

Ten Geleide

Hoewel de eindtermen voor het vreemdetalenonderwijs gericht zijn op communicatieve vaardigheden, blijkt in de praktijk dat veel leerlingen maar beperkt (leren) communiceren in de vreemde taal. Om een brug te slaan tussen de vakdidactische theorie en de vakdidactische ontwikkeling van leraren-in-opleiding in de (stage)praktijk hebben Charline Rouffet, Irene de Kleyn, Marleen IJzerman, Catherine van Beuningen en Rick de Graaff een observatie-instrument ontwikkeld dat gebaseerd is op principes van communicatief taalonderwijs, en dat het vakdidactisch handelen van leraren moderne vreemde talen (in opleiding) observeerbaar en evalueerbaar maakt. Het instrument kan zowel door lerarenopleiders en stagebegeleiders als door leraren-in-opleiding zelf gebruikt worden voor observeren, feedback geven of ontwerpen van onderwijs. Hun artikel laat zien hoe het observatie-instrument is geëvalueerd op validiteit, betrouwbaarheid en praktische bruikbaarheid en hoe het kan worden ingezet.

Camille Welie en Maarten de Korte beschrijven een hardop-denkstudie waarin eerstejaars hbo-studenten met een kleine en een grote academische woordenschat met elkaar vergeleken zijn op hun begrip van studieteksten, toepassing van leesstrategieën en leesmotivaties. Uit de resultaten blijkt dat hbo-studenten met een kleine woordenschat een lager begrip van studieteksten hadden en minder gemotiveerd waren om studieteksten te lezen. Bovendien blijken ze tijdens het

lezen minder woorden te kennen, minder onbekende woorden op te merken en minder hogere orde- inhoudsgerichte leesstrategieën toe te passen dan studenten met een grote woordenschat.

Jasmijn Stolvoort, Joyce Meuwissen, Megan Mackaaij, Teun Boekel en Elena Tribushinina brengen verslag uit van een interventieonderzoek waarin basisschoolleerlingen met een taalontwikkelingsstoornis (TOS) in het speciaal onderwijs Engelse lessen kregen via een meertalige versie van de CodeTaal-aanpak. CodeTaal is een lesmethode die taalbewustzijn vergroot en rekening houdt met de beperkingen van kinderen met TOS. De resultaten laten zien dat een- en meertalige leerlingen vooruit zijn gegaan in woordenschat en grammatica. Daarnaast hebben ze een woordleerstrategie aangeleerd waarbij ze nieuwe cognaten makkelijker konden herkennen. De leerlingen die de meeste moeite vertoonden met Engels vóór de interventie toonden de grootste leeropbrengsten. Daarnaast bleek dat leerlingen de interventielessen even leuk vonden als de reguliere lessen.

Veerle Baaijen bespreekt het proefschrift van Liselore van Ockenburg: *Synthesis writing: Teaching high school students how to read, plan, draft, and revise*. Jeroen Dera bespreekt het proefschrift van Martijn Koek: *Think twice: Literature lessons that matter*.

Namens de redactie
HELGE BONSET

Een vakdidactisch observatie-instrument voor communicatief vreemdetalenonderwijs:

Brug tussen theorie en praktijk

CHARLINE ROUFFET, IRENE DE KLEYN, MARLEEN IJZERMAN,

CATHERINE VAN BEUNINGEN & RICK DE GRAAFF

Vakdidactiek vormt een essentieel onderdeel van het curriculum van lerarenopleidingen moderne vreemde talen en principes van communicatief taalonderwijs zijn hierin een belangrijk uitgangspunt. De samenhang tussen de vakdidactische leerlijn en de (stage)praktijk is echter vaak beperkt. Dit komt mede doordat de opleiders van de lerarenopleiding en de stagebegeleiders in de beroepspraktijk onvoldoende samenwerken vanuit een gedeeld vakdidactisch referentiekader. Hoewel de eindtermen voor het vreemdetalenonderwijs gericht zijn op communicatieve vaardigheden, blijkt in de praktijk dat veel leerlingen¹ maar beperkt (leren) communiceren in de vreemde taal. Om een brug te slaan tussen de vakdidactische theorie en de vakdidactische ontwikkeling van leraren-in-opleiding in de (stage)praktijk is een observatie-instrument ontwikkeld dat gebaseerd is op principes van communicatief taalonderwijs, en dat het vakdidactisch handelen van leraren moderne vreemde talen (in opleiding) observeerbaar en evalueerbaar maakt. Het instrument kan zowel door lerarenopleiders en stagebegeleiders als door leraren-in-opleiding zelf gebruikt worden voor observeren, feedback geven of ontwerpen van onderwijs. Deze

studie laat zien hoe het observatie-instrument is geëvalueerd op validiteit, betrouwbaarheid en praktische bruikbaarheid en hoe het kan worden ingezet.

In Nederland zijn de vakdidactische competenties die nodig zijn om communicatief taalonderwijs (CTO) vorm te geven, opgenomen in de landelijke hbo-kennisbasis voor de verschillende moderne vreemde talen (tovoorde- leraar, 2017) en ingebed in de vakdidactische cursussen van de lerarenopleidingen. Hoewel leraren moderne vreemde talen (mvt) opgeleid worden om leerlingen de vreemde taal te leren teneinde te kunnen communiceren, lijkt de focus in de onderwijspraktijk vaak sterk te liggen op grammatica, correctheid en het leren van losse woorden zonder context (West & Verspoor, 2016; Graus, 2018). Deze aandacht voor grammatica- en woorden- schatoefeningen gaat ten koste van het leren toepassen van taalkennis en -vaardigheid in betekenisvolle, realistische communicatieve situaties (Long, 2009), met name waar het gaat om productief taalgebruik (spreken en schrijven). Leerlingen kennen wel grammaticaregels, maar kunnen deze maar beperkt benutten in een 'levensechte' situatie. CTO is

echter het meest motiverend en effectief als de taal betekenisvol wordt gebruikt met veel authentieke input en interactiemogelijkheden in de doeltaal (Brandl, 2008; Lightbown & Spada, 2013). Hoewel leraren-in-opleiding (lio's) de noodzaak inzien om talen op een communicatieve manier te onderwijzen, slagen zij er dikwijls niet in om deze benadering daadwerkelijk in hun dagelijkse lespraktijk te implementeren.

In dit artikel bespreken we de ontwikkeling, validering en praktische bruikbaarheid van een vakdidactisch observatie-instrument, gericht op het overbruggen van de beschreven kloof tussen theorie en praktijk in het vreemdetalenonderwijs (De Graaff, 2018).

Achtergrond en probleemstelling

Veel onderzoekers hebben gewezen op de moeite die lio's in het algemeen, en lio's mvt in het bijzonder, ondervinden bij het vertalen van hun theoretische kennis naar de lespraktijk, vooral wanneer de theorie niet overeenkomt met de werkelijkheid op hun stage-school (Farrell & Bennis, 2013; Kelly, 2018). Deze discrepantie kan te wijten zijn aan contextuele factoren zoals het aantal lessen per week, het aantal leerlingen in de klas, het beschikbare materiaal of de toetsing (Borg, 2003). Maar er kunnen ook meer conceptuele factoren meespelen, zoals de overtuigingen van collega's of begeleiders en van de lio zelf (Tang et al., 2012; Graus, 2018). Soms vinden lio's het moeilijk om principes van CTO toe te passen, omdat zij nog weinig ervaring hebben, het gevoel hebben dat zij tijdens hun stages de praktijk van hun begeleider moeten weerspiegelen, zich moeten conformeren aan specifieke verwachtingen of simpelweg de steun missen om iets nieuws te proberen (Ogilvie & Dunn, 2010). Ze missen ook vaak goede voorbeelden van vakdidactisch hande-

len in de praktijk (Graus, 2018).

Stagebegeleiders op stagescholen worden getraind om lio's te begeleiden en te beoordelen in de praktijk. Deze (meestal vakoverstijgende) trainingen zijn echter voornamelijk gericht op het begeleiden en beoordelen van de algemene pedagogisch-didactische vaardigheden van lio's en minder op hun vakdidactische competenties. Een belangrijk middel om lio's te begeleiden en hun handelen te evalueren zijn lesobservaties. Hiermee wordt inzichtelijk gemaakt hoe het handelen van de lio in de praktijk zich verhoudt tot de beoogde eindtermen, zoals geformuleerd in de landelijke hbo-kennisbases. Gevalideerde observatie-instrumenten die tegenwoordig binnen veel opleidingen door lio's, stagebegeleiders en lerarenopleiders gebruikt worden, zoals het ICALT (International Comparative Analysis of Learning and Teaching, RUG, 2019 of Maulana, Helms-Lorenz & Van de Grift, 2017) zijn vooral pedagogisch en algemeen didactisch van aard; ze bevatten indicatoren die betrekking hebben op aspecten als orde houden, het creëren van een veilig leer-klimaat, instructie geven en vragen stellen. Deze observatie-instrumenten geven weinig informatie over de vakdidactische competenties die nodig zijn om effectief communicatief vreemdetalenonderwijs vorm te kunnen geven.

Bestaande gevalideerde observatie-instrumenten, die wel expliciet gericht zijn op vreemdetalendidactiek, zijn meestal ontwikkeld in het kader van vakdidactisch onderzoek met een specifieke focus binnen de communicatieve benadering, zoals het gebruik van de doeltaal of de kwaliteit van mondelinge interactie. Weinig vakdidactische observatie-instrumenten bieden de mogelijkheid om de kwaliteit van de communicatieve aanpak van een leraar als geheel te evalueren. Het vakdidactisch observatie-instrument COLT (Spada & Fröhlich, 1995) komt het dichtst in de buurt, maar is opgesteld als een checklist

van CTO-criteria gericht op het materiaal en de houding van leerlingen en leraren. Het instrument is echter niet geschikt om het vakdidactisch handelen van lio's te evalueren. Lerarenopleiders, stagebegeleiders en lio's missen dus een eenduidig, gevalideerd observatie-instrument waarin principes van CTO vertaald worden in concrete indicatoren van effectief vakdidactisch handelen.

Een dergelijk vakdidactisch instrument zou de opleiding tot mvt-leraar kunnen versterken, omdat het een gezamenlijk vakdidactisch referentiekader zou bieden voor opleiding en stagepraktijk, en een noodzakelijke aanvulling vormen op bestaande, algemeen pedagogisch-didactische observatie-instrumenten.

Het observatie-instrument

Tussen 2018 en 2020 heeft het lectoraat Didactiek van de Moderne Vreemde Talen van de Hogeschool Utrecht (HU) daarom een vakdidactisch observatie-instrument ontwikkeld, in aansluiting op de kernprincipes van CTO. De onderzoeksgroep bestond uit vier senioronderzoekers met expertise op het gebied van vreemdetalenonderwijs en zeven vakdidactici die ook stagebegeleiders en lerarenopleiders zijn: twee voor het vak Engels, twee voor het vak Frans, twee voor het vak Spaans en één voor het vak Duits. Het instrument is in verschillende stappen ontwikkeld en gevalideerd volgens de iteratieve procedure van ontwerpgericht onderzoek (McKenney & Reeves, 2018): een analyse- en exploratiefase (inhoudsvaliditeit), een ontwerp- en constructiefase (constructvaliditeit) en een evaluatiefase, waar dit artikel specifiek over rapporteert.

Op basis van veelgebruikte literatuur over communicatief vreemdetalenonderwijs en de vertaling daarvan in de praktijk (Brandl, 2008; Bimmel, 2008; Celce-Murcia,

2008; Danielson, 2015; Larsen-Freeman & Anderson, 2011; Lightbown & Spada, 2013; Long, 2009; Savignon, 2017; Tammenga-Helmantel & Mossing Holsteijn, 2016; Dönszelmann, 2019; De Jong et al., 2019; Council of Europe, 2001) is een overzicht gemaakt van observeerbare kernaspecten van CTO. Deze vormden een eerste basis voor de inhoud van het instrument.

Na een inventarisatie van bestaande instrumenten werd ervoor gekozen het format van het ICALT-instrument (RUG, 2019; Maulana, Helms-Lorenz & Van de Grift, 2017) als uitgangspunt te nemen voor de vormgeving, omdat dit instrument tegenwoordig door veel lerarenopleidingen in Nederland wordt ingezet. In dit format zijn gedragsindicatoren opgenomen die onderverdeeld zijn onder hoofdcategorieën. Indicatoren kunnen gescoord worden op een vierpunts Likert-schaal, variërend van 1 (overwegend zwak) tot 4 (overwegend sterk). Daarna zijn de hoofdcategorieën voor het te ontwikkelen instrument bepaald, gerelateerd aan de hbo-kennisbases vakdidactiek voor de vreemde talen: een categorie voor elke taalvaardigheid afzonderlijk, een categorie voor het gebruik van de doeltaal en een categorie voor de behandeling van vormaspecten, zoals gram-matica in context (*focus on form*). Voor elke categorie zijn vervolgens gedragsindicatoren geformuleerd die aangescherpt en aangevuld werden na verschillende expertsessies binnen de onderzoeksgroep.

Om de indicatoren in het instrument zo concreet en waarneembaar mogelijk te formuleren en om een duidelijk verband te leggen met de praktijk, zijn vervolgens in de onderzoeksgroep gezamenlijk video-opnames van mvt-lessen bekeken. Het betrof uiteindelijk drie video-opnames van gehele lessen Frans, Duits en Engels opgenomen in 2017. De bespreking van de observaties leidde tot het formuleren van concretere gedragsindicatoren waaraan vakdidactische

Categorie	Indicator: de leraar...	Score
Gespreksvaardigheid/ Spreekvaardigheid	... zet functionele en/of creatieve gespreks-/ spreekactiviteiten in die aansluiten bij de belevingswerelden/of het inlevingsvermogen van de leerlingen.	1 2 3 4
	... gebruikt gespreks-/spreektaken die aansluiten bij het (ERK) niveau van de doelgroep en die bij leerlingen een passende mate van vrije productie uitlokken	1 2 3 4
Schrijfvaardigheid	... zet functionele en/of creatieve schrijfactiviteiten in die aansluiten bij de belevingswereld en/of het inlevingsvermogen van de leerlingen.	1 2 3 4
	... gebruikt schrijftaken die aansluiten bij het (ERK-) niveau van de doelgroep en die bij leerlingen een passende mate van vrije productie uitlokken.	1 2 3 4
Leesvaardigheid	... zet authentieke teksten in die aansluiten bij de belevingswereld en/of het inlevingsvermogen van de leerlingen.	1 2 3 4
	... zet functionele, op tekstbegrip gerichte leesactiviteiten in die aansluiten op het (ERK-)niveau van de doelgroep.	1 2 3 4
Luistervaardigheid/ Kijkvaardigheid	... zet authentiek luister-/kijkmateriaal in dat aansluit bij de belevingswereld en/of het inlevingsvermogen van de leerlingen.	1 2 3 4
	... zet functionele, op begrip gerichte luister-/kijkactiviteiten in die aansluiten op het (ERK-)niveau van de doelgroep.	1 2 3 4
Doeltaalgebruik	... zet de doeltaal in als leermiddel en besteedt hiermee expliciet aandacht aan taalontwikkeling.	1 2 3 4
	... stimuleert dat leerlingen de doeltaal gebruiken in communicatie onderling.	1 2 3 4
Focus on form	... richt de aandacht van leerlingen op <i>geplande wijze</i> op vormaspecten van taal (bv. grammatica, woordenschat, uitspraak) binnen een betekenisvolle, communicatieve context (geplande focus on form).	1 2 3 4
	... richt de aandacht van leerlingen op <i>ongeplande wijze</i> op vormaspecten van taal (bv. grammatica, woordenschat, uitspraak) binnen een betekenisvolle, communicatieve context (ongeplande focus on form).	1 2 3 4
Geef hier een concrete beschrijving van de geobserveerde instructie en leeractiviteiten		

1 = overwegend zwak; 2 = meer zwak dan sterk; 3 = meer sterk dan zwak; 4 = overwegend sterk

Tabel 1. Format observatie-instrument met eerste zes hoofdcategorieën en voorbeelden van gedragsindicatoren

handelingen van mvt-leraren bij het aanbieden van communicatieve activiteiten moeten voldoen (zie tabel 1)

Uiteindelijk is een aantal voorbeelden van goede praktijken toegevoegd aan elke gedragsindicator om de gedragsindicatoren te concretiseren. In tabel 2 staan enkele van deze voorbeelden behorend bij de eerste twee gedragsindicatoren van de categorie *Doeltaalgebruik*.

De huidige studie

Na de ontwerp- en constructiefase vond de evaluatiestudie plaats die in dit artikel centraal staat. Het doel van deze evaluatiefase was te toetsen in hoeverre het ontwikkelde vakdidactische instrument valide, betrouwbaar en bruikbaar is. De onderzoeksvraag die leidend is in deze studie luidt:

Wat is de kwaliteit en bruikbaarheid van een observatie-instrument gericht op het vakdidactisch handelen van mvt-leraren volgens de principes van communicatief vreemdetalenonderwijs?

Om antwoord te kunnen geven op deze hoofdvraag zijn de volgende deelvragen geformuleerd:

1. Wat is de kwaliteit van het observatie-instrument in termen van validiteit en betrouwbaarheid?
2. Wat is de kwaliteit van het observatie-instrument in termen van praktische bruikbaarheid?

Methode

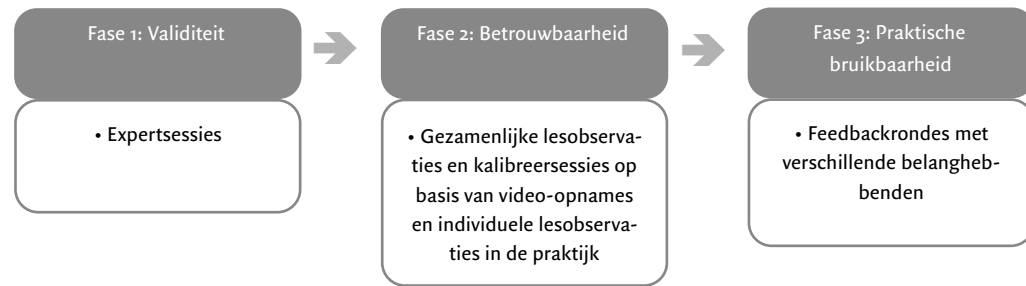
Figuur 1 op de volgende pagina illustreert het evaluatieproces dat in deze studie doorlopen is, gebaseerd op aanbevelingen van Nieveen en Folmer (2013). Het evaluatieproces bestond uit drie fases en in het proces waren verschillende stakeholders betrokken: externe en interne onderzoekers, vakdidactici, lerarenopleiders, stagebegeleiders en lio's.

Validiteit en betrouwbaarheid

In evaluatiefase 1 stond de validiteit van het instrument centraal. De conceptversie van het instrument, zoals tot stand gekomen in de ontwikkelfase, werd in deze fase aangevuld en verfijnd na overleg met diverse externe inhoudsdeskundigen die deel hebben genomen aan expertsessies van de onderzoeksgroep. In deze expertsessies zijn we in gesprek gegaan over de verschillende categorieën en gedragsindicatoren van het instrument. Input en feedbackpunten werden verzameld en gebruikt om de inhoudsvaliditeit van het instrument te verhogen.

... zet de doeltaal in als communicatiemiddel met de leerlingen	1 2 3 4	... maakt intensief gebruik van de doeltaal bij begroetingen, instructies ('pak je boek'), spontane praatjes met leerlingen, etc., zowel klassikaal als in één-op-één-interactie
... zet de doeltaal in als leermiddel en besteedt hiermee expliciet aandacht aan taalontwikkeling	1 2 3 4	... biedt doelgerichte ondersteuning om het (taal)leren van leerlingen te faciliteren (bv. met behulp van gebaren en/of selectief en bewust gebruik van de moedertaal)

Tabel 2. Voorbeelden van goede praktijk bij gedragsindicatoren uit de categorie *Doeltaalgebruik*



Figuur 1. Evaluatieproces van het ontwikkelde vakdidactisch observatie-instrument

In evaluatiefase 2 – gericht op betrouwbaarheid – hebben de leden van de onderzoeksgroep twee typen observaties en kalibreersessies uitgevoerd om zicht te creëren op de mate waarin het instrument verschillende beoordelaars in staat stelt om tot vergelijkbare conclusies te komen ten aanzien van het vakdidactisch handelen van mvt-leraren. Ten eerste werd een kalibreersessie uitgevoerd op basis van een video-opname van een gehele les Engels. Zeven leden van de onderzoeksgroep hebben deze les eerst individueel gescoord. Scores die uiteenliepen werden in de groep besproken. Om voor meer consistentie te zorgen, zijn vervolgens afspraken gemaakt en vastgelegd met betrekking tot het scoren van de indicatoren. Voor het vervolg van evaluatiefase 2 werden scholen en mvt-leraren benaderd via het netwerk van de lerarenopleiding van de Hogeschool Utrecht. Acht lesobservaties in zeven verschillende stagescholen zijn vervolgens gezamenlijk uitgevoerd door telkens twee verschillende leden van de onderzoeksgroep. Het betrof twee lessen Duits, een les Engels, drie lessen Frans en twee lessen Spaans. De geobserveerde leraren zijn gevraagd om les te geven zoals ze altijd deden. De twee observanten hebben eerst tijdens de les individueel het observatie-instrument ingevuld, en daarna nogmaals gezamenlijk na intercollegiaal overleg om

consensus te vinden bij scores die uiteenliepen. Om de interbeoordelaarsbetrouwbaarheid te bepalen vóór en na intercollegiaal overleg werd Fleiss' kappa² gebruikt.

Praktische bruikbaarheid

De versie van het instrument waarin verbeterpunten uit evaluatiefases 1 en 2 verwerkt zijn, is vervolgens voorgelegd aan verschillende praktijkdeskundigen (HU-intern en extern) tijdens feedbacksessies waarin de praktische bruikbaarheid van het instrument centraal stond. Stagebegeleiders en lerarenopleiders van de HU (n = 21), lio's mvt van de Universiteit Utrecht (n = 9) en van de Hogeschool Utrecht (n = 24), en lerarenopleiders en stagebegeleiders uit andere universitaire- en hbo-lerarenopleidingen (n = 37) hebben deelgenomen aan deze feedbackrondes. Tijdens de sessies werd het instrument eerst kort gepresenteerd en toegelicht, vervolgens werden fragmenten van een lesopname getoond en gescoord door de deelnemers en tot slot werden zowel scores als bruikbaarheid van het instrument besproken in groepen. Oordelen over de bruikbaarheid van het instrument en discussiepunten werden via evaluatieformulieren verzameld. De evaluatieformulieren waren gericht op instrumentele bruikbaarheid, en bevatten zowel open als gesloten vragen en ruimte voor opmerkingen en suggesties. Per

respondentengroep bevatten de formulieren verschillende richtvragen. Samen leverden zij informatie op met betrekking tot de volgende aspecten: algemene bruikbaarheid van het instrument; bruikbaarheid van het instrument ten behoeve van zelfevaluatie, (peer-) feedback en beoordeling; bruikbaarheid van het instrument ter ondersteuning van lesontwerp; formulering van gedragsindicatoren en voorbeelden; lay-out van het instrument. Kwantitatieve data uit de gesloten (schaal) vragen zijn beschrijvend geanalyseerd om een beeld te geven hoe de verschillende stakeholders de praktische bruikbaarheid van het instrument ervaren. Opmerkingen behorend bij de schaalvragen, antwoorden op open vragen en suggesties zijn geïnventariseerd om enerzijds de kwantitatieve resultaten te duiden en anderzijds de belangrijkste feedbackpunten te verzamelen.

Resultaten

Validiteit en betrouwbaarheid

Na consultatie van experts (evaluatiefase 1: validiteit) zijn aanvullende gedragsindicatoren toegevoegd om het vakdidactisch handelen

van lio's en leraren mvt beter in kaart te kunnen brengen. In gesprek met een aantal deskundigen die (vakdidactisch) onderzoek hebben gedaan op het gebied van taalonderwijs in de Nederlandse context, hebben wij er ten eerste voor gekozen om voor elke taalvaardigheid extra gedragsindicatoren toe te voegen die meer algemeen vakdidactisch en procesgericht het handelen van mvt-leraren beschrijven. De volgende vier criteria zijn uiteindelijk toegevoegd: verbinding van activiteiten met een communicatief leerdoel, scaffolding (stap-voor-stap begeleiding), ondersteunen bij de ontwikkeling van strategieën en het inzetten van formatieve evaluatie (zie tabel 3).

Bovendien werd op basis van de kijkkaders voor taalgericht vakonderwijs (Van Eerde et al., 2007) ook een categorie over de integratie van taalvaardigheden en bijbehorende gedragsindicatoren toegevoegd (zie tabel 4).

Na de eerste kalibreersessie op basis van de video-opname (evaluatiefase 2: betrouwbaarheid) zijn eerste algemene afspraken gemaakt met betrekking tot het consistent scoren van de gedragsindicatoren, bijvoorbeeld dat bij twijfel tussen twee scores de laagste score moet worden gekozen, volgens de richtlijnen van de ICALT (RUG, 2019).

... verbindt spreek-, schrijf-, lees- of luisteractiviteiten expliciet aan een communicatief leerdoel	1 2 3 4
... biedt leerlingen bij het (ERK-)niveau passende scaffolding om de spreek-, schrijf-, lees- of luisteraak uit te kunnen voeren	1 2 3 4
... stimuleert leerlingen strategieën in te zetten om een spreek-, schrijf-, lees- of luisteraak uit te kunnen voeren	1 2 3 4
... creëert door het inzetten van formatieve evaluatie voor zichzelf en voor leerlingen zicht op waar de leerling staat ten opzichte van het spreek-, schrijf-, lees- of luisterdoel	1 2 3 4

Tabel 3. Aanvullende procesgerichte gedragsindicatoren bij de vier taalvaardigheden

... expliciteert de samenhang tussen verschillende vaardigheden voor het bereiken van communicatieve leerdoelen	1	2	3	4
... laat verschillende vaardigheden op een geïntegreerde wijze aan bod komen in de les	1	2	3	4

Tabel 4. Gedragsindicatoren over mate van integratie van de taalvaardigheden

Bij de observaties in de praktijk (evaluatiefase 2: betrouwbaarheid) werd aangetoond dat er over het geheel voldoende overeenstemming was tussen de beoordelingen vóór intercollegiaal overleg: $\kappa = 0,602$ (95% CI, 0,501 tot 0,702), $p < 0,0001$. De interpretatie van de Fleiss' kappa waarde is gelijk aan die van Cohens kappa, waarin waarden tussen 0,40 en 0,75 kunnen worden beschouwd als een voldoende tot goede overeenstemming (Fleiss, Paik & Levin, 2003). Verschillen in scores zaten met name bij een aantal gedragsindicatoren die overwegend zwak gescoord waren door de ene beoordelaar en meer zwak dan sterk door de andere. Het betrof in de meeste gevallen gedragsindicatoren uit de categorieën van de receptieve vaardigheden, lees- en luistervaardigheid. Tijdens intercollegiaal overleg werden argumenten uitgewisseld en werd steeds bij twijfel gekozen voor de laagste score in overeenstemming met de gemaakte afspraken tijdens de gezamenlijke kalibratiesessie. Na intercollegiaal overleg was er uiteindelijk voor elke observatie volledige overeenstemming. Om interpretatieverschillen te beperken, is de definitie van de scores in de laatste versie van het instrument explicieter gemaakt. De score 1, overwegend zwak is vervangen door niet geobserveerd, de situatie vroeg er wel om, de score 2 meer zwak dan sterk door in geringe mate geobserveerd, de score 3 meer sterk dan zwak door in voldoende mate geobserveerd en de score 4 overwegend sterk door in hoge mate geobserveerd.

Praktische bruikbaarheid
Tabel 5 laat zien hoe de praktische bruikbaarheid van het instrument werd beoordeeld door stagebegeleiders en lerarenopleiders van de HU (n = 21).

De resultaten uit tabel 5 laten zien dat het instrument over het algemeen door de meerderheid van lerarenopleiders en stagebegeleiders van de HU als praktisch (49%) tot heel praktisch (16%) wordt beoordeeld. Zowel de lay-out, de categorieën met gedragsindicatoren, als het gebruik van het instrument voor beoordeling werden door ruim meer dan de helft van de respondenten praktisch tot heel praktisch bevonden.

Tabel 6 geeft weer hoe bruikbaar lio's van de HU en de UU (n=23) het instrument vinden als ondersteuning bij lesontwerp en voor zelfevaluatie en peer-feedback. De tabel toont dat bijna alle lio's (92%) het instrument praktisch (57%) tot heel praktisch (35%) bruikbaar vinden als ondersteuning bij lesontwerp. Ook voor zelfevaluatie en peer-feedback wordt het instrument gewaardeerd door een aanzienlijke meerderheid van de bevroegde lio's.

In tabel 7 en 8 op bladzijde 12 zijn de belangrijkste feedbackpunten van de verschillende stakeholders opgenomen: stagebegeleiders en lerarenopleiders enerzijds (tabel 7) en lio's anderzijds (tabel 8).

Uit de feedback in tabel 7 bleek dat een deel van de lerarenopleiders en stagebegeleiders behoefte heeft aan verduidelijking met betrekking tot de focus van het observatie-instrument. Ze benadrukken het feit dat het

	1	2	3	4	5
Algemene bruikbaarheid	0%	6%	29%	49%	16%
Lay-out	3%	7%	29%	29%	32%
Categorieën en gedragsindicatoren	0%	0%	13%	68%	19%
Praktische toepassing voor beoordeling	0%	0%	10%	51%	39%

1=Helemaal niet praktisch; 2=Niet praktisch; 3=Neutraal; 4=Praktisch; 5=Heel praktisch

Tabel 5. Waardering van lerarenopleiders en stagebegeleiders van de HU t.a.v. de praktische bruikbaarheid van het instrument

	1	2	3	4	5
Ondersteuning bij lesontwerp	0%	4%	4%	57%	35%
Praktische toepassing voor zelfevaluatie	0%	4%	9%	48%	39%
Praktische toepassing voor peer-feedback	4%	0%	13%	48%	35%

1=Helemaal niet praktisch; 2=Niet praktisch; 3=Neutraal; 4=Praktisch; 5=Heel praktisch

Tabel 6. Waardering van lio's t.a.v. de praktische bruikbaarheid van het instrument

instrument vooral focust op de communicatieve aspecten van het vreemdetalenonderwijs en daardoor geen compleet beeld geeft van het beoogde vakdidactisch handelen van lio's en leraren mvt. Daarnaast wordt het belang benadrukt van een overzichtelijke lay-out en duidelijke formulering van de gedragsindicatoren waarin gebruikt jargon geëxpliciteerd moet worden voor lio's, maar ook voor collega's uit andere vakgebieden. Een goede introductie en training worden aanbevolen. Lio's hebben aangegeven het instrument bruikbaar te vinden en hebben geen inhoudelijke verbeterpunten genoemd (tabel 8). Aan de hand van de resultaten is de lay-out van het instrument aangepast en overzichtelijker gemaakt, een aantal gedragsindicatoren is aangescherpt en beknopter geformuleerd om verschillende interpretaties te vermijden.

Extra concrete voorbeelden zijn toegevoegd om de indicatoren te illustreren en de bruikbaarheid te verhogen. Daarnaast is er een begrippenlijst toegevoegd om de specifieke vakdidactische terminologie toe te lichten.³

Conclusies en discussie

In deze evaluatiestudie is onderzocht wat de kwaliteit en bruikbaarheid is van een observatie-instrument dat het handelen van lio's en leraren mvt volgens de principes van communicatief vreemdetalenonderwijs zichtbaar en meetbaar maakt.

Discussies met inhoudsdeskundigen tijdens de expertsessies hebben gezorgd voor een hogere inhoudsvaliditeit van het instrument. Belangrijke handelingen die nodig zijn

	FEEDBACKPUNTEN	CITATEN
Algemene bruikbaarheid	Reikwijdte instrument	Het instrument zet in op vaardigheden, maar als vakdidactisch instrument zouden ook andere onderdelen (bijv. interculturele competentie) opgenomen moeten worden.
Lay-out en omvang	Behoeftte aan overzichtelijkheid	De omvang van het instrument maakt overzicht krijgen/houden soms lastig. De lay-out zou nog overzichtelijker/beter behapbaar kunnen door er aparte tabellen van te maken.
Gedragsindicatoren (formulering en relevantie)	Aanscherping in formulering nodig	Huidige formulering en indeling nog niet goed bruikbaar. Sommige termen zijn onduidelijk. Eventueel betere uitleg bij de termen (met duidelijke voorbeelden).
	Behoeftte aan begrippenlijst	Begrippen als ERK zijn voor niet-talendocenten onhanteerbaar.
Praktische toepassing voor beoordeling	Gebruik als aanvullend instrument bij beoordeling	Dit is een heel goed instrument! Echt een aanvulling voor het observeren van mvt-lessen naast het ICALT-formulier.
	Belang van training	Duidelijk is dat aan de observatie een training vooraf noodzakelijk is. Een getraind oog ziet meer.
	Meer geschikt als feedback-instrument	Relevant om bewustwording op gang te brengen. Dus in dat opzicht prima. Minder geschikt om daadwerkelijk te scoren.

Tabel 7. Feedback van lerarenopleiders en stagebegeleiders op praktische bruikbaarheid van het instrument

Ondersteuning bij lesontwerp	Het geeft goede richtlijnen waar je les aan kan voldoen, vooral gezien de voorbeelden die exemplarisch aangeven wat voor aspecten bij het ontwerp van communicatief talenonderwijs komen kijken. Het helpt me echt verder in het ontwikkelen van mijn lessen!
Praktische toepassing voor zelfevaluatie	De aspecten waaraan ik minder aandacht besteed, komen zo naar voren en motiveren mij om mij daar mee bezig te houden.
Praktische toepassing voor feedback	Genoeg zinvolle aspecten die de vaardigheden voldoende lijken af te dekken en aan de hand waarvan concrete feedback kan worden geformuleerd.

Tabel 8. Feedback van lio's op de praktische bruikbaarheid voor lesontwerp, zelfevaluatie en feedback

om CTO goed vorm te kunnen geven, zijn opgenomen in aanvullende indicatoren. Het betreft de koppeling van activiteiten aan een communicatief leerdoel, scaffolding, ontwikkeling van communicatieve strategieën en formatief handelen. Daarnaast is het belang van het geïntegreerd aanbieden van taalvaardigheden duidelijk gemaakt door de hoofdcategorie *Mate van integratie* met bijbehorende descriptor toe te voegen. Authentieke communicatieve activiteiten moeten leerlingen namelijk aanzetten tot het geïntegreerd gebruiken van verschillende taalvaardigheden, zoals het geval is bij communicatie in het echte leven: een e-mail begrijpen én beantwoorden (lees- en schrijfvaardigheid) of een discussie voeren naar aanleiding van een film (kijk-/luister- en gespreksvaardigheid) (Brown, 2005).

De Fleiss Kappa waarde, waarmee de mate van overeenstemming tussen beoordelaars is gemeten, bleek vóór intercollegiaal overleg voldoende, maar toonde volledige overeenstemming bij iedere lesobservatie na intercollegiaal overleg. Dit benadrukt het belang van een goede training en duidelijke kaders alvorens het instrument in te zetten voor beoordeling. Bredere implementatie, training en evaluatie in de praktijk zijn nodig om de interbeoordelaarsbetrouwbaarheid vóór intercollegiaal overleg te verhogen.

Uit de evaluaties om de praktische bruikbaarheid van het ontwikkelde instrument te toetsen, bleek dat de meerderheid van de verschillende stakeholders (lerarenopleiders, stagebegeleiders en lio's) het instrument een bruikbaar hulpmiddel vindt. Lio's geven aan veel te leren over de mvt-lespraktijk door inzet van het instrument bij lesontwerp, eigen observaties, observaties van derden, peer-feedback en (zelf)reflectie. Uit de feedback van lerarenopleiders en stagebegeleiders bleek wel het belang van een verduidelijking van het doel van het observatie-instrument: wordt het gebruikt als feed-

backinstrument of als aanvullend vakdidactisch instrument bij beoordeling? De helft van de lerarenopleiders en stagebegeleiders vindt het instrument praktisch bruikbaar als aanvulling op een pedagogisch-didactisch observatie-instrument zoals het ICALT.

Het instrument is voornamelijk alleen bedoeld om het handelen volgens de principes van CTO inzichtelijk te maken. Andere belangrijke pijlers van vreemdetalenonderwijs zoals aandacht voor interculturaliteit, literatuur of taalbewustzijn zijn (nog) niet opgenomen. Om aan te sluiten bij de realiteit van de veranderende samenleving en schoolpraktijk blijft het instrument in ontwikkeling. Zo zijn er inmiddels voorbeelden toegevoegd van manieren waarop leraren de meertaligheid van leerlingen kunnen benutten bij het leren communiceren in een nieuwe vreemde taal.

Inzetbaarheid in de praktijk

Het observatie-instrument wordt momenteel in de vakdidactische leerlijn van de eerste- en tweedegraads mvt-lerarenopleidingen van de HU ingezet, om (peer-)feedback te geven op lesontwerpen en lessen van (mede) studenten. Om stagebegeleiders en niet-vakspecifieke lerarenopleiders te trainen in het gebruik worden workshoprondes aangeboden op de HU en op stagescholen. Het instrument wordt inmiddels ook gedeeld met andere eerste- en tweedegraads lerarenopleidingen voor het begeleiden van lio's. Het kan ook gebruikt worden in de beroepspraktijk door individuele leraren en secties moderne vreemde talen, bijvoorbeeld voor lesvoorbereiding, intercollegiale observatie of intervisie-activiteiten. Uiteindelijk hopen we met het instrument een bijdrage te leveren aan de totstandkoming van een sterke en duurzame verbinding tussen theorie en praktijk in de lerarenopleiding en de beroepspraktijk.

NOTEN

1. Onder het woord 'leerlingen' verstaan wij taalleerders in het voortgezet onderwijs en (voorbereidend) beroepsonderwijs.
2. Fleiss' kappa meet de interbeoordelaarsbetrouwbaarheid van verschillende beoordelaars, in tegenstelling tot andere kappa's zoals Cohens kappa, die alleen de overeenstemming tussen twee vaste beoordelaars meten of van één beoordelaar tegenover zichzelf.
3. Zie begrippenlijst en laatste versie van het observatie-instrument op <https://www.hu.nl/onderzoek/projecten/onderzoek-naar-effectief-communicatief-taalonderwijs>

LITERATUUR

- iovoordeleraar. (2017). Kennisbases. Geraadpleegd op 15 februari 2023 van <https://iovoordeleraar.nl/kennisbases/publicaties>
- Bimmel, P., Canton, J., Fasoglio, D., & Rijlaarsdam, G. (Red.). (2008). *Handboek Ontwerpen Talen*. Vossiuspers UvA.
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36(2), 81–109.
- Brandl, K. (2008). *Communicative language teaching in action: Putting principles to work*. Pearson Prentice Hall.
- Brown, J. D. (2005). *Testing in language programs*. McGraw-Hill College.
- Celce-Murcia, M. (2008). Rethinking the role of communicative competence in language teaching. In E. Alcón Soler & M. P. Safont Jordá (Eds.), *Intercultural language use and language learning* (pp. 41–57). Springer.
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge University Press.
- Danielson, C. (2015). Framing discussions about teaching. *Educational Leadership*, 72(7), 38–41.
- Dönszelmann, S. (2019). *Doeltaal-Leertaal: Didactiek, professionalisering en leereffecten*. Proefschrift Vrije Universiteit Amsterdam. Uitgeverij Parthenon.
- Eerde, D. van, Hacquebord, H., Hajer, M., Pulles, M., & Raymakers, C. (2007). *Kijkwijzer voor taalgericht vakonderwijs* (brochure met dvd). SLO/Platform Taalgericht Vakonderwijs.
- Farrell, T. S., & Bennis, K. (2013). Reflecting on ESL teacher beliefs and classroom practices: A case study. *RELC Journal*, 44(2), 163–176.
- Fleiss, J. L., Paik, M. C., & Levin, B., (2003). *Statistical methods for rates and proportions*. 3rd ed. John Wiley & Sons Inc.
- Graaff, R. de. (2018). *Van Vreemde Talen naar Vereende Talen: de plek van Vreemde Talen in het onderwijs*. Hogeschool Utrecht.
- Graus, J. (2018). *Student teacher cognitions on form-focused instruction: An explanatory sequential study of Dutch undergraduate and post-graduate EFL student teachers*. Proefschrift Radboud Universiteit.
- Kelly, L. B. (2018). Preservice teachers' developing conceptions of teaching English learners. *Tesol Quarterly*, 52(1), 110–136.
- Jong, N. de, Vrind, E. de, & Beuningen, C. van. (2019). *Moderne vreemde talen*. In F. Janssen, H. Hulshof, & K. van Veen (Eds.), *Wat is echt de moeite waard om te onderwijzen? Een perspectiefgerichte benadering* (pp. 81–105). Universiteit Leiden / Rijksuniversiteit Groningen.
- Larsen-Freeman, D., & Anderson, M. (2011). *Techniques and principles in language teaching*. Oxford University Press.
- Lightbown, P. M., & Spada, N. (2013). *How languages are learned* (4th edition). Oxford University Press.
- Long, M. (2009). Methodological principles for language teaching. In M. Long, & C. Doughty (Eds.), *The handbook of language teaching* (pp. 373–394). Wiley-Blackwell.
- Maulana, R., Helms-Lorenz, M., & Grift, W. van de. (2017). *Validating a model of*

- effective teaching behaviour of pre-service teachers*. *Teachers and Teaching*, 23(4), 471–493.
- McKenney, S., & Reeves, T. C. (2018). *Conducting educational design research*. Routledge.
- Nieveen, N., & Folmer, E. (2013). Formative evaluation in educational design research. In T. Plomp & N. Nieveen (eds), *Educational Design Research* (pp. 152–169). SLO: Nationaal expertisecentrum leerplanontwikkeling.
- Ogilvie, G., & Dunn, W. (2010). Taking teacher education to task: Exploring the role of teacher education in promoting the utilization of task-based language teaching. *Language Teaching Research*, 14(2), 161–181.
- RUG. (2019). *ICALT3: International Comparative Analysis of Learning and Teaching*. Rijksuniversiteit Groningen. Geraadpleegd op 18 mei 2022 van <https://www.rug.nl/gmw/lerarenopleiding/onderzoek/psychometrisch/>
- Savignon, S. J. (2017). *Communicative competence*. In J. I. Lontos (Ed.), *The TESOL Encyclopedia of English Language Teaching*, (pp.1–7). John Wiley.
- Spada, N., & Fröhlich, M. (1995). *COLT-communicative orientation of language teaching observation scheme: Coding conventions and applications*. National Centre for English Language Teaching and Research.
- Tang, S. Y., Wong, A. K., & Cheng, & M. M. (2012). Professional learning in initial teacher education: Vision in the constructivist conception of teaching and learning. *Journal of Education for Teaching*, 38(4), 435–451.
- Tammenga-Helmantel, M., & Mossing Holsteijn, L. (2016). *Doeltaal in het moderne vreemde talen onderwijs: Gebruik en opvattingen van leraren in opleiding*. Rijksuniversiteit Groningen & Meesterschapsteam Vakdidactiek MVT.
- West, L., & Verspoor, M. (2016). An impression of foreign language teaching approaches in the Netherlands. *Levende Talen Tijdschrift* 17(4), 26–36.
- CHARLINE ROUFFET is hogeschooldocent Frans en lerarenopleider Moderne Vreemde Talen aan de Hogeschool Utrecht voor de eerste- en tweedegraads lerarenopleidingen. Ze werkt daarnaast voor het kenniscentrum Leren en Innoveren als docent-onderzoeker binnen het lectoraat Meertaligheid en Onderwijs. E-mail: charline.rouffet@hu.nl
- IRENE DE KLEYN is lerarenopleider en vakdidactica Moderne Vreemde Talen aan de Hogeschool Utrecht met ervaring in het voortgezet onderwijs. Irene is curriculumverantwoordelijke van de Bachelor lerarenopleidingen Engels en Frans. Zij heeft een leidende rol in het herontwerp van de leerlijn vakdidactiek moderne vreemde talen. E-mail: irene.dekleyn@hu.nl
- MARLEEN IJZERMAN is hogeschooldocent Frans en lerarenopleider Moderne Vreemde Talen voor de eerste- en tweedegraads lerarenopleiding aan de Hogeschool Utrecht met ruime ervaring in het voortgezet onderwijs. Zij houdt zich bezig met het herontwerp van de leerlijn vakdidactiek moderne vreemde talen. E-mail: marleen.ijzerman@hu.nl
- CATHERINE VAN BEUNINGEN is hoofddocent Talenonderwijs en Meertaligheid binnen de lerarenopleidingen van de Hogeschool van Amsterdam. Daarnaast leidt zij een programmalijn rondom kansrijke omgang met taaldiversiteit in onderwijs en opvoeding in het Centre of Expertise Urban Education. E-mail: c.g.van.beuningen@hva.nl
- RICK DE GRAAFF is hoogleraar Talen-didactiek aan de Universiteit Utrecht, en lector Meertaligheid en Onderwijs aan de Hogeschool Utrecht. Hij doet onderzoek met talentdocenten en vakdidactici in de onderwijspraktijk, om het leren en onderwijzen van talen betekenisvoller te maken. E-mail: rick.degraaff@hu.nl