

Microteaching met leerwinst als criterium¹⁾

H. J. KOOREMAN,

Centrum voor Didaktiek en Onderzoek van Onderwijs, Technische Hogeschool Twente.

J. B. WASSINK,

Christelijke Pedagogische Akademie, Almelo

Samenvatting

Een onderzoek werd gedaan naar condities waaronder één microteaching-cyclus leidt tot effectiever onderwijzen (uitgedrukt in leerwinstscores). Het verschaffen van zeer specifieke doelstellingen – 14 meerkeuzevragen, de helft van de natoetsvragen – had een zeer significant effect ($p < 0,001$). Geen verschil werd gevonden tussen een docent en student als supervisor. Leerlingenfeedback alleen leidde niet tot vooruitgang. Als de supervisors – gedurende 7 weken – ervaring hadden opgedaan met het gebruik van leerlingenfeedback en enkele leer- en onderwijsprincipes (structurering, 'knowledge of results' en bordgebruik), werd tijdens de volgende 7 weken niet alleen een winst gevonden op de toets die als leerlingenfeedback werd gebruikt ($p < 0,005$), maar ook een winst op de andere helft van de natoets (generalisatie-effekt). In één geval was dit effect net niet significant ($p = 0,053$), in het andere geval zeer significant ($p < 0,001$).

Inleiding

In een voorgaand artikel (Kooreman 1971a) werd betoogd dat als uiteindelijk criterium voor goed of slecht onderwijzen beschouwd moet worden: de leerwinst bij de leerlingen in de richting van de onderwijsdoelstellingen. De effectiviteit van microteaching – een methode die gericht is op het verbeteren van het onderwijzen – is dan ook het meest direct na te gaan door de leerwinst te meten die studenten tijdens een microteaching-cyclus weten te bewerkstelligen bij hun leerlingen. Dat deze door Gage (Gage, 1963, 1968 en Gage & Unruh, 1967) gepropageerde benaderingswijze

– die de doelstelling primair stelt en daarna de vraag stelt naar middelen en methoden en niet omgekeerd – nog geen gemeengoed is, moge blijken uit de bijdragen over microteaching op de International Conference of Educational Technology in Glasgow (Britton & Leith, 1971 en McAleese & Unwin, 1971) en de discussie die daarop volgde. De centrale vraag was volgens de inleiders: 'Wat zijn de criteria voor het al of niet goed uitvoeren van 'teaching skills'?'. Met deze criteria bedoelde men leerkrachtenvariabelen. De vraag naar het leerkrachtengedrag is echter een afgeleide vraag van de werkelijke kernvraag: 'Onder welke condities worden onderwijsdoelstellingen – beoogde leerwinst – optimaal bereikt?' (Glaser, 1962, van Gelder, 1968, De Cecco, 1968).

Het stellen van de vraag naar de doelstellingen voordat methoden en middelen aan de orde komen, is de grote winst van de afgelopen jaren (Bever, 1968). Door uit te gaan van de vaardigheden die een leerkracht 'moet' (waarom?) beheersen en zich niet af te vragen wat met die vaardigheden beoogd wordt, wordt deze vooruitgang weer teruggedraaid. De onderzoekingen naar 'teachereffectiveness' bv. hebben uitgewezen dat er geen effectief leerkrachtengedrag 'an sich' bestaat. Wat effectief is voor de ene doelstelling – en dit soort leerlingen – hoeft dit niet te zijn voor een andere doelstelling.

Probleem

Wij willen nagaan of microteaching een geschikt middel is om a.s. onderwijzers te helpen effectiever onderwijs te geven. In een vroegere publikatie (Kooreman 1971a) werden een drietal richt-

lijnen uit de literatuur naar voren gebracht die zouden kunnen leiden tot een effectief gebruik van microteaching:

1. stel de beoogde leerwinst vast;
2. koppel de theorie aan de praktijk;
3. laat de leerlingen feedback geven.

Onze onderzoeksvraag luidde nu: 'Onder welke condities leidt het gebruik van deze drie richtlijnen tot effectiever onderwijzen?'

We zullen nu nagaan welke gegevens over verschillende condities bekend zijn en op grond daarvan aangeven welke condities wij willen onderzoeken.

1. Rosenshine (1970) vergeleek de onderzoeken die de stabiliteit nagingen van de resultaten van onderwijsgedrag. Hiertoe werden de korrelaties berekend tussen leerwinstscores die een leerkracht behaalde onder de volgende drie condities:

1. dezelfde lesinhoud wordt aan andere (vergelijkbare) leerlingen gegeven;
2. een andere lesinhoud wordt aan dezelfde leerlingen gegeven;
3. een andere lesinhoud wordt aan andere (vergelijkbare) leerlingen gegeven.

Hij konkludeerde dat alleen in geval 1 de stabiliteit van het leerkrachtgedrag redelijk was; de gemiddelde korrelatie tussen leerwinstscores over zes onderzoeken was 0,52. In de twee ander gevallen werden zeer wisselende (meestal niet significante) korrelaties gevonden. Slechts wanneer dezelfde lesinhoud aan andere leerlingen wordt gegeven lijkt een eerste onderzoek naar het effect van een microteaching-cyclus zinvol; immers eventuele winst kan anders even goed worden toegeschreven aan instabiliteit van het leerkrachtgedrag (tenzij een zeer hoge significantiedrempel van bijvoorbeeld $p < 0,001$ wordt aangenomen). Wij kozen daarom voor de eerste mogelijkheid.

2. In een aantal onderzoeken werd de effectiviteit van feedback door een supervisor in in twijfel getrokken (Acheson, 1964, aangehaald in Borg, Kallenbach, Morris & Frie-

bel, 1969, Tuckman & Oliver, 1968). Vanuit de praktijksituatie vindt deze twijfel steun bij Stewig (1970). Leerlingenfeedback blijkt effectiever te zijn. Positieve veranderingen in leerkrachtgedrag werden gevonden als leerlingen de aan hen gegeven les beoordeelden (Gage, Runkel en Chatterjee, 1960 en 1963, Bryan, 1963, Tuckman & Oliver, 1968).

Al deze onderzoeken gingen echter niet uit van leerwinst als criterium, maar van bepaalde gedragingen van de leerkracht (bv. de verhouding tussen 'student-' en 'teacher-talk').

Deze criteria hoeven niet veel te zeggen over het al of niet bereiken van de onderwijsdoelstellingen. De meest direkte methode om na te gaan in hoeverre beoogde leerwinst is gerealiseerd, is niet de beoordeling van het onderwijs door de leerlingen, maar feedback over de leerprestaties van de leerlingen. Dit kan de student (leerkracht) zelf vaststellen, of het kan hem worden meegedeeld. De laatste methode werd toegepast in een onderzoek van Gropper en Lumsdaine (1961).

Een significant hogere leerwinst werd bereikt wanneer de gegevens van een na afloop afgenomen toets werden gebruikt om een tv-leraar feedback te geven op grond waarvan hij (in samenwerking met anderen) zijn onderwijs reviseerde.

Op grond van het bovenstaande wilden wij het effect onderzoeken van feedback over de bereikte leerprestaties. Vanwege de (veronderstelde) motiverende werking lieten wij de student zelf zijn onderwijsresultaat vaststellen door (de helft van) de natoets met de leerlingen na te kijken.

3. Ondanks de tot nu toe niet zo positieve resultaten over supervisie, wilden wij nagaan of zij een eigen inbreng heeft bij de eventuele verbetering van het onderwijsresultaat. Wij wilden echter twee soorten supervisie onderzoeken. Supervisie door een hoofdakte-student en supervisie door een docent. Om na te gaan of er mogelijkheden zijn hoofdakte-

studenten – na een trainingsfase – als supervisor in te schakelen waardoor micro-teaching mankracht zou kunnen besparen, omdat supervisie op de pedagogische akademies dan voor een deel door hen zou kunnen geschieden. Dit is dan tevens een goede oefening voor de in-service-training die zij later – als hoofd – aan hun aankomend personeel moeten geven.

4. Mager en McCann (1962) constateerden dat a.s. ingenieurs onderwijsdoelstellingen sneller bereikten wanneer deze expliciet werden meegedeeld. Later onderzoek heeft dit resultaat bevestigd. Zo haalden studenten in de medicijnen, aan wie nauwkeurig omschreven doelstellingen werden verstrekt, significant hogere tentamenscores dan studenten aan wie geen of vage doelstellingen werden verstrekt (Dalis, 1970).

Toegespitst op het geven van onderwijs vonden Wittrock (1962) en McNeil (1967) dat studenten aan wie verteld werd dat zij beoordeeld zouden worden op grond van de leerwinst die zij bij hun leerlingen wisten te realiseren een significant hogere leerwinst behaalden dan studenten tegen wie dit niet was gezegd of tegen wie was gezegd dat hun praktijkcijfer zou afhangen van het oordeel van een docent. Op grond van deze onderzoeken zou men verwachten dat het verschaffen van natoetsvragen, ter voorbereiding van het geven van een les, zou leiden tot beter onderwijsresultaat. Het probleem is nu echter waaraan een eventueel beter onderwijsresultaat moet worden toegeschreven. Aan het vooraf verschaffen van de lesinhoud? Aan het er weet van hebben dat na afloop meerkeuzevragen gesteld zullen worden? Of aan het vooraf verschaffen van de meerkeuzevragen? Tot nog toe werden geen significante verschillen in onderwijsresultaat gevonden tussen al of niet vooraf verschaft meerkeuzevragen. Indien vijf van de tien meerkeuzevragen – die na afloop aan de leerlingen werden gesteld – vooraf aan leerkrachten werden uitgereikt, werden deze vragen door de leer-

lingen niet beter gemaakt dan de vijf niet uitgereikte vragen (Fortune, Gage & Shutes, 1966 en Belgard, Rosenshine & Gage, 1968, beide aangehaald in Rosenshine, 1970). Deze onderzoeken lijken erop te wijzen dat het verschaffen van meerkeuzevragen weinig toevoegt aan het effect van het verschaffen van de lesinhoud (een vrij korte tekst) en het weten dat meerkeuzevragen na afloop zullen worden gevraagd. Toch wilden we deze factor controleren. Een eventueel effect van het nabespreken van de antwoorden op de helft van de natoetsvragen – leerlingenfeedback – zou indien de vragen niet bekend waren bij de voorbereiding van de eerste les kunnen worden toegeschreven aan het bekend zijn met de toetsvragen. De helft van de studenten kreeg daarom de toetsvragen die gebruikt werden bij de leerlingenfeedback, de andere helft niet.

Hypothesen

Op grond van het voorgaande stellen wij de volgende hypothesen – één hoofdhypothese en drie deelhypothesen.

1. Het gebruik tijdens de nabespreking van een eerste microteaching-les van de drie richtlijnen – stel de beoogde leerwinst vast – koppel de theorie aan de praktijk – laat de leerlingen feedback geven – leidt tot een significant ($p < 0,05$) hogere leerwinst bij het geven van de tweede microteaching-les vergeleken met de eerste keer.
1. Er is geen verschil tussen supervisie door een docent of door een (ouderejaars) student.
2. Er is geen verschil tussen het al of niet vooraf verschaffen van de helft van de natoetsvragen aan studenten die onderwijs geven.
3. Er is bij de tweede keer lesgeven een significant ($p < 0,05$) hogere leerwinst op de vragen waarover door de leerlingen feedback is gegeven vergeleken met de andere natoetsvragen.

Tabel 1 Overzicht van het aantal proefpersonen in het experiment naar soort supervisie en wel of niet vooraf verschaft van de helft van de natoets.

deelnemers	supervisie docent		supervisie student		Totaal
	BC vooraf verschaft	BC niet vooraf verschaft	BC vooraf verschaft	BC niet vooraf verschaft	
studenten	7	6	7	6	26
leerlingen	84	72	82 ¹	72	310

1. Eén leerling van de groep aan wie de 2e keer werd lesgegeven was ziek. Deze leerling en de overeenkomende leerling uit de groep van de 1e keer lesgeven werden niet in de analyse betrokken.

Procedure

de deelnemers

Bij ons eigenlijke onderzoek waren betrokken 26 eerstejaarsstudenten van de Christelijke Pedagogische Akademie te Almelo en 310 leerlingen van de 5e en 6e klassen van basisscholen uit Almelo en omgeving (zie tabel 1). Bij het proefdraaien werd bovendien nog op 2 eerstejaarsstudenten en 24 basisschoolleerlingen een beroep gedaan. Als supervisoren traden op een docent en een hoofdakte-student van de akademie.

de inhoud van de les

Aan studenten werd een door ons gemaakte tekst van ongeveer 700 woorden verstrekt, handelend over het droogleggen van de meren in Noord-Holland. Deze tekst was bedoeld als standaardinhoud voor de door de studenten te geven les. De inhoud betrof de periode van ± 1200 tot de drooglegging van de Haarlemmermeer in 1852. De tekst droeg de naam 'Leegwater' naar de centrale figuur in het verhaal.

de konstruktie van de toets

Over de bovengenoemde tekst werden 35 meerkeuzevragen geschreven die in een 5e klas werden uitgeprobeerd en op grond daarvan gerevisieerd. Vijf vragen bleken op grond van de negatieve itemrestkorrelatie of ongunstige moeilijkheidsgraad ongeschikt. De 30 overgebleven - veelal

herschreven - items werden weer afgenomen in een andere 5e klas van het basisonderwijs. Twee items bleken nog een ongunstige moeilijkheidsgraad te hebben en werden verwijderd. De homogeniteit van de 30 items was volgens de Kuder-Richardson-20 formule 0,54, niet hoog, maar voor ons doel - waar homogeniteit van de items geen vereiste is - niet ongeschikt. De 28 overgebleven items werden verdeeld over 4 groepen (A, B, C en D). De items werden geplaatst in volgorde van hun moeilijkheidsgraad. De moeilijkste vraag kwam in groep A, de daarna komende vraag in B, etc. Bij de tweede verdeling begonnen we met groep D tot A en bij de derde verdeling weer bij groep A, totdat alle vragen over de groepen verdeeld waren. Zo ontstonden vier groepen vragen die qua gemiddelde moeilijkheidsgraad vrijwel niet verschilden. De vragen lagen voor ongeveer de helft op kennisniveau en voor de andere helft op inzicht (comprehension) niveau (Bloom, 1956). Behalve feitelijke gebeurtenissen werden in de tekst voorkomende begrippen gevraagd (oktrooi, ringvaart, garantie, droogmakerij), enige vragen over de geografische ligging en een enkele extrapolatie naar het heden. Een voorbeeld van een vraag naar inzicht (comprehension) in een begrip is:

Prins Maurits verleende Leegwater oktrooi voor zijn uitvinding, d.w.z.

1. Maurits mocht de uitvinding nu gebruiken;

2. de uitvinding werd nu gebruikt bij de oorlogsvoering;
3. niemand anders kon nu meer zeggen dat hij de uitvinding had gedaan;
4. niemand anders mag de uitvinding nu ooit gebruiken.

Een voorbeeld van een vraag naar een feitelijke gebeurtenis op kennisniveau is:

Waardoor ontstonden de meeste meren in Noord-Holland?

1. doordat dijken doorbraken;
2. doordat rivieren zich insneden in het land;
3. doordat er veen werd afgegraven;
4. doordat de storm het zeewater hoog opjog.

de samenstelling van de leerlingen

Wij wilden ongeveer gelijkwaardige leerlingengroepen samenstellen om les aan te geven. Omdat wij cognitieve leerwinst maten wilden wij de verbale prestatiefactor onder controle houden. Hiertoe werd gebruik gemaakt van een Leidse stillestoets (1970) die door Kooreman was uitgebreid en met als criterium de itemrest-korrelaties – gevonden in vorig onderzoek – op 26 items was gebracht. Een voordeel is dat de toetsingprocedure (meerkeuzevragen) overeenkomt met de door ons gebruikte criteriumtoetsen. Deze toets werd enkele weken voor de microteachingles, waaraan de desbetreffende leerlingen deelnamen, op de basisscholen afgenomen. De K.R.-20 homogeniteit liep van 0,73 tot 0,81 voor vijf verschillende scholen, wat ruim voldoende is om de leerlingen in drie groepen in te delen. De schoolbevolking van een 5e en 6e klas werd verdeeld in drie ongeveer even grote, hoog, middelmatigen laag op de stillestoets scorende groepen. Indien leerlingen moesten afvallen omdat per school steeds 12-tallen nodig waren – aan 6 leerlingen werd de eerste keer lesgegeven en aan 6 andere de tweede keer – werden steeds de extreem goede of slechte weggelaten om de kans op gelijkwaardige groepen zo groot mogelijk te houden. 2 lage, 2 hoge en 2 middelmatige scoorders werden nu telkens at random gekozen om een leerlingengroep te vormen.

de instructies aan de studenten

Aan twee klassen met eerstejaarsstudenten van de Pedagogische Akademie werd door de pedagogiekleraar verteld dat er een onderzoek zou plaatsvinden naar de effectiviteit van microteaching; een methode waarmee zij reeds bekend waren. De opzet van een microteaching cyclus (zie hieronder) werd uiteengezet en hen werd verteld dat zij enkele dagen voordat zij aan de beurt waren een opdracht zouden krijgen om les te geven. Allen werden dringend verzocht niet met elkaar te praten over hun les omdat dit het onderzoek zou verstoren.

Alle studenten kregen enkele dagen voor de microteachingles de volgende instructie: 'Deze tekst over Leeghwater is de lesinhoud. Je hebt precies 13 minuten om les te geven. Je moet zelf uitmaken wat je belangrijk en wat je minder belangrijk vindt. Ook de wijze waarop je de les overbrengt wordt aan je eigen vindingrijkheid overgelaten. Hoeveel de leerlingen ervan opsteken wordt nagegaan door ze zowel voor als na je les een aantal meerkeuzevragen te laten maken. Je geeft les aan 6 leerlingen, na een nabespreking van je les kun je jezelf verbeteren door aan een vergelijkbare groep leerlingen les te geven'.

De – at random gekozen – helft van de studenten kreeg bovendien de volgende extra instructie: 'Bovendien heb je hier de helft van de meerkeuzevragen die na afloop gesteld worden aan de leerlingen om te zien wat ze van je les hebben opgestoken'.

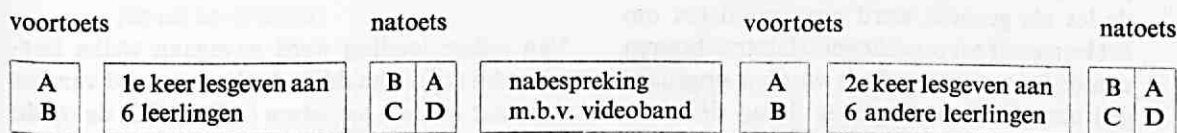
de microteaching-cyclus

Om leerwinst te kunnen vaststellen moet nagegaan worden wat het verschil is tussen de eindsituatie van een leerling – na het lesgeven – en zijn beginsituatie. De beginsituatie werd gemeten m.b.v. een voortoets (14 items), terwijl de eindsituatie werd gemeten met twee natoetsen (ook ieder 14 items).

Door middel van een at random-procedure bleek AB de voortoets te zijn en BC de toets die nabesproken werd en daarmee tevens wel of niet vooraf werd verschaft. AD was de andere natoets.

Met de BC-toets maten wij de invloed van het

Figuur 1 Schema van een microteaching-cyclus



vooraf verschaffen van deze toets en/of de invloed van leerlingenfeedback over deze toets, met de AD-toets het generalisatie-effekt.

Chronologisch is de sequentie binnen een microteaching-cyclus als volgt weer te geven (zie ook figuur 1).

1. 6 leerlingen maken de voortoets AB (± 8 minuten).
2. de student geeft de eerste keer les aan de in 1 genoemde leerlingen (13 minuten exakt).
3. de leerlingen maken de natoetsen AD en BC (± 20 minuten).
4. de student kijkt samen met de leerlingen de BC-toets na (± 5 minuten).
5. de supervisor en student bespreken en bekijken de gegeven les (± 20 minuten).
6. terwijl de student zijn tweede keer lesgeven voorbereidt, maken 6 nieuwe leerlingen de voortoets AB (± 8 minuten).
7. de student geeft de tweede keer les aan de in 6 genoemde leerlingen (13 minuten exakt).
8. de leerlingen maken de natoetsvragen AD en BC (± 20 minuten).

Een leerlingengroep werd willekeurig aan een eerste of tweede keer lesgeven toegewezen. De voortoets werd meestal gemaakt in hetzelfde lokaal waar de leerlingen les kregen. De les werd met één stilstaande kamera – zonder aanwezigheid van anderen – op een videorecorder opgenomen. Voor het op- en instellen van kamera en mikrofoon en het bedienen van de videorecorder was de amanuensis van de academie aanwezig.

Het afnemen van de natoetsen vond meestal plaats in een ander lokaal omdat een nieuwe student – bij de andere supervisor – zijn les moest geven in het opnamelokaal. De organisatie van het geheel – leerlingen halen – toetsen afne-

men – bijhouden wie aan wie les had gegeven – was in handen van een hoofdakte-student. Door deze maatregelen konden de supervisors hun volledige aandacht aan hun eigen taak besteden.

de inhoud van de nabespreking

Gedurende een ochtend werd met de supervisors proefgedraaid. Aan twee studenten werd supervisie gegeven waarbij ook de inhoud van de nabespreking werd doorgepraat. In principe bestond de nabespreking uit de volgende punten:

1. Nagaan welke BC-vragen foutief beantwoord zijn en de student vragen hoe dit verbeterd kan worden. In het gesprek tussen de supervisor en student moet gezamenlijk gezocht worden naar wegen die wel tot leerwinst zouden kunnen leiden. De supervisor zou niet wijzen op tekortkomingen in houding en gedrag van de student, maar uitgaan van de 'harde' gegevens over het resultaat van het onderwijzen. Wij wilden immers het effect nagaan van leerlingenfeedback en niet het effect van intuïtieve 'feedback' over het gedrag van de student (zie uitvoeriger in 'Konklusies').
2. Als de supervisor verbeteringen voorstelde, zou hij gebruik maken van twee leerprincipes: 'structurering' en 'knowledge of results'. Het eerste principe is een combinatie van de beide principes 'progressieve differentiatie' en 'integratieve verzoening' van Ausubel (1968) waarover Kooreman (1971b) onlangs schreef. Het tweede principe: het verschaffen van 'knowledge of results' aan de leerlingen kan geschieden door het stellen van controlevragen en het geven van feedback op de antwoorden die de leerlingen geven. Beide principes zouden niet in abstracto uitgelegd worden aan de student; de supervisor zou de

principes toepassen bij de nabespreking van de les als gezocht werd naar middelen om het leerresultaat te verhogen. Het structureren van de les zou grotendeels worden opgehangen aan de chronologische draad die door het verhaal van Leeghwater liep.

3. De supervisor zou voorts wijzen op het gebruik van bord, kaart en andere illustraties; hetgeen wij beschouwen als een onderwijsprincipe.

bijzondere omstandigheden en opmerkingen

Aan het einde van twee microteaching-cycli ontdekten we op het bord in de 6e klas een getekende ringvaart. Het bleek dat het hoofd der school – overigens geheel te goeder trouw – zijn leerlingen alvast iets meer over het onderwerp Leeghwater, waarmee ze allen gekonfronteerd zouden worden, had verteld. Gelukkig bleek hierbij geen van de door ons reeds gerekruteerde leerlingen vóór zijn bijdrage aan de microteaching aanwezig te zijn geweest. De overige ± 40 leerlingen (inklusief de 5e klas) waren, ondanks dat wij hen de stillestoets hadden afgenomen en zorgvuldig gerandomiseerd, vanzelfsprekend niet meer bruikbaar. Het was overigens de eerste school waardoor verder onheil door tijdige voorlichting kon worden voorkomen.

Voor video-opnamen kon gebruik worden gemaakt van apparatuur van de TH Twente terwijl SVO zo welwillend was videobanden – in bruikleen – ter beschikking te stellen.

Een groot gedeelte van de apparatuur werd in de periode van het onderzoek op de Pedagogische Akademie in Almelo gestolen, waardoor het onderzoek ternauwernood kon doorgaan.

Wij hebben bij zowel de leerlingen als bij de studenten gevraagd of ze niet met anderen over de les wilden praten. Wij hebben geen aanwijzingen dat het wel is gebeurd, maar als dit toch zo mocht zijn, dan controleerden wij deze faktor omdat wij *leerwinst* bij de leerlingen en *onderwijswinst* bij de studenten maten. In beide gevallen werd eindsituatie minus beginsituatie als criterium genomen.

Resultaten

Van iedere leerling werd nagegaan welke leerwinst hij (zij) behaald had. Hiertoe werd van het resultaat op de 2 natoetsen (AD en BC) de op de voortoets behaalde score (AB) afgetrokken. Een eventueel voortoetseffect (vragen die voorafgaande aan de experimentele procedure zijn gesteld, worden op de natoets beter gemaakt dan vragen die niet zijn gesteld), o.a. gekonstateerd door Berlyne (1966) en Peeck (1970) heeft op deze wijze gelijk invloed op de beide natoetsen.

De homogeniteit van de voortoets en de twee natoetsen werd berekend m.b.v. een toetsanalyse-komputerprogramma. Bij de eerste keer lesgeven ($n = 155$) leverde de Kuder-Richardson-20 formule de volgende koëfficiënten op: voortoets AB 0,14, natoets AD 0,12 en natoets BC 0,47. Voor de tweede keer lesgeven (eveneens $n = 155$) waren de koëfficiënten 0,14 voor de voortoets AB, 0,26 voor natoets AD en 0,49 voor natoets BC. Omdat geen enkele van de 84 berekende itemrest-korrelaties significant negatief was, zijn deze – lage – koëfficiënten niet toe te schrijven aan slechte items. De items waren bovendien al vooraf uitgetoetst. Er zijn twee andere duidelijk oorzaken: ten eerste beoogden wij geen homogene toets en bestaat de toets uit vragen die verschillende zaken meten (zie 'Procedure'), ten tweede werd in ons geval de homogeniteit niet berekend als iedere leerling een standaardprocedure had ondergaan (bv. dezelfde tekst lezen over Leeghwater) maar als verschillende a.s. onderwijzers op verschillende wijze een les hadden gegeven over Leeghwater. Dat de BC-toets zowel bij de eerste als tweede keer een redelijke homogeniteit heeft (0,47 en 0,49) moet dan ook verklaard worden uit het feit dat de BC-toets vooraf aan de helft van de studenten was verschaft en op deze wijze meer kans was op een lesinhoud die meer gelijk was voor de verschillende leerlingen.

De gemiddelden en standaarddeviaties van de leerwinstscores werden vervolgens berekend en staan aangegeven in tabel 2 volgens een 2^4 factoriële opzet. De vier onderzochte factoren zijn gerelateerd aan de door ons gestelde hypothesen.

Tabel 2 Overzicht van de 2⁴ faktoriële opzet. Gemiddelde en standaarddeviatie van de leerwinstscores staan per cel aangegeven.

		faktor I							
		1e keer lesgeven				2e keer lesgeven			
		faktor 1				faktor 1			
		supervisie docent		supervisie student		supervisie docent		supervisie student	
		faktor 2		faktor 2		faktor 2		faktor 2	
		BC voor- af ver- schaft	BC niet vooraf verschaft	BC voor- af ver- schaft	BC niet vooraf verschaft	BC voor- af ver- schaft	BC niet vooraf verschaft	BC voor- af ver- schaft	BC niet vooraf verschaft
		n = 42	n = 36	n = 41	n = 36	n = 42	n = 36	n = 41	n = 36
faktor 3	niet als feedbackmiddel gebruikt bij trainingsfase (AD)	m = 2,02 s = 2,11	m = 1,50 s = 1,83	m = 2,09 s = 1,84	m = 1,77 s = 1,66	m = 2,02 s = 1,94	m = 2,00 s = 2,00	m = 2,39 s = 2,39	m = 1,27 s = 2,22
	als feedbackmiddel gebruikt bij trainingsfase (BC)	m = 2,69 s = 2,16	m = 1,25 s = 1,58	m = 3,19 s = 2,51	m = 1,33 s = 1,96	m = 3,30 s = 2,17	m = 3,47 s = 2,33	m = 3,63 s = 2,16	m = 2,63 s = 2,51

faktor I: het effect van de gehele trainingsfase.

faktor 1: het verschil in effect tussen supervisie door een docent en door een student.

faktor 2: het verschil in effect tussen het al of niet vooraf verschaffen van de helft van de natoets.

faktor 3: het verschil in effect tussen het al of niet gebruiken van de helft van de natoets als feedbackmiddel bij de trainingsfase.

Een eerste indruk van de resultaten levert tabel 3 op. Verschillen in leerwinstscores tussen de tweede en eerste keer lesgeven staan aangegeven voor de factoren 1, 2 en 3. Er zijn twee zeer significante verschillen op de t-toets die duidelijk

overeenkomen. Als de BC-natoets niet vooraf is verschaft, wordt – afgezien van de soort supervisie – op de BC-toets een zeer significante vooruitgang geboekt. Omdat dit niet het geval is als BC wel vooraf is verschaft, lijkt hypothese 2 verworpen te moeten worden. Er is wel degelijk een verschil tussen het al of niet verschaffen van de natoets, zoals ook uit het vervolg zal blijken.

Een uitgebreider overzicht van de invloed van de verschillende factoren – en hun interacties – krijgt men m.b.v. een variantieanalyse over de 2⁴ faktoriële opzet (zie tabel 2). Hiertoe werd vanwege de ongelijke aantallen per cel en de eenvoudigheid van de berekening gekozen voor de 'unweighted means analysis' (Dayton 1970, blz.

Tabel 3 Verschil in leerwinstscores tussen de 2e en 1e keer lesgeven naar twee soorten supervisie, wel of niet vooraf verschaffen van de helft van de natoets (BC) op twee soorten toetsen.

Toets	supervisie docent		supervisie student	
	BC vooraf verschaft	BC niet vooraf verschaft	BC vooraf verschaft	BC niet vooraf verschaft
niet als feedbackmiddel gebruikt bij trainingsfase (AD)	0,00	0,50	0,30	-0,50
als feedbackmiddel gebruikt bij trainingsfase (BC)	0,61	2,22 ⁴	0,44	1,30 ³

Signifikante verschillen op de t-toets zijn als volgt weergegeven:

³ = $p < 0,005$

⁴ = $p < 0,001$

117 e.v.). Variantie-analyse is echter pas toegestaan als homogeniteit van de variantie is gewaarborgd. De toets van Bartlett (Dayton, 1970, blz. 32 e.v.) leverde een chi-kwadraat op bij 15 vrijheids graden van 16,7 wat op een homogene verdeling van de varianties wijst. De resultaten van de variantie-analyse staan weergegeven in tabel 4. We zullen nu achtereenvolgens de vier door ons opgestelde hypothesen nagaan.

Hypothese 1 is te toetsen door de invloed van faktor I na te gaan. Deze faktor blijkt een zeer significante invloed te hebben ($p < 0,001$) waardoor hypothese I bevestigd moet worden. De gehele trainingsfase leidt tot een significant hogere leerwinst bij de tweede keer lesgeven. Over de kondities die deze vooruitgang in het bijzonder bewerkstelligen gaan de hypothesen 1, 2 en 3.

Hypothese 1 is te toetsen door na te gaan of er een interactie is tussen faktor I en faktor 1. De bijdrage van deze interactie is niet significant. De twee soorten supervisie blijken dus niet te verschillen in hun effect op de tweede keer lesgeven waarmee hypothese 1 wordt bevestigd. Ook zijn er geen significante interacties tussen faktor 1 en de andere factoren, hoewel er wel een tendens is dat de docent meer bereikt bij studenten die de BC-toets niet vooraf hebben gekregen en de

student-supervisor het beter doet bij studenten aan wie de BC-toets wel vooraf is verschaft (interactie 1, 2).

Hypothese 2 is te toetsen door de bijdrage van faktor 2 en de interacties tussen de factoren I en 2 en de factoren I, 2 en 3 na te gaan. Zoals al bij de analyse met de t-toets bleek moet deze hypothese verworpen worden. Faktor 2 heeft een zeer significante invloed ($p < 0,001$) wat erop wijst dat het vooraf verschaffen van de helft van de natoets tot een zeer grote leerwinstverhoging leidt. De interactie tussen faktor I en 2 is niet significant, waaruit gekonkludeerd kan worden dat er geen verschil in effect te constateren is tussen het verschaffen van de BC-toets voor het geven van de eerste les en het verschaffen van de toets tijdens de trainingsfase. Beide leiden tot een grote - niet significant verschillende - vooruitgang. Tenslotte blijkt uit de interactie tussen de factoren I, 2 en 3 dat de grootste leerwinst gemaakt wordt bij de tweede keer lesgeven op de BC-toets wanneer deze toets niet vooraf is verschaft ($p < 0,05$).

Hypothese 3 is niet te toetsen met behulp van de variantie-analyse. Weliswaar is er een zeer significante invloed van faktor 3 ($p < 0,001$) en zijn er significante interacties tussen I, 2 en 3, maar dit mag niet alleen worden toegeschreven aan het

Tabel 4 Resultaat van de variantie-analyse op grond van een ongewogen gemiddelde analyse van de 24 faktoriële opzet (zie tabel 2).

Bron	kwadratensom	vrijheidsgraden	gemid. kwadratensom	F-ratio	p-waarde
faktor I	56,69	1	56,69	11,96	< 0,001
faktor 1	0,006	1	0,006	0,001	n.s.
faktor 2	89,26	1	89,26	18,83	< 0,001
faktor 3	98,83	1	98,83	20,85	< 0,001
interaktie I 1	7,27	1	7,27	1,54	n.s.
interaktie I 2	10,72	1	10,72	2,26	n.s.
interaktie I 3	43,29	1	43,29	9,13	< 0,005
interaktie 1 2	14,94	1	14,94	3,15	n.s.
interaktie 1 3	0,33	1	0,33	0,07	n.s.
interaktie 2 3	10,72	1	10,72	2,26	n.s.
interaktie I 1 2	10,34	1	10,34	2,18	n.s.
interaktie I 1 3	2,68	1	2,68	0,57	n.s.
interaktie I 2 3	18,38	1	18,38	3,88	< 0,05
interaktie 1 2 3	1,46	1	1,46	0,31	n.s.
interaktie I 1 2 3	0,06	1	0,06	0,01	n.s.
foutenbron	2864,77	604	4,74		

gebruik van BC als feedbackmiddel bij de nabespreking. Het – zeer significante – effect van het vooraf verschaffen van de BC-toets is hier ook in verdisconteerd. De enige mogelijkheid om het

effect van het gebruik van BC als feedbackmiddel te toetsen is daarom na te gaan of, als de BC-toets vooraf is verschaft, er toch een grotere leerwinst wordt bereikt bij de tweede keer lesgeven dan bij

de eerste keer. Uit tabel 3 blijkt dat dit niet het geval is. De verschillen in leerwinstscores tussen BC en AD zijn niet significant. Hypothese 3 moet daarom worden verworpen.

Op grond van de mededelingen van de docent en de student dat zij steeds meer gewend raakten aan deze nieuwe methode van supervisie – waarbij niet uitgegaan moest worden van het gedrag van de leerkracht voor de klas, maar van de resultaten die de leerlingen bereikten – werd nagegaan of in het tweede gedeelte van het onderzoek leerlingenfeedback in combinatie met supervisie wel een positief effect had. Omdat er geen verschil in supervisie kon worden aangetoond, werd afgezien van het maken van onderscheid tussen supervisie door docent en door student. De resultaten staan weergegeven in tabel 5. Als BC vooraf verschaft is, blijkt in het tweede gedeelte van het onderzoek de combinatie leerlingenfeedback met supervisie een zeer significant effect te hebben ($p < 0,005$). Bovendien blijkt er een sterke tendens te bestaan dat ook toets AD hiervan meeprofiteert. Het gevonden verschil levert een t-waarde op van 1,64, terwijl de significantiedrempel voor $p = 0,05$ op 1,67 ligt. Deze tendens dat de AD-toets beter wordt gemaakt ondanks het feit dat er geen feedback over gegeven is, wordt nog versterkt door de zeer significante vooruitgang op de AD-toets wanneer BC niet vooraf is

verschaft ($p < 0,001$). Dat BC in dit laatste geval de grootste vooruitgang boekt, moet dan ook worden toegeschreven aan het gekombineerde effect: vooraf verschaffen van de toets en leerlingenfeedback met supervisie.

Hypothese 3 moest verworpen worden. Leerlingenfeedback 'an sich' bleek geen significante invloed te hebben. Wel bleek het gebruik van de leerlingenfeedback gegevens door twee supervisors, die ervaring hadden opgedaan met deze vorm van supervisie, tot significante leerwinstvermeerdering te leiden. En dit niet alleen op de toets die gebruikt werd als leerlingenfeedback, maar ook op de andere natoets. In het ene geval was dit verschil echter (net) niet significant.

Konklusies

Microteaching met leerwinst als criterium kan tot een verbetering van het onderwijzen leiden. We zullen de door ons gevonden condities die tot effectiever onderwijzen hebben geleid achtereenvolgens nagaan.

1. Het verschaffen van meerkeuzevragen die het onderwijsresultaat meten van de student, was de conditie die tot de grootste leerwinst leid-

Tabel 5 Verschil in leerwinstscores in het tweede gedeelte van het onderzoek tussen de tweede en eerste keer lesgeven naar wel of niet vooraf verschaffen vna de helft van de natoets (BC) op twee soorten toetsten.

Toets	BC vooraf verschaft (n = 35)	BC niet vooraf verschaft (n = 36)
niet als feedbackmiddel gebruikt bij trainingsfase (AD)	0,82*	1,06 ⁴
als feedbackmiddel gebruikt bij trainingsfase (BC)	1,32 ³	2,39 ⁴

Signifikante verschillen op de t-toets (in het voordeel van de tweede keer lesgeven) zijn als volgt weergegeven:

³ = $p < 0,005$

⁴ = $p < 0,001$

* Voor dit verschil werd een t-waarde van 1,64 gevonden; het significantieniveau voor $p = 0,05$ ligt bij $t = 1,67$.

de. Deze bevinding komt overeen met het in een vorig artikel (Kooreman 1971a) gekonstateerde, dat 'de beoogde leerwinst (moet) worden vastgesteld en wel zó concreet dat na afloop gemeten kan worden of de beoogde leerwinst ook werkelijk is gerealiseerd'. Het onderzoek van Fortune, Gage en Shutes (1966) en Belgard, Rosenshine en Gage (1968) – beide aangehaald door Rosenshine (1970) – dat wij nog niet kenden bij ons voorgaand artikel, deden ons echter vermoeden dat niet het vooraf verschaffen van de meerkeuzevragen maar het verschaffen van de lesinhoud en het weten dat de les geëvalueerd wordt m.b.v. meerkeuzevragen de belangrijkste variabelen waren. Zij vonden nl. geen verschil tussen de 5 wel en de 5 niet vooraf verschaft meerkeuzevragen. Wij vonden echter dat wanneer 14 van de 28 meerkeuzevragen vooraf worden verschaft, dit tot een significant hogere leerwinst op de verschaft vragen leidt dan wanneer geen meerkeuzevragen worden verstrekt. Dat wij wel en het door Rosenshine aangehaalde onderzoek geen verschil vond kan worden toegeschreven aan de verschillende onderzoekcondities.

Wij vertelden de betrokken studenten dat de aan hen verstrekte vragen na afloop aan de leerlingen zouden worden gesteld; de instructies aan de studenten van de door Rosenshine genoemde onderzoeken zijn ons niet bekend. Het kan zijn dat in deze onderzoeken vragen niet als een vorm van geoperationaliseerde, beoogde leerwinst zijn aangeboden, maar bv. als voorbeeldvragen. Wij vermoeden dat de diskrepantie is toe te schrijven aan verschillende instructies aan de studenten maar sluiten de invloed van andere condities – waarvan wij er twee noemen – zeker niet uit. Zij vergeleken niet, zoals wij, het wel of niet vooraf verschaffen van meerkeuzevragen aan leerkrachten, maar het verschil tussen wel en niet vooraf verschaft meerkeuzevragen wanneer aan alle leerkrachten 5 meerkeuzevragen als voorbeeld van criteriumitems werden verstrekt. Boven-

dien gebruikten wij bijna driemaal zoveel meerkeuzevragen. Onze resultaten komen wel overeen met onderzoek dat een gunstig effect lieten zien voor het expliciet meedelen van specifieke – in tegenstelling tot vage – doelstellingen (Mager & McCann, 1962 en Dalis, 1970).

2. Wij vonden geen significante verschillen tussen supervisie door een docent en door een student. De generalisatiewaarde bij twee supervisors is natuurlijk beperkt, maar het zou er op kunnen wijzen dat – getrainde – hoofdstudenten als supervisors dienst kunnen doen. Op deze wijze zouden de mankrachtproblemen die microteaching oproept opgelost kunnen worden. Bovendien is dit voor de hoofdstudenten een goede oefening voor de in-servicetraining die zij later – als hoofd – zullen moeten geven.

Men kan ons resultaat echter ook interpreteren als een aanwijzing dat pedagogiekleraren bijgeschoold moeten worden omdat zij er niet in slagen effectiever supervisie te geven dan hoofdstudenten.

3. Leerlingenfeedback alleen bleek geen significante invloed te hebben. Als echter de supervisors gedurende een zevental weken ervaring hadden opgedaan met deze voor hen nieuwe methode van supervisie dan bleek de combinatie van leerlingenfeedback en supervisie tot veel betere onderwijsresultaten te leiden. Niet alleen was er een direct effect aantoonbaar op de toets die gebruikt werd als leerlingenfeedback (BC), maar ook bleek er een generalisatie-effect naar toets AD. Dit laatste effect was zeer duidelijk als BC wel vooraf was verschaft. Op grond van deze resultaten lijkt het zeer belangrijk de supervisors te trainen. Leerlingenfeedback blijkt pas effectief gebruikt te worden als de supervisor ervaring heeft opgedaan met deze nieuwe vorm van supervisie. Tijdens de trainingfase waarin twee keer supervisie werd gegeven – konstateerden wij dat zowel docent als student er moeite mee hadden *niet over de*

houding van de student-onderwijzer voor de klas te spreken, bv. 'je staat gebogen, te dicht bij de kinderen, je doet zo zenuwachtig'. De bedoeling van de supervisie was een zakelijk gesprek over de manier waarop de leerlingen de lesinhoud beter zouden kunnen begrijpen en niet een gesprek over de houding van de student-onderwijzer. De student moet zich richten op zijn leerlingen en de lesinhoud en niet teruggeworpen worden op zijn eigen houding voor de klas. Naar de mening van de docent en de student gelukte de – door ons bedoelde – supervisie steeds beter. Wij interpreteren daarom het betere onderwijsresultaat in de tweede helft van het onderzoek als het gevolg van een meer centraal stellen van de leerlingen en de leerinhoud. Als niet over de houding (e.d.) van de leerkracht werd gesproken kon een zakelijk gesprek tot aanwijzingen voor verbetering van het onderwijzen leiden, waarbij de supervisor beschikte over een aantal relevante leer- en onderwijsprincipes (structurering, controlevragen, bordgebruik), waardoor theorie en praktijk konden worden gekoppeld.

Training van de supervisors zal vooral moeten plaatsvinden omdat momenteel nog sterk de nadruk wordt gelegd op de houding van de leerkracht en niet op het bereiken van onderwijsdoelstellingen. Ook zal de training moeten bestaan uit het leren hanteren van leer- en onderwijsprincipes, waarbij de eerste geleverd zullen moeten worden door een theorie (of theorieën) van het leren en de laatste door een theorie (of theorieën) van het onderwijzen. Voor theorieën van het leren verwijzen wij naar Van Parreren (1969/1970) en Ausubel (1968). Over de theorie van de laatste en de relevantie ervan voor het onderwijzen is onlangs geschreven door Kooreman (1971b). Voor theorieën van het onderwijzen verwijzen wij naar Gage (1964), Bruner (1967) en Van Gelder (1968).

Tenslotte wijzen wij erop dat de bovengenoemde resultaten zijn bewerkstelligd na één trainingsfase. Het meeste onderzoek bevat

meerdere trainingsfasen en vindt dan minder of geen vooruitgang (bv. Kieviet, 1971).

Noot

1. Dank is verschuldigd aan prof. dr. C. F. van Parreren en drs. G. M. G. M. van Lieshout voor hun suggesties ter verbetering van vroegere concepten en aan E. J. Gouman, W. Post en A. J. Smelt voor hun bijdrage aan de uitvoering van het experiment.

Literatuur

- Ausubel, D. P.,
Educational psychology: a cognitive view,
Holt, Rinehart & Winston, inc., New York 1968
- Berlyne, D. E.,
Conditions of prequestioning and retention of meaningful material,
Journal of educational psychology,
1966, 57 (3), blz. 128-132
- Bevers, J. A. A. M.,
Formuleren van onderwijsdoelstellingen,
I.O.W.O., Nijmegen 1968
- Bloom, B. S. (ed.),
Taxonomy of educational objectives: Handbook I cognitive domain,
David McKay Company, inc., New York 1956
- Borg, W. R., W. Kallenbach, M. Morris & A. Friebel,
Videotape feedback and microteaching in a teacher training model,
The journal of experimental education,
1969, 37 (4), blz. 9-16
- Britton, R. J. & G. O. M. Leith,
Some uses of mathematics and microteaching in teacher training, in: het *Konferentieboek van de International Conference of Educational Technology*, Glasgow 1971
- Bruner, J. S.,
Toward a theory of instruction,
Belknap press, Cambridge 1967
- Bryan, R. C.,
Reactions to teachers by students, parents and administrators,
Western Michigan University, Kalamazoo 1963
- Dalis, G. T.,
Effect of precise objectives upon student achievement in health education,

- The journal of experimental education*,
1970, 39 (2), blz. 20-23
- Dayton, C. M.,
The design of educational experiments,
McGraw Hill, inc., New York 1970
- De Cecco, J. P.,
The psychology of learning and instruction: educational psychology,
Prentice Hall, Englewood Cliffs 1968
- Gage, N. L., P. J. Runkel & B. B. Chatterjee,
Equilibrium theory and behavioral change: an experiment in feedback from pupils to teachers,
University of Illinois 1960
- Gage, N. L., P. J. Runkel & B. B. Chatterjee,
Changing teacher behavior through feedback from pupils: an application of equilibrium theory,
in: Charter, W. W. & N. L. Gage (eds.), *Readings in the social psychology of education*, Boston 1963
- Gage, N. L.,
Theories of teaching,
in: Hilgard, E. R. (ed), *Theories of learning and instruction*, University of Chicago Press, Chicago 1964
- Gage, N. L. & W. R. Unruh,
Theoretical formulations for research on teaching,
Review of educational research,
1967, 37, blz. 358-370
- Gage, N. L.,
An analytical approach to research on instructional methods, in: *Research and development toward the improvement of education*, Palo Alto 1968
- Gelder, L. van,
Kind, school, samenleving,
Wolters-Noordhoff N.V., Groningen 1968
- Glaser, R.,
Psychology and instructional technology, in;
Glaser, R. (ed.), *Training research and education*,
John Wiley & Sons, inc., New York 1962
- Gropper, G. L. & A. A. Lumsdaine,
The use of student response to improve televised instruction: an overview,
Pittsburg 1961
- Kieviet, F. K.,
Onderzoek van een nieuwe methode in de opleiding van leerkrachten: microteaching,
Pedagogische Studiën,
1971, 48 (5), blz. 205-212
- Kooreman, H. J.,
Naar een effectief gebruik van microteaching,
Pedagogische Studiën,
1971a, 48 (2), blz. 89-94
- Kooreman, H. J.,
De theorie van Ausubel: een leertheorie voor onderwijskundigen,
Pedagogische Studiën, 1971b, 48(9) blz. 402-411
- Mager, R. F. & J. McCann,
Learner-controlled instruction,
Palo Alto 1962
- McAleese, W. & D. Unwin,
A simulated teaching environment as part of a teacher training program,
in: het *Konferentieboek van de International Conference of Educational Technology*, Glasgow 1971
- McNeil, J. D.,
Concomittants of using behavioral objectives in the assessment of teacher effectiveness,
The journal of experimental education,
1967, 36, blz. 69-74
- Parreren, C. F. van,
Psychologie van het leren,
Van Loghum Slaterus, Arnhem, deel I, 1969, deel II, 1970
- Peeck, J.,
Effect of prequestions on delayed retention of prose material,
Journal of educational psychology,
1970, 61 (3), blz. 241-246
- Rosenshine, B.,
The stability of teacher effects upon student achievement,
Review of educational research,
1970, 40 (5), blz. 647-662
- Stewig, J. W.,
What should college supervisors do?
Journal of Teacher Education,
1970, 21 (2), blz. 251-257
- Tuckmann, B. W. & W. F. Oliver,
Effectiveness of feedback to teachers as a function of source,
Journal of educational psychology,
1968, 59 (4), blz. 297-301
- Wittrock, M. C.,
Set applied to student teaching,
Journal of educational psychology,
1962, 53 (4), blz. 175-180