de eigen klas, stellen lesgevers vast, is de gestandaardiseerde toets die wordt aangeboden ongetwijfeld zinvol voor bepaalde doeleinden en in bepaalde situaties. Maar wat bijvoorbeeld als op het einde van de rit blijkt dat een cursist niet geslaagd is? Het is dan te laat om er iets uit te leren.

Zowel de lesgever als de cursist hebben behoefte om de eigen didactiek respectievelijk de eigen leerwijze bij te sturen.

Hieronder geef ik een proeve van uitwerking van 'breder' evalueren.

De term 'breder evalueren' geeft daarbij aan dat bestaande evaluatievormen hun plaats vinden in een breder programma van evalueren dat nog onvoldoende is ingevuld.

Instrumenten op maat: een voorbeeld graag

Het voorbeeld is ontleend aan een observatie-instrument uitgewerkt bij het programma De Lirt. De Lirt is een functioneel en behoeftegericht voortraject Nt2 voor anderstalige volwas-

sen, ontwikkeld aan Centrum voor Taal en Migratie (Katholieke Universiteit Leuven) in opdracht van de Vlaamse Dienst voor Arbeidsbemiddeling en Beroepsopleiding (VDAB). Het bereidt deelnemers op een taakgerichte manier voor op een beroepsopleiding voor administratieve beroepen⁽¹⁾ bij dezelfde VDAB.

We proberen het observatie-instrument hier kort te beschrijven. Daarna zullen we het gebruiken om te illustreren: hoe gaat het CTM concreet te werk bij de ontwikkeling van een instrument voor bredere evaluatie?

Het voortraject Nt2 neemt minimaal 400 uren in beslag. Dat is lang voor volwassen werklozen die zo snel mogelijk aan het werk willen. Van sommige cursisten voelt een lesgever bij de start al intuïtief aan: 'Ik twiifel of deze cursist het wel zal halen.' VDAB wil vermijden dat zo'n cursist die het moeilijk heeft, pas op het einde van het Nt2 traject (pas na 400 uur dus) te horen krijgt: u heeft een probleem en mag niet doormoeten naar de beroepsopleiding.

Wat evalueren, waarom en hoe?

Er was dus behoefte aan een instrument om zo vlug mogelijk in het traject problemen in het taalleerproces op te sporen en indien mogelijk bij te sturen. Na 100 uur zou VDAB kunnen beslissen over voortzetting van het gekozen Nt2-traject, dan wel heroriëntering. Dit op basis van het antwoord op de volgende vragen:

- Hoe verloopt het taalleerproces van de risicocursist in die eerste 100 uur?
- Is het op basis daarvan haalbaar en zinvol, voor de cursist en de VDAB, dat de cursist het traject voortzet?
- Indien ja, zijn daar bepaalde voor-

waarden aan verbonden? Bijvoorbeeld: nog bijkomende ondersteuning met betrekking tot bepaalde zwakke punten.

Naast de ruimere visie op tweedetaalverwerving omschreven in de inleiding, lagen de volgende uitgangspunten ten grondslag aan de manier waarop het instrument is opgezet:

In De Lirt worden cursisten op een cyclische manier geconfronteerd met relevante, functionele taken die hen stimuleren in het bereiken van de einddoelen. In elk van die taken zit een kloof. Een doelbewuste ondersteuning⁽²⁾ van de lesgever helpt cursisten om deze kloof te overbruggen. Als het leerproces goed verloopt, worden cursisten hierdoor zelfstandiger in het aanpakken van taken en het overbruggen van mogelijke kloven daarin. De ondersteuning van de lesgever zal met andere woorden dus stilaan minder intensief worden.

Vandaar onze keuze in de opzet van het instrument: de evolutie van het taalleerproces in kaart brengen via het registreren van de intensiteit en de aard van de ondersteuning die een risicocursist bij het uitvoeren van vergelijkbare cruciale taken in de les nodig heeft.

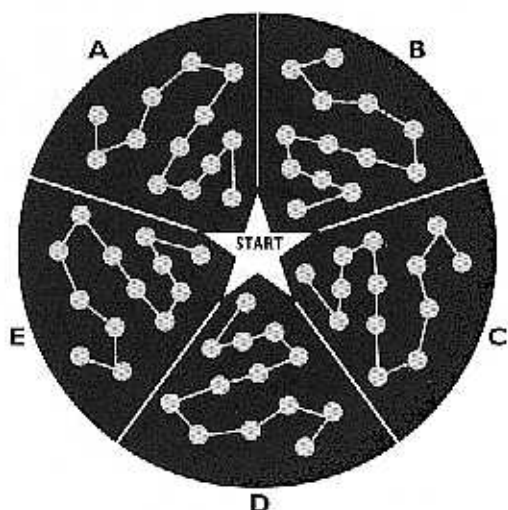
Daartoe selecteerden we uit de eerste 100 uren 47 observatiemomenten die we konden onderbrengen in 5 categorieën van cruciale, steeds weerkerende typetaken. De omschrijving van die vijf categorieën van taken, kan je terugvinden op illustratie 1.

Voor elk van deze observatiemomenten werd een observatiefiche uitgeschreven. Daarop staat aangegeven tijdens welke handelingen van de cursist de lesgever hem moet observeren en welke gegevens hij moet registreren.

Op deze observatiefiches moet de lesgever invullen/inkleuren of een risicocursist bij het uitvoeren van de taak nood had aan:

- heel intensieve ondersteuning (bolletje rood inkleuren);
- redelijk wat ondersteuning (bolletje oranje inkleuren);
- weinig of geen ondersteuning. (bolletje groen inkleuren).

Ook moet hij informatie noteren over de aard van de ondersteuning die de cursist nodig had. Met 'aard' van ondersteuning bedoelen we: Waar liep de cursist op vast? Om welke taalelementen ging het? Was de kloof te groot wegens een gebrek aan kennis van de wereld? Liep de cursist vast



A LUISTEREN/SCHRIJVEN

Luisteren naar één of meer korte dialogen en de info daaruit overbrengen in een voorgestructureerd document.

B LEZEN/LUISTEREN/SCHRIJVEN

Een kort berichtje schrijven om iets mee te delen of te vragen (op basis van geschreven of mondelinge info en instructie).

C LEZEN/SCHRIJVEN

Uit korte informatieve tekst(en) info halen en die in een voorgestructureerd document invullen.

D LEZEN/LUISTEREN/SPREKEN

In een (telefoon)gesprek vooraf gekende info geven of vragen.

E LEZEN/SPREKEN

In een gesprek met een medecursist mondeling info rapporteren op basis van een schriftelijke bron.

omdat hij niet doorhad waar hij de nodige informatie moest gaan zoeken? Kon hij zich niet in het vereiste perspectief verplaatsen als hij een telefoongesprek moest voeren?

De intensiteit van de ondersteuning (rood, oranje of groen) kleurt de lesgever later in op een overzichtsdocument over de 100 eerste lessen. Dat overzichtsdocument werd 'De leercirkel' gedoopt (zie illustratie). Elk bolletje in de leercirkel geeft een observatiemoment weer.

In het overzichtsdocument wordt al vrij snel duidelijk: met welke categorieën van taken heeft de cursist het vooral moeilijk en met welke niet?

Beslissingen nemen met het instrument

Als na 100 uur de hele leercirkel is ingekleurd, krijg je een overzicht dat je toelaat te beoordelen: is er evolutie in de intensiteit van de ondersteuning of niet? Op basis daarvan kan dan een beslissing tot voortzetting of heroriëntatie worden genomen.

We beschrijven hier kort enkele van de mogelijke scenario's.

Scenario 1: de inkleuring evolueert over het algemeen van rood (zeer intensieve ondersteuning) naar minder intensief (oranje en/of groen). Op basis van de verzamelde observatiegegevens blijkt dus dat een cursist bij vergelijkbare taken minder en minder ondersteuning nodig heeft. Met andere woorden: er is een gunstige evolutie in het taalproces; de cursist kan met kans op slagen zijn traject in De Lift verderzetten.

Scenario 2: de inkleuring blijft over het algemeen rood. De intensieve ondersteuning van de lesgever levert voor de cursist geen leerwinst op. De kans is zeer klein dat de cursist binnen de gegeven omstandigheden zal slagen in zijn traject. Het is dan ook weinig zinvol - en weinig realistisch - een dergelijke intensieve ondersteuning voor de rest van het traject te blijven voorzien. In overleg met de trajectbegeleider moet dan gezocht worden naar een andere weg voor de cursist om zijn doel te bereiken. Of moet de cursist geholpen worden om zijn doelen bij te stellen.

Enkele handvatten

Bij het opzetten van een dergelijk instrument bleek het voor ons een goede werkwijze om je invalshoek te verruimen met betrekking tot de volgende aspecten:

- het tijdsaspect;
- aard van de verzamelde informatie;
- betrokken beoordelaars;
- manieren van verzamelen van informatie.

Je schuift als het ware met deze aspecten als handvatten over een as van beperkter naar breder. (zie schema hieronder).

We beschrijven het schuiven met betrekking tot deze aspecten één na één en illustreren dit aan de hand van het observatie-instrument bij De Lift. Daarbij toetsen we meteen ook hoe het schuiven over deze assen bijdraagt aan en strookt met een ruimere visie op taallessen.

Het tijdsaspect

De meeste evaluatie-instrumenten beperken zich tot een éénmalige momentopname.

Door op verschillende momenten tijdens de cursus de leerder op bepaalde aspecten te beoordelen en te vergelijken kan je de betrouwbaarheid van je informatie verhogen. Bovendien verzamel je informatie die voor de bijsturing van het leerproces relevanter kan zijn.

Via het observatie-instrument worden beslissingen over heroriëntatie of bijkomende ondersteuning niet genomen op basis van het kunnen uitvoeren van een taak op één moment. Wel op basis van een vergelijking van de ondersteuning die een cursist nodig had in het uitvoeren van cruciale vergelijkbare taken gedurende de eerste 100 lessen.

Dit continu observeren en evalueren maakt het meteen ook mogelijk om instructie en evaluatie sterker met elkaar te verweven. Een lesgever merkt zo bijvoorbeeld na twee of drie vergelijkbare taken op dat de cursist het moeilijk heeft omdat hij niet gericht op zoek gaat naar informatie in een document maar systematisch alles probeert te lezen. In een volgende vergelijkbare taak kan de lesgever zijn ondersteuning dan gericht hierop afstemmen.

Aard van de verzamelde informatie

Meestal geeft een toets informatie over de talige prestaties van een leerder. We kunnen daarnaast ook informatie verzamelen over andere aspecten zoals kennis van de wereld, cognitieve elementen, attitude, motivatie, thuisituatie van de leerder. Dit kan een ander licht werpen op het al dan niet

vertonen van de gewenste taalvaardigheid; bovendien kan het helpen om de ondersteuning van de leerder gericht en efficiënter te maken.

Zo wordt in het observatie-instrument niet alleen kwantitatieve informatie verzameld: Hoeveel hulp had de cursist nodig? En: evolueert hij daarin bij een vergelijkbare opdracht? Tegelijkertijd wordt ook gekeken naar de aard van de ondersteuning. Lesgevers moeten ook registreren: wat was het signaal om individuele hulp aan de cursist te geven, welke gerichte vragen heb je gesteld om crichter te komen waarop de cursist vastliep, welke acties heb je ondernomen om de cursist zelf tot een oplossing te laten komen? Was deze diagnose de juiste, met andere woorden: heeft je gerichte ondersteuning effect gehad? Via het observatie-instrument kwamen een lesgever en een cursiste bijvoorbeeld tot de volgende vaststelling: de cursiste slaagde er nooit in een tele-

BREDERE EVALUATIE: schuiven op 4 assen

(Verbreding van aantal invalshoeken schuiven steeds in functie van de kernvragen: *Wat? Waarom? Hoe?*)

Eénmalige momentopname

TIJD Doorheen de tijd, op verschillende momenten

Correctie van toets door taallesgever

BETROKKEN
BEOORDELAARS Niet alleen de lesgever, maar ook
Self-assessment, Peer assessment
Meer dan één observator
Een team van betrokkenen (waaronder ook de
taallesgever zelf) beoordeelt de informatie en neemt
een beslissing

Genereert enkel info over kennis van taalelementen

AARD VAN DE
VERZAMELDE
INFORMATIE Niet enkel over het al dan niet kennen van bepaalde
taalelementen maar ook
talig functioneren in het uitvoeren van realistische,
relevante taken
m.b.t. taalproces
m.b.t. andere aspecten die taalproces beïnvloeden:
Cognitieve vaardigheden, kennis van de
wereld, attitude, leerstrategieën, persoonlijkheid en
persoonlijke situatie van de leerder, ...

Louter registreren van het al dan niet slagen in aflevering van gevraagde eindproduct

MANIER VAN
VERZAMELEN VAN
DE INFORMATIE Gebruik maken van alle mogelijke gelegenheden die
zich aandienen:
Rechtsreeks observeren tijdens taalproces
Video-opnames
Gesprekken met leerder, peers, begeleider, ...
Bijhouden van een logboek

foongesprek tot een goed einde te brengen; zij begon steeds totaal ongepland en impulsief aan het telefoongesprek. Lesgever en cursist spraken dan ook af bij een volgende telefoondopdracht eerst expliciet stil te staan bij: 'Wie belt hier naar wie?' en 'Wat moet ik eigenlijk zeggen'.

Lesgevers geven aan dat ze sinds het gebruik van het instrument veel duidelijker detecteren waarmee de cursist het moeilijk heeft. Ze beseffen dat hun ondersteuning tevoren vaak niet verder ging dan het uitleggen van moeilijke woorden of het herhalen van de instructie.

Betrokken beoordelaars

Naast de lesgever of testafnemer, moeten ook de leerder zelf en andere betrokkenen (tweede observator, medecursisten, trajectbegeleiders, ...) een rol in het beoordelingsproces kunnen opnemen. Dit wil niet zeggen dat we zomaar automatisch alle heil van zelf- of peerevaluatie verwachten. Het is vooral opnieuw een manier om onze blik te verruimen: door zowel de invalshoek van de leerder, de lesgever en andere betrokkenen samen te brengen, kunnen we de betrouwbaarheid van onze informatie verhogen. Bovendien leidt het betrekken van de leerder in de evaluatie ertoe dat hij mee de verantwoordelijkheid kan opnemen.

In het observatie-instrument is het weliswaar de lesgever die de observatiefiche invult. Maar deze observatiefiche doet meteen ook dienst als basis voor overlegmomenten. In de eerste plaats kan dit overleg helpen om de puzzel vanuit de leerder verder aan te vullen met gegevens die de lesgever vanuit zijn positie in de les niet kan zien. Voor een concreet voorbeeld hiervan zie het punt 'manieren van verzamelen van informatie' verderop. Daarnaast helpen deze overlegmomenten de cursist om een duidelijker zicht op zijn eigen leerproces te krijgen. Sommige - en vaak ook net de minst sterke - cursisten hebben het hier heel moeilijk mee: ze hebben de indruk dat het eigenlijk allemaal wel goed loopt; of ze hebben het moeilijk om te voelen waarop ze vastlopen. Daardoor hebben ze ook geen enkel houvast om er iets aan te doen. Door op basis van de observaties van de lesgever te bekijken: 'Wat liep er mis' en 'Waarom kan dat liggen' krijgen ze concretere aanknopingspunten om

er zelf iets aan te doen. Denk maar aan de cursiste van hierboven in de telefoongesprekken.

Door de cursist al vanaf het begin in deze evaluatie te betrekken, kan hij ook beter zijn rol opnemen in het beslissingsmoment na 100 uur. Hij zal gedurende de 100 uur duidelijker voelen: 'Dit lukt me nu al veel beter.' Of hij zal zelf hebben ondervonden: 'Dit is echt niets voor mij.' De beslissing is dan veel minder een verdriet dat na 100 uur plots over hem valt.

Manier van verzamelen van informatie

Lesgevers beseffen dat niet enkel de georchestreerde toetsmomenten relevante informatie kunnen opleveren. Ze hebben meestal wel intuïtief een zicht op hun cursisten op basis van rechtstreekse observaties tijdens de les, video opnames, gesprekken met leerders, logboeken. Een instrument opzetten waarin deze manieren van informatieverzameling een expliciete plaats krijgen, kan hierin meer systematisch en consistentie brengen.

In het observatie-instrument is niet gekozen voor aparte toetsmomenten: de lesgever observeert de cursist tijdens de lessen zelf. Tegelijkertijd wordt hij ook aangemoedigd om alle dingen die hem opvallen te noteren. Of om de obligate woordenschat- of grammarcarnersen te leggen naast de observaties uit de les.

De overlegmomenten met de cursist kunnen ook weer andere informatie opleveren. Of de lesgever helpen om zijn beeld bij te stellen. Zo observeerde een lesgever dat een risicocursiste zeer intensieve ondersteuning bleef nodig hebben in alle taken. De cursiste kwam vaak te laat, leek weinig betrokken en weinig geconcentreerd. In de rubriek 'Wat is je vandaag nog opgevallen bij de cursist?' van de observatiefiche noteerde de lesgever op een dag: 'een deel van de notities van de cursiste zijn verscheurd.' Een eerste indruk van de lesgever had kunnen zijn dat de cursiste erg nonchalant was. In een overlegmoment vertelde de cursiste echter dat haar echtgenoot niet wilde dat zij Nederlandse les volgde. Daarom had hij haar notities verscheurd. Dit wierp meteen een heel andere blik op het gebrek aan concentratie van de cursiste.

Lesgevers gevraagd!

Uit het voorbeeld hierboven mag blijken dat bij bredere evaluatie nog meer maatwerk komt kijken. Lesgevers zijn de meest aangewezen personen om - naast de leerders zelf - hierin een belangrijke rol op te nemen. Zij hebben daarvoor behoefte aan op maat ontwikkelde instrumenten. Of althans aan concrete voorbeelden waarop ze zich kunnen inspireren: voor elke behoefte specifieke instrumenten ontwikkelen, lijkt weinig haalbaar.

Bovendien is het stellen van een bredere diagnose - zelfs als je een concreet uitgewerkt instrument hebt - voor lesgevers niet zomaar evident.

Twee redenen waarom ook moet geïnvesteerd worden in ondersteuning van lesgevers. Het gaat dan om de ontwikkeling van leerkrachtvaardigheden die zowel de instructie als de evaluatie ten goede komen: iemand ondersteunen om vanuit een bredere invalshoek te evalueren, betekent meteen ook dat je hem helpt om een betere lesgever te worden.

Katrien Bultynck

De auteur is wetenschappelijk medewerker aan het Centrum voor taal en Migratie (Katholieke universiteit Leuven) en vooral actief in de ontwikkeling van lesmaterialen en instrumenten voor evaluatie in de sectie Volwassenonderwijs van het Centrum.

Literatuur

- D. Allwright & K. Bailey, *Focus on the language classroom: An introduction to classroom research for language teachers*. Cambridge, Cambridge University Press, 1991 & 1995.
J.M. Norris, J.D. Brown, T. Hudson & J. Yashinski, *Designing Second Language Performance Assessments*. University of Hawaii at Manoa, 1998.
P. Van Petegem & J. Vonhof, *Evaluatie op de testbank. Een handboek voor het ontwikkelen van alternatieve evaluatiemethoden*. UM Departement Didactiek en Pedagogiek, 2002.
Shohamy Eloni, *The power of tests: a critical perspective on the uses of language tests*. ALTE, Salamanca, November 13, 2002.

Noten

1. Zie bijvoorbeeld het artikel van Marleen De Schutter over De Lift in het mininummer van Les 2004.
2. Zie bijvoorbeeld het artikel van Anne-Mieke Janssen - Van Dienen over voordelen en beperkingen van gestandaardiseerde toetsen in Les 110, april 2001.
3. Voor een uitgebreidere beschrijving van De Lift zie ook het artikel van Marleen De Schutter in het mininummer van Les 2004.
4. Met 'ondersteuning' bedoelen we dan elke hulp van de lesgever (of medecursist) om de cursist de taak te laten uitvoeren. Niet alleen het uitleggen van moeilijke woorden, maar ook: vragen stellen om de cursist op het juiste pad te zetten, de onwetende leerder spelen om reacties bij de cursist uit te lokken, zijn aandacht op belangrijke elementen richten, ...