

O N D E R W I J S

HET MAKEN VAN EEN TANDHEELKUNDIG BEHANDELPLAN

11. EVALUATIE VAN DE CURSUS

H. W. KERSTEN
R. W. R. STEURES
T.H. J. M. TROMP

*Uit de vakgroep Cariologie, Endodontologie en Pedodontologie van de Universiteit van Amsterdam.
Uit het Centrum voor Onderzoek van het Wetenschappelijk Onderwijs van de Universiteit van Amsterdam.*

Trefwoorden: Onderwijs – Behandelplan – Probleemoplossen

1. Inleiding

In een vorig artikel (Kersten e.a., 1980) werd verslag gedaan van een cursus die sinds 1975 aan derdejaarsstudenten wordt gegeven en waarin het leren maken van het behandelplan centraal staat. De cursus werd jaarlijks bijgesteld aan de hand van de opgedane ervaringen. In dit artikel komt de evaluatie van de cursus ter sprake aan de hand van de resultaten van de opeenvolgende versies tot en met het cursusjaar 1978-1979.

2. De cursus 1975-1976

2.1. Inrichting

De cursus vond plaats gedurende vier mid-dagen in november 1975. Alle studenten namen er op hetzelfde moment aan deel. De cursus bestond uit het opstellen van een behandelplan voor zeven geëxtraheerde elementen. Tussentijds werd de studenten geen informatie over de door hen gemaakte oplossingen verstrekt. De evaluatie van de cursus werd toegespitst op de volgende vragen:

1. Heeft 'instructie' effect?
2. Heeft 'oefening' effect?

Daartoe werd de groep deelnemende stu-denten aselekt verdeeld in twee groepen:

- een Syllabus + groep;
- een Syllabus - groep.

De S⁺-groep kreeg het volledig instructie-materiaal, syllabus en invulformulier, in

één keer aangeboden, met de opdracht dit ter plaatse te bestuderen. Tijdens de twee-de zitting kreeg deze groep een drietal ele-menten aangeboden om daarvoor een be-handelplan op te stellen, tijdens de derde en vierde zitting twee elementen. We be-schouwen dit als de 'gebruikelijke' volgor-de.

In de S⁻-groep begonnen de studenten met het oplossen van drie problemen (P) waar-bij ze de invulformulieren en alleen de toe-lichting daarop (hoofdstuk 3 uit de sylla-bus) tot hun beschikking hadden.

Tijdens de tweede zitting moesten zij de resterende hoofdstukken uit de syllabus bestuderen. De derde en vierde zitting wa-ren gelijk aan die van de S⁺-groep. We beschouwen dit als de 'onbruikelijke' volgorde. De afbeeldingen 1 en 2 geven de volgorde waarin het instructiemateriaal werd aangeboden en de interactie tussen instructie en oefening, schematisch weer. De hypothesen, geformuleerd in termen van het gemiddelde aantal goede antwoor-den per groep, waren respectievelijk:

1. voor de variabele instructie: S⁺ > S⁻;
2. voor de variabele oefening: P₇ > P₁;
3. voor de interactie tussen instructie en oefening: instructie en oefening varië-ren onafhankelijk.

In afbeelding 3 wordt deze hypothese gra-fisch toegelicht.

2.2. Resultaten

Een variantie-analyse met herhaalde me-

bijeenkomst groep	1	2	3	4
S ⁺ -groep	I	<u>P₁</u> , P ₂ , P ₃	P ₄ , P ₅	P ₆ , <u>P₇</u>
S ⁻ -groep	<u>P₁</u> , P ₂ , P ₃	I	P ₄ , P ₅	P ₆ , <u>P₇</u>

Afb. 1. De volgorde waarin het cursusmateriaal in de cursus 1975-1976 werd aangeboden.

I = het aanbieden van instructiemateriaal.

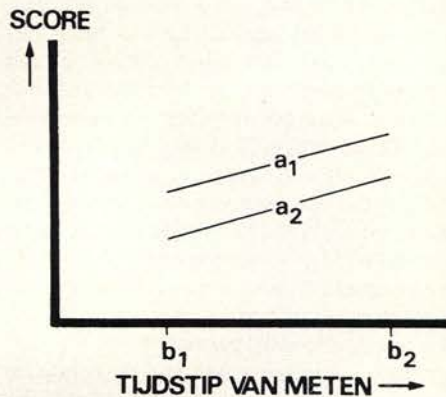
P = het aanbieden van een probleem; voor de onderstreepte problemen werd het percentage goede antwoorden berekend en statistisch bewerkt.

Samenvatting:

In dit tweede artikel wordt verslag ge-daan van de evaluatie van de cursus 'Het maken van een tandheelkundig behan-delplan'. Voor ieder van de vier geëvalu-eerde cursusjaren werd onderzocht of de cursus leerwinst tot gevolg had. Deze leerwinst bleek de eerste twee cursusja-ren niet significant te zijn. Na enkele wijzigingen in de cursusopzet, m.n. in het geven van feedback over de door de student gemaakte oplossingen, was er in de volgende jaren wel sprake van statis-tisch significante leerwinst. Bij de laatste cursus werd de instructie uitgebreid met een videoprogramma waarin een zoge-naamde 'domme student' de cursus doorloopt. Het ligt in de bedoeling dat de kijker door het Observeren van een Niet-ervaren probleem-Oplosser, zelf fouten leert te vermijden. De keuze voor de ver-anderingen in de cursus aangebracht en het invoeren van het ONO-programma wordt in dit artikel besproken.

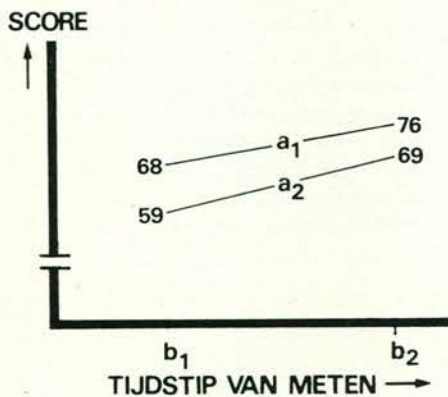
oefening instructie	P ₁ (b ₁)	P ₇ (b ₂)
S ⁺ (a ₁)	a ₁ b ₁	a ₁ b ₂
S ⁻ (a ₂)	a ₂ b ₁	a ₂ b ₂

Afb. 2. Combinaties van instructievolgorde en aanbiedingstijdstip van de problemen (cursus 1975-1976).



Afb. 3. Hypothetische interactie van instructie en oefening a₁ = S⁺-groep; a₂ = S⁻-groep; b₁, enz. = tijdstippen waarop de scores werden bepaald.

tingen op de variabele 'oefening' werd uitgevoerd op de resultaten bij het eerste en zevende probleem (tijdstip b_1 en b_2) van 26 studenten: 13 uit de S^+ -groep en 13 uit de S^- -groep (zie bijlage). Deze studenten waren aselect gekozen uit de studenten die de cursus volgens voorschrift en volgens rooster hadden doorlopen. De scores kwamen tot stand door van iedere oplossing het percentage correct door de student genomen beslissingen te berekenen, in vergelijking met beslissingen genomen in de standaardoplossing. Alleen die beslissingen waarvoor een criterium bestond, werden vergeleken (zie afb. 4).



Afb. 4. Het gemiddeld percentage goede antwoorden per groep studenten op de twee tijdstippen van meten (cursus 1975-1976).

Het hoofdeffect instructie (hypothese $S^+ > S^-$) was significant in termen van gemiddelde $a_1 = (68 + 76)/2 > a_2 = (59 + 69)/2$.

Het hoofdeffect oefening (hypothese $P_7 > P_1$) was niet significant,

$b_2 = (76 + 69)/2 > b_1 = (68 + 59)/2$.

Ook het interactie-effect was niet-significant.

2.3. Discussie

Het resultaat van een weliswaar bescheiden significant instructie-effect werd uitgelegd ten gunste van het gebruik van het instructiemateriaal, i.c. de syllabus. Daarentegen werd het resultaat van een niet-significant oefeneffect niet zonder meer opgevat als een argument tegen de oefening in het oplossen van een reeks problemen. Steun voor het voortzetten van de cursus werd gevonden in de resultaten van een enquête en de overige reacties van studenten.

Deze enquête werd gehouden middels vragenlijsten aan het einde van de eerste en aan het einde van de laatste zitting. Er bestonden vier verschillende vragenlijsten, namelijk voor de eerste en vierde zitting van respectievelijk de S^+ -groep en de S^- -groep. Het antwoord kon worden gegeven door een keuze te maken uit één van de

volgende antwoordcategorieën:

ja!; ja; ±; nee; nee!

Vragen die aan de orde kwamen waren o.a.:

– het doel van het oplossen van de oefenproblemen is mij duidelijk;

– de toelichting bij het oefenformulier is mij duidelijk;

– is de syllabus begrijpelijk;

– ik vond de aangeboden problemen moeilijk.

Voor elk type vragenlijst werd per vraag een chi-kwadraattoets uitgevoerd om na te gaan in hoeverre de gegeven antwoorden afwaken van de op grond van het toeval te verwachten antwoorden. De hypothese was dat op de vragen een bevestigend antwoord werd gegeven. Bij de S^+ -groep waren 16 antwoorden op de 20 vragen (van de beide enquêtes gezamenlijk) significant in de richting zoals voorspeld in de hypothese. Bij de S^- -groep waren 9 antwoorden op de 23 vragen (van de beide enquêtes gezamenlijk) significant. Het grote verschil in het aantal significante vragen tussen de beide groepen is gelegen in het feit dat de S^- -groep een cursusopzet kreeg voorgeschoteld die de ongebruikelijke volgorde van het cursusmateriaal inhield. Hierdoor werden b.v. vragen naar het nut van de syllabus, negatief beantwoord. Uit de open vraag van de enquête kwam sterk naar voren dat de student meer informatie wenste over de correctheid van de gemaakte oplossingen. Deze wenselijkheid was ook door de constructeurs van de cursus al ingezien, maar kon op het moment van de eerste cursus nog niet worden geëffectueerd.

3. De cursus 1976-1977

3.1. Inrichting

Op grond van de bevindingen bij de voorafgaande cursus werden de volgende wijzigingen aangebracht:

– Groepsindeling: om de organisatie van de cursus te vereenvoudigen, en om meer aandacht te kunnen besteden aan elk van de cursisten afzonderlijk, werden de studenten verdeeld over vier groepen, bestaande uit ongeveer 24 studenten. Iedere groep doorliep gedurende vier middagen de cursus.

– Instructie: voor de aanvang van de cursus dienden alle studenten de syllabus te bestuderen. Deze maatregel werd genomen omdat in de vorige cursus immers was aangetoond dat bestuderen van de syllabus tot betere resultaten leidt.

– Feedback: telkens nadat de student een oplossing had gemaakt kon hij deze vergelijken met de standaardoplossing van dat betreffende probleem. Hierbij werd de student niet geïnformeerd over de hoogte van de score van de door hem gemaakte oplossing.

– X-foto's: behalve de sonde stond de student voor ieder probleem een röntgenfoto ter beschikking.

– Toets en retentiesituatie; de derde middag gold als toetsmoment; de student mocht geen hulp inroepen van de instructeur of van een medestudent. Tijdens de vierde zitting werd nagegaan hoeveel kennis de student nog bezat omtrent het in de cursus geleerde. Dit werd de retentiezitting genoemd. Deze vond enige weken later plaats dan de toets, maar de omstandigheden waaronder de student de problemen diende op te lossen, waren bij de toetszitting (T) en de retentiezitting (R) de zelfde. Waar de toets een eventueel leer-effect aansluitend op de oefenfase laat zien, geeft het retentiemoment inlichtingen omtrent het leer-effect op de wat langere termijn.

Bij de inrichting van de cursus werd wederom rekening gehouden met een evaluatie. Deze keer wilden wij op de volgende vragen een antwoord:

1. Heeft feedback effect?

2. Heeft oefening effect?

Hiertoe werd aan twee van de vier groepen studenten feedback gegeven, waardoor een F^+ -groep (met feedback) en een F^- -groep (zonder feedback) ontstonden. In de afbeeldingen 5 en 6 zijn de volgorde waarin het cursusmateriaal werd aangeboden en de interactie tussen feedback en oefening schematisch weergegeven. De hypothesen waren respectievelijk:

1. voor de variabele feedback: $F^+ > F^-$;

2. voor de variabele oefening: $T_1 > R_1 > P_1$;

3. voor de interactie tussen feedback en oefening: bij de oefening een grotere leerwinst voor de F^+ -groep; en op tijdstip b_3 (retentie) gemiddeld hogere scores voor beide groepen dan op tijdstip b_1 (zie afb. 7).

3.2. Resultaten

Ten behoeve van de analyse van de resultaten werden uit de studenten die de cursus volgens voorschrift en volgens rooster hadden doorlopen, voor beide groepen 15 personen aselect gekozen.

Op de resultaten van het eerste oefenprobleem, het eerste toetsprobleem en het eerste retentieprobleem (respectievelijk tijdstip b_1 , b_2 en b_3) van de studenten uit de steekproef werd een variantie-analyse met herhaalde metingen op de variabele oefening toegepast (zie afb. 8 en ook de bijlage). Het hoofdeffect feedback was significant, de F^+ -groep scoorde in alle fasen van de cursus gemiddeld hoger dan de F^- -groep. Het hoofdeffect oefening en het interactie-effect waren niet significant.

3.3. Discussie

De analyse geeft een bevestiging van de

bijeenkomst groep	1	2	3	4
F ⁺ -groep	<u>P₁</u> *	P ₂ , P ₃ , P ₄	<u>T₁</u> , T ₂ , T ₃	<u>R₁</u> , R ₂ , R ₃
F ⁻ -groep	<u>P₁</u>	P ₂ , P ₃ , P ₄	<u>T₁</u> , T ₂ , T ₃	<u>R₁</u> , R ₂ , R ₃

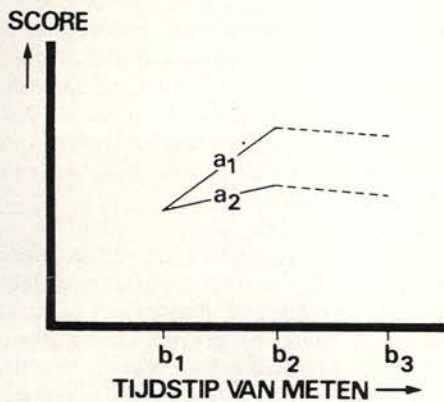
Afb. 5. De volgorde waarin het cursusmateriaal werd aangeboden in de cursus 1976-1977.

P = het aanbieden van een oefenprobleem.

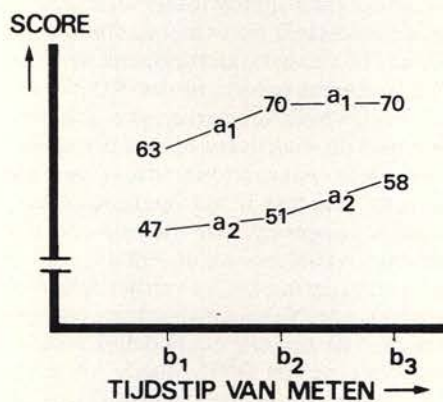
T = het aanbieden van een toetsprobleem.

R = het aanbieden van een retentieprobleem.

Voor de onderstreepte problemen werd het percentage goede antwoorden berekend en statistisch bewerkt.



Afb. 7. Hypothetische interactie tussen feedback en oefening a₁ = F⁺-groep; a₂ = F⁻-groep; b₁, enz. = tijdstippen waarop de scores werden bepaald (cursus 1976-1977).



Afb. 8. Het gemiddelde goede antwoorden per groep studenten op drie tijdstippen van meten (cursus 1976-1977).

eerste hypothese. Wanneer tussentijds na iedere oplossing feedback wordt verstrekt over die oplossingen heeft dit in het leerproces een grotere leerwinst tot gevolg dan wanneer men de feedback niet geeft.

Of leerwinst gerelateerd is aan oefening is niet aangetoond. Indien het hoofdeffect in zijn totaliteit niet significant is, is het weinig zinvol om de significantie van onderdelen van de hypothese, respectievelijk T₁ > P₁ en R₁ > P₁, te onderzoeken.

De F⁺-groep en de F⁻-groep hebben onder gelijke condities het eerste probleem opgelost. Een vraagteken moet daarom worden gezet bij de hoge gemiddelde P₁-score in de F⁺-groep. Een verklaring hiervoor zou kunnen zijn dat er tijdens de eerste zitting van de F⁺-groep van een onregelmatigheid sprake was, in die zin dat na het vergelijken door studenten van standaardoplossingen met eigen gemaakte oplossingen, verbeteringen zijn aangebracht in de eigen oplossing. Met de ervaring van die eerste zitting is zulks bij latere zittingen voorkomen. De gang van zaken tijdens de eerste zitting

heeft geen consequenties voor de hoogte van de scores van latere zittingen, zodat een vergelijking tussen de T₁ en R₁, van de F⁺-groep met de T₁ en R₁ van de F⁻-groep gerechtvaardigd is.

Aan het eind van de eerste, derde en vierde zitting werden de studenten vragenlijsten voorgelegd. Deze vragenlijsten stemden vrijwel overeen met die van het jaar ervoor en werden op overeenkomstige wijze bewerkt (zie 2.3.). De antwoorden boden in het algemeen ondersteuning aan de cursus en haar plaats in het onderwijsprogramma. Op de in totaal 22 vragen voor de F⁺-groep waren 17 antwoorden significant in de voorspelde richting, terwijl dit bij 13 antwoorden op 21 vragen voor de F⁻-groep het geval was.

4. De cursus 1977-1978

4.1. Inrichting

Ook dit jaar waren de studenten verdeeld in vier groepen van ongeveer 22 personen. In

oefening feedback	P ₁ (b ₁)	T ₁ (b ₂)	R ₁ (b ₃)
F ⁺ (a ₁)	a ₁ b ₁	a ₁ b ₂	a ₁ b ₃
F ⁻ (a ₂)	a ₂ b ₁	a ₂ b ₂	a ₂ b ₃

Afb. 6. Combinaties tussen feedback- en niet-feedback-groep en de aanbiedingsvolgorde van de problemen (cursus 1976-1977).

vier middagen moesten oplossingen worden gemaakt voor in totaal 12 problemen. De veranderingen ten aanzien van het vorige jaar zagen er als volgt uit:

- Alle studenten kregen na iedere gemaakte oplossing hierover feedback door het laten bekijken van de standaardoplossing.
- Er werden geen vragenlijsten meer voorgelegd, omdat de conclusies verbonden aan de resultaten van voorafgaande jaren reeds geëffectueerd waren.
- De X-foto's kwamen te vervallen. De foto's bleken de studenten meer in verwarring te brengen dan dat ze voor hen een hulp waren bij het onderzoek en het stellen van de diagnose; over het algemeen werd te veel waarde aan de X-foto's gehecht.
- Er werd aan de cursus de eis verbonden dat iedere student een voldoende moest halen voor de toets.

Bij de evaluatie werd onderzocht of oefening effect heeft.

Met betrekking tot deze vraag werden op drie momenten in de cursus de scores van een aantal studenten berekend. In schema zag de cursus eruit zoals weergegeven in afbeelding 9.

bijeenkomst			
1	2	3	4
<u>P₁</u> , P ₂ , P ₃	P ₄ , P ₅ , P ₆	<u>T₁</u> , T ₂ , T ₃	<u>R₁</u> , R ₂ , R ₃

Afb. 9. Volgorde waarin het cursusmateriaal werd aangeboden in de cursus 1977-1978.

P = het aanbieden van een oefenprobleem.

T = het aanbieden van een toetsprobleem.

R = het aanbieden van een retentieprobleem.

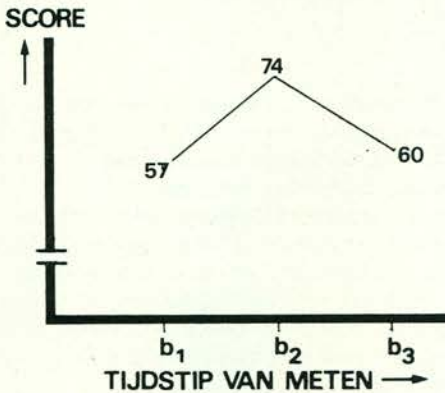
Voor de onderstreepte problemen werd het percentage goede antwoorden berekend en statistisch bewerkt.

De vierde bijeenkomst werd enige weken later gehouden dan de eerste drie bijeenkomsten.

De te onderzoeken hypothese was: $T_1 > R_1 > P_1$.

4.2. Resultaten

Een 1-factor variantie-analyse met herhaalde metingen op de variabele oefening werd uitgevoerd op de scores van 29 studenten, aselect gekozen uit de groep studenten die de cursus volgens voorschrift en volgens rooster hadden doorlopen (zie bijlage). De gemiddelde scores worden gegeven in afbeelding 10.



Afb. 10. Het gemiddeld percentage goede antwoorden op de drie tijdstippen van meten (cursus 1977-1978).

Het effect oefening was significant. Uitsplitst over de twee hypothese-onderdelen was het gemiddelde percentage goede antwoorden voor het eerste toetsprobleem (tijdstip b_2) significant hoger ($T_1 > P_1$), dan voor het eerste oefenprobleem (tijdstip b_1), terwijl het verschil tussen het eerste retentieprobleem (tijdstip b_3) en het eerste oefenprobleem niet significant was ($R_1 < P_1$). Van de in totaal 88 deelnemers zakten 39 studenten voor de eerste toetsmogelijkheid. Een tweede toets werd onder dezelfde condities afgenomen als de eerste. Naar keuze konden de studenten vóór hun tweede toets eerst nog oefenen of hun resultaten met een instructeur bespreken. Bij studenten die voor de tweede maal zakten werd een mondelinge toets afgenomen.

4.3. Discussie

Uit de resultaten blijkt dat wanneer de gemiddelde score voor het eerste toetsmoment wordt vergeleken met de gemiddelde score voor het eerste oefenmoment, er leerwinst kan worden vastgesteld. Dat dit niet het geval is bij vergelijking van het eerste retentieprobleem met het eerste oefenprobleem kan misschien gedeeltelijk worden verklaard uit het feit dat aan de retentiezittingen geen sancties waren verbonden en de studenten dientengevolge minder gemotiveerd waren om tot een goed resultaat te komen dan in de toetszitting.

5. De cursus 1978-1979

5.1. Inrichting

Nadat de cursus drie achtereenvolgende jaren was gehouden, waren er drie argumenten om de cursus voor het vierde jaar nogmaals te veranderen.

1. Het leek waarschijnlijk dat het instructiemateriaal voor de cursus verbeterd zou kunnen worden door een specifieke leersituatie te gebruiken welke beschreven is door Vaags (1975). Vaags heeft uit onderzoek bij het technisch wetenschappelijk onderwijs kunnen vaststellen dat voor het leren oplossen van problemen één bepaalde leersituatie duidelijk effectiever was dan drie andere onderzochte leersituaties. Hij maakte onderscheid tussen de volgende leersituaties: de OMO-situatie, de IKO-situatie, de IRS-situatie en de ONO-situatie. In de OMO-situatie (observatie van een modeloplossing) wordt aan de student getoond hoe iemand anders op perfecte wijze een probleem oplost. In de IKO-situatie (instructie over kenmerken van objecten) laat men de student het oplosproces doorlopen door hem voortdurend te vragen hoe de volgende fase in het oplosproces kan worden aangepakt. De IRS-situatie (instructie over responses op stimuli) is er op gebaseerd dat in elke fase van het oplosproces naar specifieke antwoorden wordt gevraagd, die zonnodig onmiddellijk worden gecorrigeerd. De ONO-situatie (observatie van een niet-ervaren oplosser) tenslotte, laat iemand zien die probeert een probleem op te lossen, daarbij dwaalwegen inslaat en wordt gecorrigeerd, beloond en gestraft. Het bleek dat in de ONO-situatie studenten de minste tijd nodig hadden om een probleem tot een oplossing te brengen terwijl met deze leersituatie ook kwalitatief het beste resultaat werd bereikt. De hypothese is dat door het observeren van het zoekende gedrag van de niet-ervaren oplosser geleerd wordt foutieve wegen te vermijden.

2. Er was gebleken dat bij de studenten, een duidelijke behoefte bestond om de oplossingen van de aangeboden problemen met de instructeur te bespreken; de omstandigheden waaronder de cursus plaatsvond lieten dat echter niet toe.

3. Het aantal invulfouten tenslotte dat bij de eerste oefenproblemen door de studenten werd gemaakt, leverde het laatste argument ten gunste van verandering. De indruk werd gewekt dat een gedeelte van de oefenstof werd gebruikt om tot een 'administratieve' beheersing van de methode te komen.

De boven geschetste mogelijke verbeteringen werden op de volgende wijze geëffectueerd.

Ad 1. De beschreven ONO-situatie werd verwezenlijkt in een videoprogramma waarmee een diapresentatie parallel liep.

In dit programma ziet men een onervaren student bezig met het oplossen van het tandheelkundige probleem, t.w. het behandelplan voor de restauratie van een gebitselement met een plastisch vulmateriaal. Het gedrag van de student wordt geprezen dan wel gecorrigeerd door een deskundige. Op video speelt zich de discussie tussen oplosser en deskundige af en wordt het invullen van het formulier getoond. Gelijktijdig wordt op een dia het desbetreffende element getoond.

Ad 2. Er werd in de vierde cursus op een andere wijze in feedback voorzien. Behalve feedback door het inkijken van de standaardoplossing vond eveneens terugkoppeling plaats in een discussie tussen een student of een aantal studenten en de instructeur. Om dit mogelijk te maken werden de studenten verdeeld in groepen van 8 à 10 personen. De groepen werden begeleid door twee instructeurs, zodat het aantal studenten per instructeur ten aanzien van vorige jaren werd gehalveerd. Door de discussie in zo'n kleine groep kan de student kennismaken van het oplosgedrag van zijn medestudenten en van de mening van de instructeurs.

Ad 3. Voorafgaand aan het eerste oefenprobleem (P_1) werd een inleidend videoprogramma getoond, waarin aan de hand van voorbeelden het gebruik van het invulformulier wordt uitgelegd. Verwacht werd dat door een betere instructie en begeleiding er minder oefenproblemen nodig zouden zijn om tot dezelfde beheersing van de stof te geraken als in voorgaande jaren. Ook werd verondersteld dat door die betere instructie en begeleiding de student voor het oplossen van een probleem minder tijd nodig zou hebben, zodat de cursus met de helft kon worden bekort tot twee halve dagen. Deze besparing op onderwijs-contacturen maakte het mogelijk het aantal studenten per instructeur te verkleinen tot de reeds genoemde verhouding. Echter bedacht moet worden dat deze gewijzigde opzet van de cursus aan de begeleidende instructeur andere eisen stelt, waardoor de voorbereidingstijd van de instructeur wordt verhoogd.

Met deze veranderingen kwam de cursus 1978/79 er als volgt uit te zien: de studenten werden verdeeld in groepjes van 8 à 10. Voor iedere groep was $\frac{3}{4}$ van een hele dag beschikbaar en twee weken later nog eens een halve middag. De cursus begon met het vertonen van het inleidende videoprogramma, waarna de cursisten gingen werken aan het eerste oefenprobleem. Vervolgens werd het ONO-programma getoond, waarna de drie resterende oefenproblemen werden afgewerkt. Later op de dag vond een groepsdiscussie plaats waarin problemen werden besproken die zich moeilijk in een geprogrammeerde cursus laten inpas-

sen en waarbij relaties werden gelegd met problemen die zich voor kunnen doen in een behandelplan voor een patiënt. Zo kwam bijvoorbeeld ter sprake het wel of niet behandelen van een bepaald defect en de omstandigheden die op zo'n beslissing van invloed zijn. Bij die beslissing komen zoveel overwegingen aan bod dat wanneer zo'n probleem in een geprogrammeerde cursus wordt opgenomen, het formulier zo uitgebreid moet zijn dat het onoverzichtelijk wordt, terwijl het bovendien voor de student dan nog niet mogelijk is middels het formulier zijn argumentatie kenbaar te maken. Daarnaast is het niet mogelijk met de formulieren van de standaardoplossing feedback te geven over de argumentatie die tot de juiste oplossing leidt. Tijdens de bijeenkomst die twee weken na de cursusdag plaatsvond, werd de toets afgenomen. Tussen de cursusdag en het moment van toetsen bestond gedurende een halve middag de gelegenheid extra te oefenen.

Afbeelding 11 is een schematische weergave van de volgorde waarin het cursusmateriaal werd aangeboden. De evaluatie van de cursus spitste zich toe op de volgende vragen:

1. Heeft de cursus als geheel effect, m.a.w. wordt er leerwinst geboekt?
2. Heeft de ONO-instructie effect?
3. Heeft extra oefenen effect?

De hypothesen waren respectievelijk:

1. $T_1 > P_1$.
2. $P_2 > P_1$.
3. Voor de variabele extra oefening: het percentage geslaagden is groter bij extra oefening.

5.2. Resultaten

Een 1-factor variantie-analyse met herhaalde metingen werd uitgevoerd op de scores van 26 studenten, aselect gekozen uit de studenten die de cursus volgens voorschrift en volgens rooster hadden doorlopen (zie bijlage).

De gemiddelde scores worden gegeven in afbeelding 12.

Het gemiddelde percentage goede antwoorden over het eerste toetsprobleem (tijdstip b_3) was significant hoger dan voor het eerste oefenprobleem (tijdstip b_1). De eerste hypothese werd derhalve bevestigd: er is sprake van leerwinst. Wat de ONO-instructie betreft, werd gevonden dat voor P_2 (tijdstip b_2) gemiddeld hoger werd gescoord dan voor P_1 . Het verschil bleek niet significant te zijn.

Binnen het totaal van 77 studenten hadden er 61 gebruik gemaakt van de extra oefening (het betreft hier in alle gevallen vier extra oefenproblemen). Van die 61 studenten zakten 12 voor de toets (20%). Van de overige 16 studenten, die niet extra oefen-

bijeenkomst						
1			gelegenheid voor extra oefenen			
			2			
uitleg invulformulier	P_1	ONO- videoband	P_2, P_3, P_4	groeps- discussie	P_5, P_6, P_7, P_8	T_1, T_2, T_3

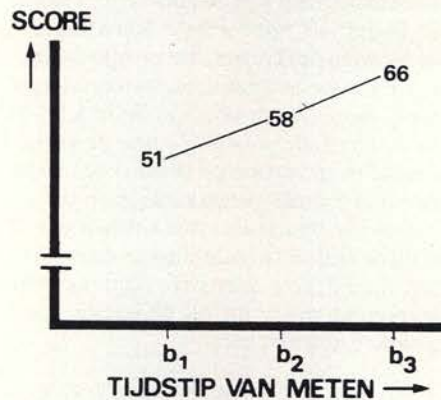
Afb. 11. Volgorde waarin het cursusmateriaal werd aangeboden in de cursus 1978-1979.

P = het aanbieden van een oefenprobleem.

P_{5-8} = de facultatieve aanbieding van (maximaal) vier extra oefenproblemen.

T = het aanbieden van toetsproblemen.

Voor de onderstreepte problemen werd het percentage goede antwoorden berekend en statistisch bewerkt.



Afb. 12. Het gemiddeld percentage goede antwoorden op de drie tijdstippen van meten (b_1 , b_2 en b_3).

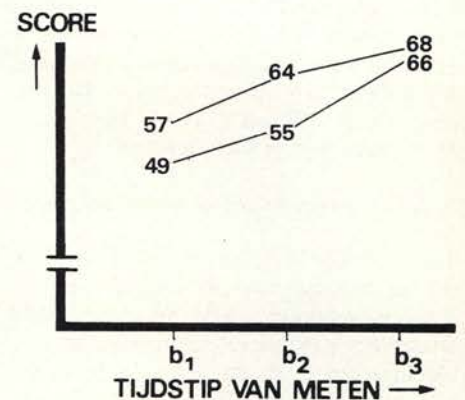
den zakten er vier (25%). Het verschil bleek niet significant te zijn.

Van de in totaal 16 gezakte studenten behaalden twee studenten ook bij de tweede toetsgelegenheid niet de vereiste score.

5.3. Discussie

In dit laatste cursusjaar is niet, zoals in vorige jaren, het effect van oefening onderzocht, maar het effect van de cursus als geheel. Deze wijziging is aangebracht omdat het effect van oefenen alleen niet meetbaar was doordat tijdens de cursus de ONO-instructie en de groepsdiscussie plaatsvonden, die beide wel invloed op het niveau van de toetsoplossing hebben maar niet op het niveau van de eerste oefenoplossing. Uit de constatering dat de gemiddelde score voor het eerste toetsprobleem significant hoger was, dan voor het eerste oefenprobleem kan worden afgeleid dat er leerwinst uit de cursus werd behaald.

Ten aanzien van de vraag of de ONO-instructie effect had werd vastgesteld dat het verschil in hoogte tussen de scores die werden behaald voor het eerste en tweede oefenprobleem niet significant was. Dit ligt in de lijn der verwachting als men bedenkt dat er weliswaar een significant maar geen spectaculair verschil was tussen de scores



Afb. 13. Het verschil in scores tussen de groep studenten die facultatief vier extra oefenproblemen oploste (onderste lijn) en de groep studenten die dat niet deden (bovenste lijn). Extra oefening betekent hier in alle gevallen vier extra oefenproblemen.

van P_1 en T_1 . Tussen b_2 en b_1 werd geoefend op één element en tussen b_3 en b_2 werd geoefend op 3 tot 7 elementen (afhankelijk van deelname aan de extra oefening). In beide fasen van de cursus vindt ongeveer evenveel leerwinst plaats (zie afb. 12). Het ligt voor de hand te veronderstellen dat de relatief grote leerwinst tussen b_2 en b_1 mede tot stand is gekomen door de ONO-instructie die in die fase van de cursus werd gegeven. De onderzoekopzet was echter zodanig dat dat niet kan worden aangetoond.

In afbeelding 13 wordt het verschil getoond tussen de gemiddelde scores van de groep studenten die niet extra oefende en de groep studenten die wel extra oefende. Terwijl de groep studenten die extra oefende aanzienlijk meer leerwinst behaalde dan de groep studenten die niet extra oefende, is er nauwelijks verschil tussen de gemiddelde scores voor het eerste toetsprobleem van de respectieve groepen (tijdstip b_3). Dit is in overeenstemming met de waarneming dat er ook niet significant meer of minder studenten zakten al naar gelang er wel of niet extra was geoefend. De oorzaak van de

tegenstrijdigheid van een grotere leerwinst en geen beter resultaat ligt in het feit dat de groep studenten die extra oefende gemiddeld veel lager scoorde over het eerste en tweede oefenprobleem in vergelijking met de groep studenten die niet extra oefende (zie afb. 13; tijdstip b_1 en b_2). Het is mogelijk dat de groep die niet extra oefende is samengesteld uit studenten die in de oefenfase constateerden dat zij de aangeboden problemen beter oplossen dan hun medestudenten en daardoor de conclusie trokken dat voor hen extra oefening overbodig was. Een blijkbaar terechte conclusie!

6. Algemene discussie

Bij het construeren van de cursus zijn we te werk gegaan volgens een schema dat is ontworpen voor het ontwikkelen van onderwijs waarin probleemoplossen centraal staat (Vaags, 1975). Dit schema bestaat uit de vijf hieronder genoemde fasen.

1. Het begrenzen van het probleemgebied

Tijdens het ontwerpen van de cursus werd bij het bepalen van het probleemgebied rekening gehouden met de voorkennis van studenten en met de relevantie van het probleemgebied in de studie en in de latere beroepsuitoefening. De vorm waarin het onderwijs wordt gepresenteerd werd mede bepaald door de resultaten van de analyse van het probleemoplossen zoals dat door onervaren studenten wordt gedaan (Kersten e.a., 1980). De vorm wordt verder gegeven door de mogelijkheden van tijd, mankracht en ervaring in het ontwikkelen van onderwijs.

2. Het kiezen van een verzameling oefenen toetsproblemen

Binnen de begrenzing van het probleemgebied werden oefen- en toetsproblemen geselecteerd. Er werden bij het kiezen van de verzameling toetsproblemen geen andere criteria aangelegd dan bij het selecteren van de oefenproblemen. Wel werden die elementen geselecteerd die de minste aanleiding gaven tot discussie tussen de beoordelaars onderling.

3. Het maken van oplossingen door deskundigen

De standaardoplossingen werden verkregen door de oefen- en toetsproblemen door tandartsen, zonodig in onderling overleg, te laten oplossen.

4. Het bepalen van criteria voor de beoordeling van de door de studenten geproduceerde oplossingen

Die beslissingen in de oplosroute die van essentieel belang zijn bij het op de juiste

wijze afwerken van de stappen werden als criteria bij de beoordeling gebruikt (b.v. een beslissing t.a.v. een afwijkend fissuurpatroon wordt niet als essentieel aange-merkt).

5. Het maken van instructiemateriaal

In deze fase van het schema van Vaags is in eerste instantie de syllabus gemaakt. Het streven de cursus voortdurend te verbeteren heeft tenslotte geleid tot het maken van het videoprogramma.

De statistische bewerking van de resultaten van de achtereenvolgende cursussen zijn uitgevoerd om inzicht te verkrijgen in de doelmatigheid van de cursus, met name wat betreft de hoeveelheid leerwinst die kon worden verkregen. De conclusies uit de statistische bewerkingen werden steeds zoveel mogelijk verwerkt in de inrichting van de volgende cursus. Zo is in de vorige paragrafen van dit artikel te zien hoe min of meer ingrijpende veranderingen in de inrichting van de cursus ertoe hebben geleid dat bij de laatste twee cursussen een statistisch significante leerwinst kon worden aangetoond, waar dit bij de eerste twee cursussen nog niet kon.

Er zijn een aantal veranderingen in de cursus die vooral hebben bijgedragen tot deze grotere leerwinst. Als eerste moet worden genoemd het geven van feedback aan de student over de door hem gemaakte oplossing.

Met het verdelen van de studenten in kleine groepjes (cursus 1978-1979) werd deze feedback verder uitgebouwd. De student kan nu niet alleen zien waar hij fouten in de oplossing heeft gemaakt maar ook, door de mogelijkheid een discussie met de instructeur(s) en medestudenten aan te gaan, te horen krijgen waarom zijn oplossing niet goed is.

Een tweede belangrijke verandering in de cursus was het optimaliseren van de criteria die voor de beoordeling van de student bruikbaar waren gebleken. Tijdens de eerste jaren dat de cursus werd gegeven bevonden zich onder die criteria sommige die niet voldoende nauwkeurig waren omschreven en op uiteenlopende manier door de onderscheiden beoordelaars werden gebruikt. Daarom werden bijvoorbeeld in het derde cursusjaar beslissingen ten aanzien van kleurafwijkingen niet meer als criteria voor de beoordeling gehanteerd. Dat beslissingen in de oplosroute die te onnauwkeurig zijn om als criterium te worden gebruikt, toch blijven gehandhaafd, vindt zijn oorzaak in het feit dat zonder die beslissingen niet in alle gevallen een juiste beslissing in een volgende stap kan worden genomen.

Een voor de hand liggende manier voor het vergroten van leerwinst is het aanbieden

van meer oefenstof. We zijn echter van mening dat uitbreiding van de oefening in het kader van de cursus, weinig invloed zal hebben op de leerwinst op lange termijn. Uitbreiding van de oefening kan beter worden gezocht in een herhaling van de cursus in een vorm waarin andere en uitgebreidere aspecten van de tandheelkunde als probleem worden aangeboden.

Wat ons inziens wel grote invloed heeft gehad op het vergroten van de leerwinst is het invoeren van een verplichte minimumscore per toets-element en voor de toets-elementen gezamenlijk. Het bleek dat het ontbreken van sancties bij het niet bereiken van een voldoende niveau de cursus een te vrijblijvend karakter geeft. In het algemeen blijkt helaas een cijfer voor de student interessanter te zijn dan de vraag of kennis van het desbetreffende studie-onderdeel van wezenlijk belang is.

Het bepalen van de grens voldoende/onvoldoende bracht een aantal problemen met zich mee. Het blijkt namelijk dat er altijd een aantal triviale 'administratieve' fouten worden gemaakt, die van invloed zijn op het aantal juiste door de student genomen beslissingen, maar die geen rol mogen spelen in de beslissing voldoende/onvoldoende. Aanvankelijk waren wij de mening toegedaan dat voor een voldoende niveau, de grens op 85% moest liggen. Dit betekende dat voor een voldoende, 85% van de als criteria aangemerkte beslissingen juist moet zijn. Maar toen bleek dat deskundigen die vooraf niet bekend waren met de cursus deze score ook niet haalden, is dit herzien. Een 100%-score wordt alleen bereikt door deskundigen die bekend zijn met de cursus en onder optimale omstandigheden het probleem oplossen (met de mogelijkheid collega's en literatuur te raadplegen). Soortgelijke ervaringen zijn ook opgedaan bij een evaluatie van de studieprestaties in de Maastrichtse artsopleiding (Greep, 1979).

Uiteindelijk is vastgesteld dat per element minimaal 45% van de als criteria aangemerkte beslissingen juist moeten zijn en dat voor de drie toets-elementen gezamenlijk deze score gemiddeld hoger moet zijn dan 55 (Kersten, 1980).

De beschreven cursus voorziet naar onze mening in een leemte die er bestond in het onderwijs, enerzijds wat de oefening in het systematisch verwerken van gegevens betreft, anderzijds met betrekking tot het oefenen in het herkennen en beoordelen van afwijkingen zoals b.v. cariës. Het beoordelen van de afwijkingen bij een enkel gebits-element en het maken van een plan voor de behandeling ervan verschilt in wezen niet van het beoordelen van de afwijkingen in een gebit en het maken van een plan voor de geïntegreerde behandeling ervan, anders dan dat er meer relaties zijn tussen de afzonderlijke aspecten van het probleem. Het probleemgebied is uitgebreider, maar

de informatieverwerking en de besluitvorming geschieden op gelijke wijze.

Title: Developing a dental treatment plan: II. Evaluation of the course.

Summary:

In this second article, the evaluation of the course in dental treatment planning is reported. For each of the evaluated four years, it was investigated if any learning could be demonstrated. This learning proved to be statistically not significant with regard to the first two years that the course was held.

After some modifications in the structure of the course, particularly in the way feedback was given to the students concerning their problem-solving activity, there was evidence of statistically significant learning in the following two years. In the last course, the course instruction was extended by the addition of a video-programme in which a so-called 'weak' student proceeds through the course. The assumption is that by Observation of a Non-experienced problem-Solver, one can learn to avoid mistakes. The choice for modifying the course and the introduction of the ONS-programme is discussed in this article.

Literatuur:

1. Greep, J. M. (1979): Het onderwijs aan de medische faculteit in Maastricht. Medisch Contact 35: 1107-1114.
2. Kersten, H. W., Steures, R. W. R., Tromp, Th. J. M. (1980): Het maken van een tandheelkundig behandelplan. I: Constructie van een cursus. Ned. Tijdschr Tandheelkd 87: 134-139.
3. Kirk, R. E. (1968): Experimental design: Procedures for the behavioral science. Brooks/Cole, Belmont (Cal.).
4. Vaags, D. W. (1975): Leren oplossen van problemen in de technische mechanica. Onderzoek van Onderwijs 4: 3-8.

Juni 1980.

Adres: Louwesweg 1,
1066 EA Amsterdam.

Bijlage

Variantie-analysetabellen van de tot nu toe gehouden cursussen.

Opmerking: In de hierna volgende tabellen zijn de getallen afgerond. Voor meer informatie dan in deze bijlage is bijeengebracht, wordt verwezen naar Kirk (1968, m.n. blz. 131 e.v. en 245 e.v.).

1. De cursus van 1975-1976

Uitgevoerd werd een 2-factor variantie-analyse, waarvan factor B met herhaalde metingen; factoren A en B gefixeerd, 'factor' Personen aselekt.

Tabel I. Variantie-analyse op gegevens van de cursus 1975-1976.

Variatiebron	Kwadraten-som	Aantal vrijheidsgraden	Gemiddelde kwadraten-som	F	P
Instructie (A)	777	1	777	2,91	.10
Personen binnen groepen	6414	24	267		
Problemen (B)	991	1	991	1,92	n.s.
AB	4	1	4	<1	
B × personen binnen groepen	12407	24	517		

2. De cursus van 1976-1977

Het type uitgevoerde variantie-analyse is gelijk aan dat van de eerste cursus.

Tabel II. Variantie-analyse op gegevens van de cursus 1976-1977.

Variatiebron	Kwadraten-som	Aantal vrijheidsgraden	Gemiddelde kwadraten-som	F	P
Feedback (A)	5554	1	5554	19,0	.01
Personen binnen groepen	8184	28	292		
Problemen (B)	1234	2	617	2,6	.10
AB	179	2	89	<1	
B × personen binnen groepen	13452	56	240		

3. De cursus van 1977-1978

Uitgevoerd werd een 1-factor variantie-analyse met herhaalde metingen; factor gefixeerd, 'factor' Personen aselekt.

Tabel III. Variantie-analyse op gegevens van de cursus 1977-1978.

Variatiebron	Kwadraten-som	Aantal vrijheidsgraden	Gemiddelde kwadraten-som	F	P
Tussen problemen	5019	2	2509	8,22	.01
Tussen personen	7802	28	279	<1	
Residu	17088	56	305		

4. De cursus van 1978-1979

Het type uitgevoerde variantie-analyse is gelijk aan dat van de derde cursus.

Tabel IV. Variantie-analyse op gegevens van de cursus 1978-1979.

Variatiebron	Kwadraten-som	Aantal vrijheidsgraden	Gemiddelde kwadraten-som	F	P
Tussen problemen	2899	2	1449	3,43	.05
Tussen personen	11361	25	454	1,08	n.s.
Residu	21116	50	422		