

**WERKGROEP NATUURKUNDE-DIDAKTIK**

**verslag van de konferentie**

**woudschoten**

**EVALUATIE**

WERKGROEP NATUURKUNDE-DIDAKTIEK

VERSLAG VAN DE 9E KONFERENTIE WOUDSCHOTEN

20 - 21 DECEMBER 1974

EVALUATIE

LAB. VOOR VASTE STOF

SORBONNELAAN 4

UTRECHT

FEBRUARI 1975

WERKGROEP NATUURKUNDE-DIDAKTIK

Laboratorium voor Vaste Stof

Sorbonnelaan 4

UTRECHT - De Uithof

Tel. 030 - 531179

BESTUUR

*voorzitter:*

H.P. Hooymayers                      Sorbonnelaan 4                      Utrecht

*sekretaris:*

C. Floor                                  Rooseveltlaan 560B                      Utrecht

*leden:*

H.F. van Aalst                          Sorbonnelaan 4                          Utrecht

H.A. Créton                              Sorbonnelaan 4                          Utrecht

J. van Lidth de Jeude                  Mgr. van de Weteringstraat 53                  Utrecht

J.Ph. Steller                              Technische Hogeschool, Postbus 513                  Eindhoven

ORGANISATIE KONFERENTIE

H.A. Créton

C. Floor

H.P. Hooymayers

H. Veurink

## VOORWOORD

Evenals voortgaande jaren wordt de Woudschotenkonferentie van 1974, die gewijd was aan het thema 'evaluatie', gevolgd door een schriftelijk verslag, dat aan alle leden wordt toegezonden.

De Werkgroep blijft hiermee een vier jaar geleden ingeslagen weg bewandelen; niet uit angst om er vanaf te wijken, maar omdat de leden van de Werkgroep een dergelijk verslag op hoge prijs blijken te stellen. Op een aan alle 350 leden toegezonden enquête werden 150 antwoorden ontvangen. Nagenoeg zonder uitzondering hielden de antwoorden in dat het verslag geheel (!) of gedeeltelijk gelezen was en dat men ook dit jaar weer graag een verslag zou ontvangen. Ook via andere kanalen (zie bijvoorbeeld Faraday 44, 4, p.89) werd de behoefte aan een dergelijk verslag bevestigd.

Daarom zijn wij blij U hierbij het verslag van de Woudschotenkonferentie 1974 te kunnen presenteren.

Het verslag valt in vier delen uiteen.

In een *inleidend gedeelte* wordt gesproken over het doel van de conferentie en de keuze van het thema; verder wordt algemene informatie over de conferentie verschaft (blz. 1 t/m 4).

Het *tweede deel* bevat de volledige tekst van de ter conferentie uitgesproken voordrachten (blz. 5 t/m 54).

De werkzaamheden in de discussiegroepen komen in *deel drie* aan de orde. Achtereenvolgens treft U per discussieopdracht aan:

- de formulering van de opdracht
- de bijbehorende (voor-)informatie
- het verslag van de discussiegroep(en) die zich met de opdracht heeft (hebben) bezig gehouden

Tenslotte bevat *deel vier* een aantal bijlagen. De belangrijkste hiervan zijn de teksten van de voordrachten, die op de wintervergadering van Velines (thema: schriftelijk eindexamen) zijn gehouden door J.Smit (inspekteur vwo/havo) en B.Th. Berendts (vertegenwoordiger C.M.L.N.).

Wij danken de schrijvers en Hans van Aalst van de sectie natuurkunde van Velines voor hun bereidwillige medewerking het opnemen van deze teksten in dit verslag mogelijk te maken.

Namens de Werkgroep  
Natuurkunde-Didaktiek

C. Floor, sekretaris



DEEL 1: ALGEMENE INFORMATIE OVER DE KONFERENTIE

INHOUD

	blz.
1. Het thema: 'evaluatie'	1
2. Het doel van de conferentie	1
3. Programma	1
4. De slotevaluatie	3
5. De conferentie in cijfers	3



## DEEL 1: ALGEMENE INFORMATIE OVER DE KONFERENTIE

### 1. Het thema: 'evaluatie'

In de loop van 1973 is aan de leden van de Werkgroep Natuurkunde-Didaktiek een enquête verzonden, waarin hen werd gevraagd mogelijke thema's voor volgende Woudschotenkonferenties te noemen. Na het thema 'werkvormen', waaraan de vorige conferentie gewijd was, werd het thema 'evaluatie' het meest genoemd. Ook in de plenaire slotevaluatie van de conferentie over werkvormen werd dit thema met nadruk ingebracht (zie verslag Werkvormen, pag. 35). Daar bovendien het bestuur van de Werkgroep een bezinning op het hoe en waarom van evaluatie als zeer zinvol ziet werd besloten dit onderwerp tot konferentietema te kiezen.

### 2. Het doel van de conferentie

De conferentie had tot doel de deelnemers een beeld te geven van de mogelijkheden en beperkingen van een aantal evaluatietechnieken. Daarnaast was het doel de discussie op gang te brengen over een aantal andere aspecten van evaluatie, aan de hand van vraagstellingen als:

- wat kan een leerling, die eindexamen natuurkunde heeft gedaan
  - hoe kan het experiment in het schoolonderzoek worden ingebouwd
  - wat is de betekenis van het cijfer, hoe komt het cijfer tot stand
- etc.

### 3. Programma

Door middel van vier voordrachten - een enkele maal gevolgd door een korte plenaire discussie - werden de verschillende onderwerpen ingeleid. In discussiegroepen kon verder van gedachten gewisseld worden over punten die in de voordrachten aan bod waren gekomen. Daarnaast werd in een aantal groepjes gediscussieerd aan de hand van vraagstellingen, die andere aspecten van evaluatie belichtten, dan door de sprekers waren aangedragen.

Deelnemers uit verschillende discussiegroepen rapporteerden elkaar aan het einde van de tweede konferentiedag over het besprokene in hun subgroep met behulp van aan de muur bevestigde flappen, op de zogenaamde 'rapportagemarkt'. Konkreet zag het programma er als volgt uit:



WERKGROEP NATUURKUNDE-DIDAKTIK  
in samenwerking met Velines

Woudschotenkonferentie 1974: EVALUATIE

*vrijdag 20 december*

- 14.30 - 15.00 ontvangst en koffie  
15.00 - 15.15 opening door de voorzitter van de Werkgroep Natuurkunde-Didaktiek,  
Dr. H.P. Hooymeyers  
15.15 - 15.30 inleiding door de voorzitter van de conferentie, Drs. J.W. Solberg,  
directeur C.I.T.O.  
15.30 binnenkomst laatkomers\*  
15.30 - 16.15 Beoordeling: zekerheden en twijfels; wat wil en kan men meten van  
leerlingen die eindeksamen hebben gedaan met natuurkunde in hun  
pakket door Dr. S. Auer, docent algemene didaktiek, Pedagogisch  
Didaktisch Instituut van de G.U. Amsterdam  
16.15 - 16.45 thee  
16.45 - 17.30 Schoolonderzoek, onder andere hoe kan het schoolonderzoek bijdragen  
tot een beoordeling van andere essentiële kwaliteiten en verworven-  
heden van de leerling dan op het centrale schriftelijke werk moge-  
lijk is door Dr. J.Ph. Steller, lector natuurkunde-didaktiek T.H.  
Eindhoven  
17.30 - 17.45 groepsindeling  
18.00 diner  
19.30 - 21.00 groepsdiskussies  
(20.15 koffie)

*zaterdag 21 december*

- 8.15 ontbijt  
9.00 - 10.00 Psychometrische aspecten van de meerkeuzetoets; validiteit, betrouw-  
baarheid, p- en a-waarden, itemhomogeniteit ( $r_{it}$ ); een aantal voor-  
beelden van vier-keuzetests, hun p-, a- en  $r_{it}$ -waarden bij verschil-  
lende schooltypes door Drs. P.K.H. Broekman, hoofd afdeling 8- en  
γ-vakken C.I.T.O.  
10.00 - 10.30 koffie  
10.30 - 11.30 Inleiding over evaluatie in een breder kader geplaatst door  
Dr. J.S. ten Brinke, docent algemene didaktiek, Pedagogisch Didak-  
tisch Instituut van de Leraarsopleiding, Utrecht  
11.30 - 11.45 aansluitende discussie  
11.45 - 12.15 groepsindeling naar opdracht  
12.30 lunch  
13.30 - 14.45 groepsdiskussies  
14.45 - 15.15 thee  
15.15 - 16.00 rapportage (markt)  
16.00 - 16.15 evaluatie van de conferentie en afsluiting

\* De subsidie voor de conferentie mag alleen worden verstrekt als de conferentie in de vrije  
tijd plaatsvindt. Voor degenen, die van ver moeten komen, of wiens lessen die dag laat  
eindigen, is deze tweede binnenkomst ingelast.

Na afloop van het programma van de eerste dag werden door D.P. Jax Mulder enkele films gedraaid, die door een groot aantal deelnemers werden bekeken. Ook de gebruikelijke boekentafel genoot grote belangstelling.

#### 4. De slotevaluatie

In de plenaire slotevaluatie bleek dat de deelnemers de volgende konferentie bij voorkeur weer op vrijdag en zaterdag voor Kerstmis op Woudschoten gehouden zien worden. Planning van plenaire diskussies, uitloopmogelijkheden voor de groepsdiskussies en de manier van rapportage van de diskussiegroepen dienen nader bekeken te worden.

Inmiddels waren op een flap in de hal de volgende suggesties voor toekomstige konferentietema's genoteerd:

- differentiatie binnen klasseverband
- media in het natuurkundeonderwijs:
  - . wat versta je eronder
  - . welk gebruik (voorbeelden)
- affektieve (mis)vorming in de natuurwetenschappelijke vakken
- maatschappij en natuurkundeles
- 'andere werkvormen', maar dan praktisch

#### 5. De konferentie in cijfers

Voor de konferentie hadden zich 195 deelnemers aangemeld. Helaas moesten 45 van hen worden afgewezen, daar het aantal konferentiegangers om organisatorische redenen tot 150 beperkt moest worden.

Een aantal deelnemers was reeds voor de konferentie in diskussiegroepen bijeen gekomen om een aspekt van evaluatie te bespreken.

De belangstelling voor de verschillende diskussietema's was als volgt verdeeld:

tema:	aantal aanmeldingen:
<i>vrijdag</i>	
1 analyse C.M.L.N.-doelstellingen	24
2 analyse van vragen met behulp van de categorieën van Klopfer	13
3 de categorieën van Klopfer konkreet in de lespraktijk	20
4 maken van vragen met behulp van de categorieën van Klopfer	11
5 cijfergegevens	40
6 praktikumbeoordeling	8
- overige	1

tema:	aantal aanmeldingen:
<i>zaterdag</i>	
5 cijfergeven	6
7 beoordeling van leerlingen	7
8 itemanalyse	51
9 evalueren in een breed kader	16
10 evalueren van complexe natuurwetenschappelijke doelen	5
11 problemen bij beoordelingslijst van leerlingen	10
- overige	3

Slechts een enkeling behoefde in een groep van zijn tweede keuze geplaatst te worden.

---

## DEEL 2: DE VOORDRACHTEN

### INHOUD

	blz.
1. Voordracht Auer:	
'Evaluatie: zekerheden en twijfels'	5
1.1. Inleiding en verantwoording	5
1.2. Het profiel van de eindexamenkandidaat met natuurkunde in het pakket	6
1.3. Examendoelen en examenvragen	8
1.4. Evaluatievorm en leerresultaat	15
Bijlage bij voordracht Auer	20
2. Voordracht Steller:	
'De rol van het schoolonderzoek in de summatieve evaluatie'	22
3. Voordracht Broekman:	
Psychometrische aspecten van de meerkeuzetoets	32
4. Voordracht Ten Brinke:	
Ontwikkelingen in het evaluatiedenken en wat we daar als leraar mee zouden kunnen doen	49



DEEL 2: DE VOORDRACHTEN

1. Voordracht Auer: 'Evaluatie: Zekerheden en twijfels'

1.1. Inleiding en verantwoording

Dit betoog voldoet in een opzicht aan de eisen, die door de volksmond aan goede dingen worden toegedicht. Het bestaat uit drie delen.

In het eerste gedeelte worden meer vragen gesteld, dan beantwoord. Die vragen betreffen in feite de plaats van het vak natuurkunde binnen die gedeelten van het VWO/AVO welke vooral als een voorbereiding op de overgang naar het beroepsvoorbereidend- of begeleidend onderwijs kunnen worden gezien, de universiteiten en hogescholen daaronder nadrukkelijk begrepen. Ook al betekent dat laatste misschien voor menigeen een onverantwoordelijke rek in de definitie van beroepsvoorbereidend onderwijs. Die vraag is in feite niet los te maken van de meer algemene vraag naar de functie die de eindfase van de diverse vormen van voortgezet onderwijs moet hebben. Op dit terrein zijn weinig principiële uitspraken gedaan, waardoor bijvoorbeeld een groot deel van de invulling der Mammoetstructuren in den blinde is geschied. In het eerste gedeelte van mijn betoog zal ik nagaan welke vragen in feite beantwoord moeten zijn alvorens kan worden besloten aan welke eisen prestaties van examenkandidaten zouden moeten voldoen.

Al bestaat er een grote ijdelheid in de meer principiële regionen (van een vacuum zou ik niet willen spreken) dat heeft bedrijvigheid op het werkterrein der examenprogramma's in de verschillende schoolvakken niet verhinderd. Integendeel. De bestaande vrijheid heeft onder meer de verscheidene moderniseringsprogramma's ertoe gebracht nieuwe programma's het licht te doen zien. Wat trouwens ook hun opdracht was. Ook de Commissie Modernisering Leerplan Natuurkunde is in juni 1974 gereedgekomen met haar Rapport (zie bijl.blz.20). Op een belangrijk gedeelte van dit rapport is het woord "nieuw" inderdaad van toepassing, omdat aan eindexamenkandidaten werkelijk ook andere, bovendien scherper omschreven eisen worden gesteld. Aan de hand van voornamelijk buitenlandse voorbeelden (2, 3 en 4) zal worden gedemonstreerd in hoeverre deze eisen inderdaad "anders" zijn. Het laatste gedeelte van mijn betoog zal gewijd zijn aan de vorm, waarin de afsluitende beoordeling - de summatieve evaluatie - wordt gegoten.

Met die vorm ben ik om verschillende redenen niet erg gelukkig. Die redenen zal ik U uiteenzetten. Daaraan zal ik - heel voorzichtig - enkele suggesties tot verandering koppelen. Steller zal dat in zijn voordracht ongetwijfeld in bepaalde opzichten verder uitwerken.

1.2. Het profiel van de eindexamenkandidaat met natuurkunde in het pakket.

Het woord profiel is ontleend aan de formulering waarmee Hooy-mayers mij deze zomer uitnodigde, in Woudschoten te komen spreken. Indien men het woord profiel relateert aan het huidige gebruik, dat bij universitaire en andere vakatures momenteel van dit begrip gemaakt wordt, dan worden terstond de mogelijkheden en moeilijkheden duidelijk. De profielen die ten behoeve van de vervulling der voornoemde vakatures worden opgesteld, zijn vaak van een ontroerende vaagheid. Toch omvatten die profielen tevens het eisenpakket en tevens een wensenlijst ten aanzien van de vervullen der vakature. Een goed profiel omvat dus een scherp eisenpakket en een duidelijke wensenlijst, het opstellen ervan is alles behalve gemakkelijk. Toch betreft het in de hier bedoelde gevallen een vakature voor een bepaalde functie, waarin dus een reeks vooraf bekende naast een aantal onvoorspelbare werkzaamheden moet worden verricht.

Wat soms gelukt in zulk een specifieke situatie moet nu vertaald worden naar een nog niet erg specifieke fase van een algemene opleiding met een weliswaar eindig, maar zeer groot aantal mogelijke vervolgopleidingen. De vraag naar het profiel van een eindexamenkandidaat met een bepaald vak (hier: natuurkunde) in zijn pakket blijkt nu een kapstokvraag te zijn, waaraan zich vele haken - in de vorm van afzonderlijke vraagstellingen - bevinden:

- Welke eisen kunnen we en willen we stellen aan leerlingen, die leerlingen die na hun "algemene vorming" de overstap naar een meer beroepsgerichte opleiding willen maken?
- In hoeverre kunnen we en willen we die eisen hard maken (minimumeisen) en in hoeverre als wensen beschouwen?

- Kunnen we en willen we een evenwichtige spreiding van onze eisen en verlangens tot stand brengen ten aanzien van kennis, vaardigheden en attitudes (kennen, kunnen en zijn) of stellen we vooral eisen aan de cognitieve ontwikkeling van de leerlingen?
- Wie dienen betrokken te zijn bij de formulering van deze eisen en verlangens?

In deze vragen is van vakken eigenlijk nog geen sprake geweest. Toch bestaat in ons voortgezet onderwijs een historisch gegroeide vakindeling en het onderwijs in deze vakken vervult zeker ook meer specifieke functies. Zo wordt een leerling, die later technische wetenschappen, een der natuurwetenschappen of medicijnen wil gaan studeren of een hoger technische opleiding wil volgen, bijna gedwongen om natuurkunde in zijn 'keuze'pakket op te nemen. Toch rijzen dan verdere vragen.

- Welke functies wil men afzonderlijke vakken laten vervullen ten aanzien van de realisatie van algemene eisen en wensen?

- Welke specifieke eisen en verlangens dienen met behulp van het onderwijs in bepaalde vakken of vakgebieden vervuld te worden?

U ziet, dat het stellen van vragen gemakkelijker is dan de beantwoording ervan. Toch meen ik, dat ook de vragensteller een functie heeft, die in ons huidige onderwijs misschien wat ondergewaardeerd wordt. Soms zijn we zo driftig bezig met het zoeken naar oplossingen, dat we ons onvoldoende realiseren wat eigenlijk de problemen zijn - of, anders geformuleerd, wie van bepaalde veranderingen het meest profiteren.

Intussen zit aan het principiële doordenken van problemen ook het gevaar, dat te weinig aandacht wordt geschonken aan de situatie van vandaag en morgen. Daarvoor bent U niet naar Woudschoten gekomen. Anderzijds kan zonder moeite worden aangetoond, dat in de huidige onderwijspraktijk dergelijke vragen of niet aan de orde komen, dan wel slechts impliciet worden beantwoord. Zo is een uiterst belangrijke vraag ten aanzien van het natuurkunde onderwijs bij mijn weten nimmer in ons land expliciet in discussie gekomen. Het is de vraag naar de functie ervan: vooral het doen verwerven van kennis en specifieke vaardigheden met het oogmerk deze te laten functioneren in bepaalde vervolgopleidingen, of in



de eerste plaats ontwikkeling van meer algemene vaardigheden en attitudes die een veel breder terrein van bruikbaarheid bezitten. Toch zou een antwoord op deze vraag kunnen helpen om enerzijds het onderwijs in het eigen vakgebied functioneler te maken, anderzijds een bijdrage kunnen leveren tot de aansluitingsproblematiek VWO-WO, HAVO-HBO en MAVO-MBO, die nu nogal eens versluierd dreigt te worden met oneigenlijke argumenten.

Antwoorden op de bovengestelde vraag konden in het verleden slechts indirect verkregen worden, namelijk door analyse van examens, eventueel aangevuld door analyse van lessen voorzover men namelijk van mening kon zijn, dat in deze lessen nog andere doelen worden nagestreefd dan in de examens werden getoetst. Het eerste is meermalen gedaan, het laatste - systematisch en gericht op natuurkundè-onderwijs - nimmer. Juist omdat er in de aanwijzingen voor het onderwijs eigenlijk slechts sprake is geweest van leerinhouden en niet of uiterst spaarzaam van gedragsdoelen, zou het niet op examen-training gerichte onderwijs sterk uiteen kunnen lopen.

### 1.3. Examendoelen en examenvragen

Sinds kort beschikken we over extra feitenmateriaal om er de vraag naar de functie van het natuurkunde-onderwijs mee te lijf te gaan: het rapport - examen-programma's van de Commissie Modernisering Leerplan Natuurkunde, dat zeer onlangs voor publikatie is vrijgegeven. Dit rapport bevat namelijk - naast de obligate leerstoflijsten - een duidelijke poging, te omschrijven welke handelingsstructuren nu eigenlijk van leerlingen op de verschillende eindniveaus (MAVO, HAVO, VWO) verwacht worden.

U zou zelf kunnen discussiëren over de vraag wáár CMLN op de as van de dipool algemeen - specifiek is terechtgekomen. U hebt daartoe de lijst van gedragsdoelen (bijlage, zie blz. 20) ter beschikking. Menige omschrijving zal op U - hoezeer ook gepoogd is een zodanige formulering te kiezen, dat eruit af te lezen is wat van kandidaten verwacht wordt - een weinig vertrouwde indruk maken.

Ik zal daarom aan de hand van een reeks voorbeelden van vragen proberen U een idee te geven in hoeverre achter deze formulering iets nieuws schuilgaat. Enkele opmerkingen over de gekozen voorbeelden - uit Bloom; Bloom, Hastings en Madaus; en uit Sandbergen - moet ik vooraf wel maken, opdat geen onnodige misverstanden ontstaan.

- alle gegeven voorbeelden zijn meerkeuzevragen. Dat houdt mijnerzijds geen voorkeur voor deze toetsingsvorm in, maar geeft alleen aan, dat buiten de ook U bekende bronnen uitsluitend dergelijke voorbeelden te vinden waren.
- de vragen zijn niet gemaakt om de door C.M.L.N. gekozen doelen (zie bijlage 1) te illustreren. De aansluiting is dan ook zeker niet perfect. Wel is ze hopelijk voldoende om een redelijke indruk van de vermoedelijke toekomstige examensituatie te geven.

De formulering van A1

' bij een beschrijving van een fysisch verschijnsel kunnen herkennen met welke begrippen of wetten het verschijnsel in verband gebracht kan worden'

kan wellicht al dadelijk leiden tot andere vraagstellingen dan bij de huidige examenpraktijk gebruikelijk is. Een voorbeeld hiervan is het volgende probleem:

Stel je voor dat je een waarnemingsspel speelt, waarbij je vijf punten krijgt voor het doen van de beste waarneming van verschillende verschijnselen. Je hebt een kleine hoeveelheid van een onbekende stof in een gesloten metalen kistje met een kleine opening in het deksel. Door aan de opening te ruiken, kun je een mufte geur in het kistje opmerken en deze geur is er nog steeds als je na een uur opnieuw ruikt.

Welke van de volgende zinnen geeft het beste aan, wat je op grond van deze waarneming over de stof weet:

- a. de stof is een vloeistof of een gas, maar geen vaste stof;
- b. de stof is een vloeistof of iets dat gemakkelijk in vloeistof overgaat;
- c. het is een vaste stof of een vloeistof, maar geen gas;
- d. het is een gas of iets dat gemakkelijk een gas vormt.

De doelstellingen B1 en B2

'Indien de uitkomsten van een experiment in tabelvorm zijn gegeven, met behulp hiervan een grafiek kunnen tekenen'

'uit grafieken gegevens kunnen aflezen'

worden - als één doelstelling - als expliciet in het "Definitieve Programma" genoemd. Ze worden sinds 1968 regelmatig getoetst, ook op mavo-niveau, zodat hier geen specifieke voorbeelden te hoeven worden gegeven. Overigens kan men een B2-vraag vaker aantreffen dan een vraag van type B1.

Ook B3

'de waarde van een grootte kunnen berekenen met behulp van een aan de leerling bekende relatie'

is een bekende, zelfs een oude bekende, omdat veel vraagstukken uit het verleden worden gezien als een toetsing van dit doel, al of niet in zinvolle combinatie met A1. Natuurlijk kan dit doel niet worden gemist als we onze leerlingen in aanraking willen brengen met de kwantitatieve aspecten van de natuurkunde.

Voorbeelden zijn overbekend en talrijk.

Dit geldt niet voor C1 en C2

'de afloop van een variatie op een bekend experiment kunnen voorspellen met behulp van bekende begrippen en wetten'

en

'kunnen nagaan of de uitkomsten van een variatie op een bekend experiment in overeenstemming zijn met bekende wetten'

Een zeer eenvoudig voorbeeld van C1 is het volgende

Annie speelde met een bellenblaaspijpje. Toen de bel even groot was als op de tekening, nam ze de pijp uit de mond. Wat gebeurde er daarna met de bel?

- a. Hij werd een tijdje groter en bleef daarna even groot.
- b. Hij werd een tijdje kleiner en bleef daarna even klein.
- c. Hij werd steeds kleiner en verdween tenslotte in de pijp.
- d. Hij bleef op de pijpekop zitten zonder van grootte te veranderen.
- e. Hij werd steeds groter en barstte tenslotte uit elkaar.

De volgende vraag is een moeilijker voorbeeld van C1, zij het dat hier ook nog een oordeel wordt gevraagd.

Twintig kerstboomlampjes zijn in serie geschakeld. Het nadeel is dat - zodra een lampje lostrilt of doorbrandt - alle lampjes uitgaan en men niet weet welk lampje hiervan de oorzaak is.

Nu stelt iemand voor tussen de aansluitingspunten van elk lampje een draadje met een weerstand te schakelen. Als nu een lampje doorbrandt loopt de stroom door het draadje, zodat alle lampjes blijven branden, behalve het kapotte.

Ben je het eens met dit voorstel? Motiveer je antwoord.

Iemand anders stelt voor de twintig lampjes twee aan twee parallel te schakelen.

Welke voor- en nadelen heeft die oplossing?

Van C2 vond ik geen volkomen aansluitend voorbeeld. De eerste vraag uit het volgende probleem komt wel in de buurt. Het eist echter niet zozeer het nagaan of de uitkomsten van de proeven in overeenstemming zijn met de bekende relatie tussen ladingstekens en krachtrichting - al zou in principe de uitkomst van de proeven in tegenspraak met deze relatie kunnen zijn - maar meer hoe de plussen en minnen over de lichamen verdeeld moeten worden, opdat het resultaat in overeenstemming is met bekende wetten. Dit komt vaker voor.

Volgens afspraak is de lading die een glazen staaf verkrijgt door deze met een wollen lap te wrijven positief (+). Een leerling voert met vijf lichamen (I, II, III, IV, V) enkele proeven uit en verkrijgt hieruit de volgende gegevens:

I stoot een met een wollen lap gewreven glasstaaf af, maar trekt IV aan.

II trekt V aan, maar stoot III af.

IV stoot II af.

De kracht tussen I en II is een eenheid als hun onderlinge afstand 4 cm bedraagt. De kracht tussen II en III is 8 eenheden, als hun onderlinge afstand 1 cm bedraagt.

1. Welk van de volgende conclusies kan de leerling betreffende het ladingsteken van de lichamen terecht trekken?
  - a. I, V positief; II, III, IV negatief
  - b. I, II positief; III, IV, V negatief
  - c. II, III positief; I, IV, V negatief
  - d. I, II, IV positief; III, V negatief
  - e. geen van deze conclusies is correct
2. Hoe groot zal de kracht tussen I en II zijn als hun onderlinge afstand 1 cm bedraagt?
  - a.  $1/4$  eenheid
  - b. 4 eenheden
  - c. 8 eenheden
  - d. 16 eenheden
  - e. geen van deze antwoorden is correct
3. Wat is de verhouding van de ladingen op I en II?
  - a. 4 : 1
  - b. 2 : 1
  - c. 1 : 1
  - d. 1 : 2
  - e. geen van deze antwoorden is correct

Men zou kunnen stellen, dat de doelstellingenlijst hier aanvulling verdient met:

'uit de resultaten van een variatie op een bekend experiment conclusies kunnen trekken, uitgaande van de veronderstelling dat deze resultaten in overeenstemming zijn met bekende wetten'.

In feite gaat de hier bedoelde activiteit in de praktijk vaak aan de in C2 genoemde vooraf. Pas als dit niet lukt, kan men tot de conclusie geraken, dat de uitkomsten blijkbaar niet in overeenstemming te brengen zijn met bekende wetten. Ook het volgende probleem is in feite een toetsingsvoorbeeld van de zoëven gegeven tussendoelstelling C1'.

De figuur (niet weergegeven) stelt een doos voor met 4 aansluitingspunten P, Q, R en S. Men doet de volgende waarnemingen:

1. tussen P en Q bestaat een zekere weerstand
2. de weerstand tussen P en R is tweemaal zo groot als de weerstand tussen P en Q
3. tussen Q en S is de weerstand verwaarloosbaar

Welke van de getekende circuits komt overeen met deze waarnemingen, aangenomen dat de getekende weerstanden even groot zijn?

De onderdelen b en c van het probleem met de vijf geladen geleiders toetsen doelstelling D1:

'aan de hand van bekende wetten kunnen nagaan welke van enkele gegevens voorspellingen over de afloop van een tevoren niet bekend experiment in overeenstemming zijn met de werkelijke afloop ervan'

al zijn daarbij toch twee opmerkingen te maken:

- deze toetsing gaat uit boven het maveniveau
- als de wet van Coulomb bekend is, kan waarschijnlijk in dit geval nauwelijks worden gesproken van een tevoren niet bekend experiment

Het zal binnen het bestek van deze voordracht niet mogelijk zijn de door C.M.L.N. ook voor HAVO en VWO geformuleerde doelen op de voet te blijven volgen.

Omdat B4

'gegevens uit enkele grafieken met elkaar in verband kunnen brengen'

zowel als B7

'gegevens, die voor het oplossen van een probleem nodig zijn, kunnen selecteren en, zonodig door berekening, een probleem ermee kunnen oplossen'

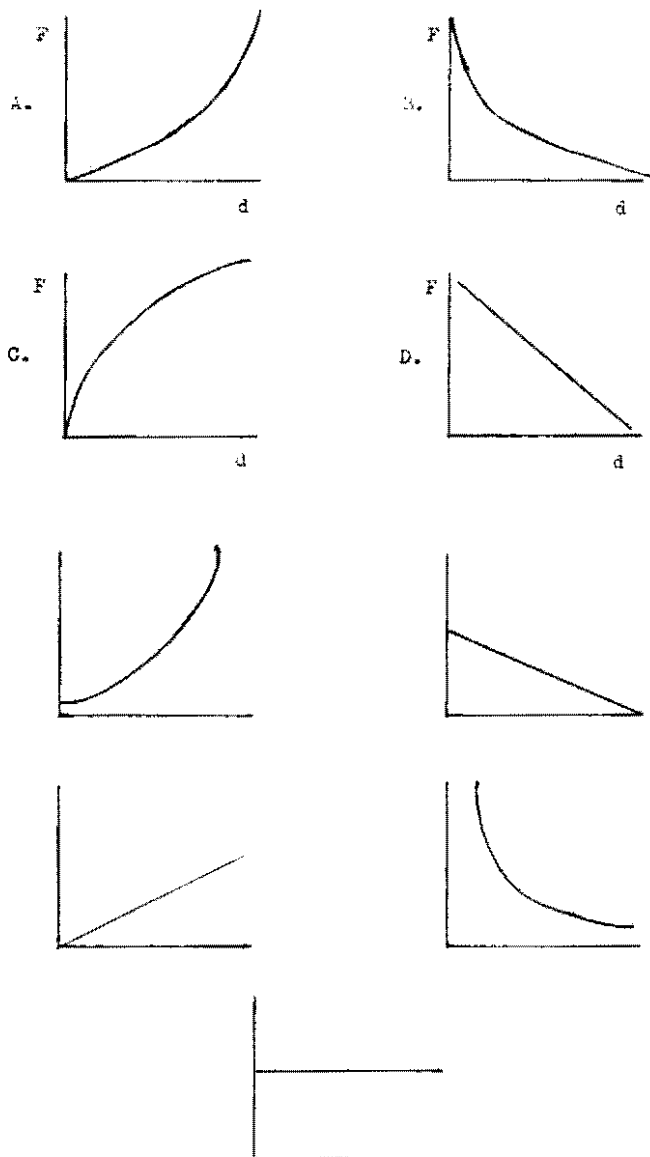
niet sterk van bekende examenproblemen afwijken, volsta ik ermee te wijzen op het tabellenboekje en op p.V, p.T en V.T. diagrammen en ga over op enkele voorbeelden bij B5

'een in wiskundige vorm gegeven relatie kunnen overbrengen in een grafiek'

De wet van de zwaartekracht van Isaac Newton wordt algebraïsch aldus voorgesteld:

$$F = G \frac{Mm}{d^2}$$

Hierin is F de kracht, M en m = twee massa's, G is een constante en d stelt de afstand tussen de twee massa's voor. Aangenomen dat M en m, evenals G, constant blijven, welke van de volgende grafieken toont dan hoe F verandert als de afstand d tussen de massa's varieert?



Van B6

'in niet te gecompliceerde gevallen een door een grafiek voorgestelde relatie kunnen vertalen in formulevorm'

vond ik in de door mij geraadpleegde bronnen althans voor natuurkunde

geen geslaagd voorbeeld, al zou uit het volgende voorbeeld door vertaling van de tabelgegevens in grafieken gemakkelijk een dergelijk voorbeeld geconstrueerd kunnen worden:

Bij een proef wordt een bal door het ontspannen van een schroefveer langs een ruw horizontaal vlak geschoten. Het verband tussen de afstand waarover de veer aan het begin is samengedrukt (c), de afstand waarover de bal rolt (s) en de tijdsduur gedurende welke de bal in beweging blijft (t), is vastgelegd in onderstaande tabel:

c (cm)	0,50	1,00	2,00	4,00
t (sec)	0,25	0,50	1,00	2,00
s (cm)	5,0	20	180	320

Welk verband bestaat er tussen de samendrukking (c) en de afstand (s)? k is een constante?

- a.  $s = kc$                       d.  $s = \frac{k}{c}$   
b.  $s = kc^2$   
c.  $s^2 = kc$                       e.  $s = \frac{k}{c^2}$

In de formulering van C.M.L.N. is vrijwel steeds sprake van experimenten en niet van waarnemingsresultaten. Misschien wijst dit op een aardse georiënteerdheid.

In de astronomie worden aan sterren, planeten en andere objecten zeer veel waarnemingen gedaan, maar zelden worden er experimenten mee verricht. Ik wilde mijn voorbeeldenreeks besluiten met twee voorbeelden, die uit de astronomie afkomstig zijn en goede illustraties van enkele doelstellingen zijn, als het 'experiment' zeer ruim wordt geïnterpreteerd.

De volgende waarnemingsresultaten omtrent het zonnestelsel zijn gegeven:

1. elke dag gaat de zon in het oosten op en in het westen onder
  2. 's nachts lijken de sterren rond de Poolster te draaien
  3. zonsverduisteringen treden met regelmatige tussenpozen op
- Welke van deze waarnemingsresultaten kan niet worden verklaard met behulp van een model van het zonnestelsel, waarbij de zon rond de aarde draait:

- a. alleen 1 en 2  
b. alleen 2 en 3  
c. alleen 1 en 3  
d. 1, 2 en 3  
e. geen enkele

Dit is een voorbeeld van het zoeken naar logische samenhang tussen een waarnemingsresultaat en een bekende theorie. Iets dergelijks wordt vermoedelijk bedoeld bij E3:

'kunnen nagaan of een veronderstelling logisch uit een bekende theorie voortvloeit'

terwijl de aktiviteit die bij het volgende voorbeeld wordt verlangd:

U hebt enige kennis verworven over de aarde en haar bewegingen zoals deze werkelijk bestaan. In deze opgave moet U de gevolgen omschrijven van enige geheel denkbeeldige omstandigheden. U moet de items van de hieronder gegeven hoofdletters voorzien en wel met

- A. als het item waar zou zijn, indien de aardas loodrecht op de evenaar stond in plaats van schuin
- B. als het item waar zou zijn, indien de baan van de aarde een cirkel was en niet een ellips
- C. als het item waar zou zijn, indien de aarde naar het westen om haar as draaide in plaats van naar het oosten
- D. als het item waar zou zijn, indien de aarde maar de helft van haar doorsnede had, maar haar normale massa behield
- E. als het item waar zou zijn, indien de aarde geen maan had

Wij nemen aan, dat slechts één van de imaginaire omstandigheden tegelijk plaats grijpt.

- 1. alle zonedagen zouden van gelijke lengte zijn
- 2. de bekende voorwerpen zouden viermaal zo zwaar zijn als thans
- 3. de hemelequator en de ecliptica zouden identiek zijn
- 4. de zon zou in het oosten ondergaan
- 5. men zou een andere poolster moeten kiezen
- 6. de aantrekkingskracht zou viermaal zo groot zijn
- 7. de snelheid van de aarde in haar baan om de zon zou gedurende het jaar steeds hetzelfde zijn
- 8. we zouden veel minder over de eigenschappen van de zon weten
- 9. dag en nacht zouden het gehele jaar door van gelijke lengte zijn op alle breedten

vrij goed overeenkomt met de aktiviteit, die wordt verlangd bij E5:

'kunnen nagaan of de uitkomst van een toetsingsexperiment een theorie ondersteunt of (gedeeltelijk) weerlegt'

#### 1.4. Evaluatievorm en leerresultaat

Uit de ervaringen, opgedaan in de Commissie Schoolonderzoek van C.M.L.N. (6, 7, 8) is overigens al gebleken, dat dit schoolonderzoek, gekoppeld aan demonstratieproeven of praktikum, veelal betere gelegenheden biedt dan een schriftelijk examen, om complexe doelstellingen te toetsen. Ik zal op deze schoolonderzoeken hier niet verder ingaan, maar verwijs U gaarne naar de voordracht van Steller en naar enkele publikaties.

Het feit, dat met het schoolonderzoek betrekkelijk gemakkelijk toetsings-situaties konden worden opgebouwd, waarin veel meer aspecten van de natuur-wetenschappelijke onderzoeksmethode aan de orde konden komen dan in vele jaren examenpraktijk is gelukt, wijst op een bijzondere geschiktheid van het middel.

Die geschiktheid heeft overigens meer dimensies:



- de vorm waarin de evaluatie plaatsvindt sluit veel dichter aan bij de gewenste handelingsstructuren zoals deze bij natuurwetenschappelijk onderzoek optreden
- er is meer zekerheid of een verschijnsel, een experiment, een relatie of wet, een onderdeel van een theorie aan de betreffende leerlingen volkomen vertrouwd, enigszins bekend of volkomen onbekend is
- een zekere sturing of aanvulling tijdens het evaluatiegebeuren is mogelijk. Bij examens moet alles voorzien zijn, wat een gekunstelde stijl, maar ook gekunstelde situaties uitlokt.

De laatste nadelen van examensituaties zijn iedere leraar uit eigen ervaring bekend.

Op het eerste aspect wil ik tot slot van dit betoog nog ingaan.

In de literatuur over toetsen heeft het begrip (inhouds) validiteit ruime aandacht gekregen. Ruw geformuleerd komt het erop neer, dat een toets een grote validiteit heeft, indien deze meet, wat men meten wil. Bij objectieve studietoetsen is aan dit aspect meestal in ruime mate aandacht besteed. Indirecte toetsing, d.w.z. toetsing of een bepaalde doelstelling is bereikt in de veronderstelling, dat wie die doelstelling bereikt heeft, een geheel andere, minder gemakkelijk meetbare ook wel bereikt zal hebben, komt dan weinig voor.

Objectieve toetsen roepen echter een geheel ander gevaar op: namelijk dat de vorm waarin de toetsing plaatsvindt een zodanige terugkoppeling op het onderwijs geeft, dat de gewenste handelingsstructuren in het gedrang komen.

Ik zal U daarvan een voorbeeld uit een geheel ander vakgebied geven, namelijk uit het moedertaalonderwijs. Het is U zeker bekend, dat tot de geliefde evaluatiemiddelen het opstel behoorde. Met dergelijke opstellen kunnen vele aspecten van produktief-schriftelijke taalvaardigheid worden getoetst, echter slechts onder zeer bijzondere omstandigheden - die in de praktijk nooit worden benaderd - redelijk objectief. Ik ga nu voorbij aan de uiterst relevante vraag welke aspecten van produktief-schriftelijk taalgebruik meer nadruk in ons onderwijs zouden moeten krijgen, maar beperk me tot het vormaspect waarover ik zo-even sprak. Een Nederlands onderzoeker (9) heeft namelijk aangetoond - en ik heb geen reden om aan zijn bewijsvoering te twijfelen - dat met objectieve studietoetsen een groot aantal aspecten van produktief - schriftelijk taalgebruik even valide kunnen worden getoetst.

Ik laat U nu twee fragmenten van deze toetsen zien. De vragen bij het eerste fragment spreken voor zichzelf, bij het tweede fragment is het de bedoeling dat de kandidaat de onderstreepte woorden en zinsdelen in de marge verbetert.

Een reisje naar de maan in het jaar 2000

1. Vannacht droomde ik dat ik in het jaar 2000 leefde en
2. dat onze meester tegen de klas zei: "Volgende week gaan we
3. een reisje naar de maan maken. Jullie mogen zelf de raket
4. besturen!
5. Ik wou pilote zijn. De jongens hadden daar eerst niet
6. veel zin in en vroegen: "Ben je wel slim genoeg om al die
7. knoppen te bedienen?" Toen ik zei dat ik de huiscomputer
8. wel aardig kon bedienen, hadden ze geen bezwaar meer. Op
9. woensdagavond werden we door een autobus opgehaald een week
10. later. We waren na tien minuten op de lanceerbasis. Voordat
11. we instapten, wensten onze ouders ons veel plezier. In de be-
12. dieningsruimte was zorgvuldig nagegaan of alles werkte. De
13. raket was binnen net een paleis, zo prachtig was hij. Het
14. duurde nog maar een half uur of we zouden starten. Dat was
15. een vreselijk spannende tijd.
16. Vanmiddag had ik verstoppertje gespeeld met mijn broer-
17. tje en zusje. Ik moest hen zoeken, maar dat viel niet mee. Ze
18. hadden zich goed verstopt. Mijn broertje was in een boom ge-
19. klommen en mijn zusje was in een lege ton gaan zitten. Na lang
20. zoeken vond ik hen.
21. Vijf, vier, drie, twee .... een .... nul. Daar gingen we!
22. Het ging sneller en sneller. We waren helemaal achteroverge-
23. drukt.

Wat kun je het beste doen met: niet veel zin in (regel 5/6)?

- A. zo laten staan
- B. vervangen door: geen aardigheid in
- C. vervangen door: niet veel vertrouwen in
- D. vervangen door: niet wat wantrouwen in

Wat kun je het beste doen met: wel aardig (regel 8)?

- A. zo laten staan
- B. vervangen door: aardig knap
- C. vervangen door: bijzonder
- D. vervangen door: nogal

Wat kun je het beste doen met: Op woensdagavond ....een week later (regel 8/10)?

- A. zo laten staan
- B. vervangen door: Na een week, op woensdagavond
- C. vervangen door: Op een week na, op woensdagavond
- D. vervangen door: Over een week, op woensdagavond

Wat kun je het beste doen met: sneller en sneller (regel 22)?

- A. zo laten staan
- B. vervangen door: hoe hoger hoe hoger
- C. vervangen door: hoe hoger hoe sneller
- D. vervangen door: hoe sneller hoe sneller

Verse broodjes en verse vis .....

Een haas en een vos trokken samen door het land. Het was een strenge winter. De velden waren besneeuwd en zelfs de muisjes bleven onder hun holletjes. "Afschuwelijk, wat een weer", zei de vos en de haas tegen elkaar. "Mijn maag knort van de honger", zei de vos, "want ik zou best iets willen eten". Ook de haas was erg hongerig. Hongerig en en verdrietig liepen ze verder. Toen opeens zagen zij een meisje naderen. Uit het mandje steeg de geur van heerlijke verse broodjes op. Ze droeg een mandje aan haar arm.

De vraag is nu, waar vervanging van opstellen door dit soort toetsen toe zou leiden. Wellicht tot onderwijs dat goede correctors, maar slechte verslaggevers oplevert. De kans dat ook in het onderwijs de correctie-arbeid het zelf vorm geven aan gedachten gaat overheersen als dit de 'eindtermen' worden, lijkt me in elk geval niet denkbeeldig.

Zo is ook de vraag of onze examens, hoezeer daarbij ook gepoogd zou worden om aspecten van wetenschappelijk denken en onderzoeken tot inzet van het evaluatieproces te maken, in feite niet afleiden van de manier waarin dergelijk denken en onderzoeken in de werkelijkheid moet plaatsvinden.

Aan de andere kant betekent een toetsing van vaardigheden en attitudes met betrekking tot natuurwetenschappelijk onderzoek in een vorm die aansluit bij het handelen en streven in de praktijk zeker een verlies aan objectiviteit in de door psychologen gedefiniëerde betekenis. Men kan resultaten van het door leerlingen uitgevoerd onderzoek vaststellen, men kan het proces volgen en beoordelen, maar men zal er veel moeite mee hebben de prestaties der leerlingen in te schalen, laat staan uit te maken wat 'voldoende' en wat 'onvoldoende' zijn.

Is dat een bezwaar? Ja, voorzover een gediplomeerde deficienties vertoont, die hij of zij bij een vervolgopleiding niet met relatief weinig inspanning in relatief korte tijd wegwerken kan. Vermoedelijk zouden dergelijke deficienties veel kans maken gesignaleerd en tamelijk gemotiveerd bestreden te worden in een onderwijsleerproces, dat gericht is op of minstens gelegenheid geeft tot eigen onderzoek. Maar utopie is soms een slechte raadgever. Laten we slechts constateren dat het bezit van werkelijk noodzakelijke basisvaardigheden niet persé aan het eind van de leerweg behoeft te worden vastgesteld. Het kan ook eerder.

Daarom pleit ik vandaag voor:

- een of meer onderzoeken naar het bezit van zorgvuldig geselecteerde basisvaardigheden, ruim voor het einde van de leerweg, desnoods in de vorm van meerkeuzetoetsen
- een onderwijsleerproces, sterk gericht op actief natuurwetenschappelijk denken en handelen van de leerlingen, met een semi-continue eva-

luatie van proces en produkt

Ik beseft dat ik daarmee - zij het impliciet - toch een aantal keuzen heb gedaan ten aanzien van de in het eerste deel van mijn voordracht aangesneden problematiek. Liever dan deze te expliciteren en U door de eenrichtingscommunicatie iets op te dringen, wil ik U prikkelen zelf een keuze te maken.

#### Literatuur:

1. Commissie Modernisering Leerplan Natuurkunde. Rapport 1974.
2. B.S. Bloom (Ed.) Taxonomie van een aantal in het onderwijs en de vorming gestelde doelen: I Het cognitieve gebied (vertaling), Rotterdam 1971.
3. B.S. Bloom, V.Th. Hastings, G.F. Madaus: Handbook of Formative and Summative Evaluation of Student Learning, New York 1971.
4. S. Sandbergen: Nederlandse Resultaten in Tekstbegrip en Natuurwetenschappen, Deel II, R.I.T.P., 1974.
5. W. Zandstra, Oefenvragen Natuurkunde, Purmerend, 1973.
6. Commissie Schoolonderzoek C.M.L.N., verslag 1971-1972.
7. Commissie Schoolonderzoek C.M.L.N., verslag 1972-1973.
8. Commissie Schoolonderzoek C.M.L.N., Afstudeerproject T.H.Eindhoven 1973
9. H. Wesdorp: Het meten van de produktief-schriftelijke taalvaardigheid Muusses, 1974.

Uit: Commissie Modernisering Leerplan Natuurkunde, rapport 1974, deel VII

VOORSTELLEN BETREFFENDE DE EXAMENPROGRAMMA'S VOOR MAVO-3, MAVO-4, HAVO EN VWO

*Voorwoord*

De in dit deelrapport vermelde onderwerpen (feiten, terminologieën, begrippen, afspraken, wetten, theorieën en toepassingen) dienen te behoren tot het kennispakket van de examenkandidaten. Voor nadere detaillering, diepgang en motivering verwijzen wij naar de betreffende deelrapporten voor de onderscheiden schooltypen.

In de examenopgaven zal in het algemeen het hanteren van eenheden uit het praktische stelsel (S.I.-eenheden) geëist worden.

Eventueel afwijkende eenheden zullen in de opgaven verklaard moeten worden.

Herhaaldelijk is achter een onderwerp tussen haakjes de toevoeging 'kwalitatief' geplaatst. Het is daarbij de bedoeling niet een volledige wiskundige beschrijving te geven van de verschijnselen, maar wel om de onderlinge verbanden duidelijk te laten uitkomen - dus of een grootheid groter/kleiner of meer/minder wordt bij een verandering van een ermee samenhangende grootheid. Als voorbeeld kan hier genoemd worden de buiging van licht aan een spleet, waar wel het verband: kleinere spleetwijdte → grotere afbuighoek gegeven moet worden (liefst uit experimenten), maar geen berekening van het kwantitatieve verband. Vaak zal bij de uitleg gebruik gemaakt kunnen worden van een schets of van een grafiek ter verduidelijking van de bespreking.

Aan de hand van het kennispakket zal in het examen worden nagegaan in hoeverre de kandidaten vaardigheid vertonen in het gebruik van deze kennis.

van MAVO-kandidaten wordt verwacht, dat zij:

- A.1. bij een beschrijving van een fysisch verschijnsel kunnen herkennen met welke begrippen of wetten het verschijnsel in verband gebracht kan worden;
- B.1. indien de uitkomsten van een experiment in tabelvorm zijn gegeven, met behulp hiervan een grafiek kunnen tekenen;
  2. uit grafieken gegevens kunnen aflezen;
  3. de waarde van een grootheid kunnen berekenen met behulp van een aan de leerlingen bekende relatie;
- C.1. de afloop van een variatie op een bekend experiment kunnen voorspellen met behulp van bekende begrippen en wetten;
  2. kunnen nagaan of de uitkomsten van een variatie op een bekend experiment in overeenstemming zijn met bekende wetten;
  3. een bekend verschijnsel of experiment met behulp van een bekende theorie kunnen verklaren;

- D.1. aan de hand van bekende wetten kunnen nagaan welke van enkele gegeven voorspellingen over de afloop van een tevoren niet bekend experiment in overeenstemming zal zijn met de werkelijke afloop ervan.

Van HAVO-kandidaten wordt bovendien verwacht, dat zij:

- B.4. gegevens uit enkele grafieken met elkaar in verband kunnen brengen;
5. een in wiskundevorm gegeven relatie kunnen overbrengen in een grafiek;
  6. in niet te gecompliceerde gevallen een door een grafiek voorgestelde relatie kunnen vertalen in formulevorm;
  7. gegevens die voor het oplossen van een probleem nodig zijn kunnen selecteren en, zo nodig door berekening, een probleem ermee kunnen oplossen;
  8. kunnen interpoleren en extrapoleren en in het laatste geval inzien, wanneer zij daarbij extra voorzichtig moeten zijn;
  9. bij berekeningen waarbij gebruik wordt gemaakt van meetwaarden, de grootte-orde van de onzekerheid in de uitkomst kunnen schatten;
- D.2. de afloop van een nieuw experiment kunnen voorspellen, gebruik makend van bekende begrippen en wetten;
3. kunnen aangeven hoe een gegeven veronderstelling op een binnen de leerstof gelegen terrein experimenteel getoetst zou kunnen worden;
  4. een veronderstelling op een binnen de leerstof gelegen terrein experimenteel kunnen verifiëren (deze eis moet in het schoolonderzoek worden getoetst);
  5. aan de hand van bekende wetten nagaan of een gegeven voorspelling over de afloop van een tevoren niet bekend experiment al dan niet in overeenstemming is met de werkelijke afloop ervan.

Van VWO-kandidaten wordt voorts verwacht, dat zij:

- E.1. de juistheid van onbekende formules kunnen toetsen met behulp van een dimensieberekening;
2. kunnen nagaan of door een meetexperiment op grond van de daarbij bereikte nauwkeurigheid een gegeven relatie wordt bevestigd of weersproken;
  3. kunnen nagaan of een veronderstelling logisch uit een bekende theorie voortvloeit;
  4. de afloop van een nieuw experiment met behulp van een bekende theorie kunnen voorspellen;
  5. kunnen nagaan of de uitkomst van een toetsingsexperiment een theorie ondersteunt of (gedeeltelijk) weerlegt;
  6. kleine veranderingen in een bekende theorie kunnen formuleren die de uitkomst van een toetsingsexperiment verklaren;
  7. een eenvoudige redenering kunnen opbouwen op een onbekend terrein der natuurkunde indien de structuur daarvan analoog is aan die van een bekend terrein, onder uitdrukkelijke verwijzing naar die analogie.

2. Voordracht Steller: 'De rol van het schoolonderzoek in de summatieve evaluatie'

"Men kan de aanwijzing van onderwerpen in een lijst vergelijken met de aanwijzing van een geschikt geacht speelveld voor een sportwedstrijd", C.M.L.N., Discussienota m.a.v.o., pag.8.

De vergelijking is zeer ongelukkig gekozen of men wil hier bewust het dwingende karakter van de leerstoflijst gecamoufleerd presenteren. Bij sportwedstrijden pleegt men nl. het sportveld dwingend voor te schrijven. Rusland is gediskwalificeerd bij de voorronden voor het wereldkampioenschap voetballen omdat de spelers weigerden te spelen in het voorgeschreven stadion (concentratiekamp) in Santiago. Bovendien heb ik er bezwaar tegen de leerstoflijst direct in verband te brengen met een sportwedstrijd i.p.v. sportbeoefening. Laat mij nu aannemen dat de onderwerpenlijst niet meer is dan "de aanwijzing van een geschikt geacht speelveld". Nu kiest een mavoreraar een ander speelveld dat hem geschikter lijkt. Dat speelveld is kleiner maar biedt juist daardoor gelegenheid alle spelers een beter overzicht te bieden en daardoor bepaalde aspecten van het spel beter, intensiever te beoefenen (let wel: beoefenen). Bovendien ligt dat speelveld in de buurt waar de spelers wonen, het is dus makkelijker te bereiken. M.a.w. de mavoreraar kiest onderwerpen die binnen de dagelijkse belevingswereld van de leerlingen liggen. Helaas zullen op dat kleinere veld andere aspecten niet tot hun recht komen. Hoe moet dat nu bij de wedstrijd die beslist over promotie of degradatie? Op welk veld wordt die gespeeld?

M.i. op twee velden, een thuiswedstrijd en een uitwedstrijd; het schoolonderzoek en het centraal schriftelijk examen. De vergelijking gaat nog verder. In de praktijk blijkt dat thuiswedstrijden gemiddeld genomen beter aflopen dan uitwedstrijden. U kunt dit ongetwijfeld zelf wel vertalen. De consequenties voor leerstoflijst en examenpraktijk wil ik graag bij een andere gelegenheid uitdiepen.

Laten wij ons bepalen tot ons onderwerp "het schoolonderzoek". En laten wij ons niet bezighouden met de uitwassen die naar men zegt optreden, niet ingaan op het misbruiken van het schoolonderzoek doordat de leraar bewust als "home referee" gaat optreden; ook niet met de maatregelen die men daartegen zou kunnen en moeten nemen, maar alleen met de voordelen die het schoolonderzoek kan bieden voor een juiste en rechtvaardige beoordeling van het kennen en kunnen, van begrip en inzicht der leerlingen.

Waarin ligt het eigen karakter van het schoolonderzoek waardoor de kwaliteiten van de leerling in een ander licht worden geplaatst en uit een ander gezichtspunt beoordeeld? Niet milder, niet strenger maar anders. Tot de volkomen legitieme voordelen van het schoolonderzoek behoren:

1. De vragen worden gesteld op de bekende en vertrouwde wijze waarop de docent dat altijd in z'n proefwerken deed en doet.
2. De leerling weet hoe lang een korte toelichting moet zijn, hoe gedetailleerd een redenering of een berekening moet worden opgeschreven.
3. De leerling weet hoe ruw een schets mag zijn en hoe precies een diagram moet worden getekend.
4. De stof is beperkt tot enkele gebieden van de gehele leerstof en de leerling weet of kan weten waar de leraar de accenten heeft gelegd, wat deze belangrijk vond en waar hij dus waarschijnlijk over zal vragen.

Tot zover een aantal, ik herhaal het, volkomen legitieme voordelen, die ongetwijfeld zullen tenderen naar een hoger cijfer, maar ook zullen tenderen naar een rechtvaardiger cijfer doordat er minder kans is op verkeerd begrijpen van de vraag; te summier of juist te uitgebreid beantwoorden van de vraag, enz. De tendens naar een juister en rechtvaardiger beoordeling en niet noodzakelijkerwijze naar een hoger cijfer, kan aanzienlijk vergroot worden als de leraar in de onderwerpen die hij in de les speciaal heeft uitgediept bewust dieper graaft met zijn vragen. Ook voor het schoolonderzoek moet gelden dat de hoge cijfers alleen verdiend kunnen worden door het goed beantwoorden van vragen die naast kennen en kunnen ook inzicht in en wendbaarheid van die kennis en kunde toetsen.



De milde leraar zal steeds bij het maken van de opgaven en het corrigeren van het werk voor ogen moeten houden dat een schoolonderzoek, zelfs als het toevallig in december valt, geen Sinterklaasfeest is, geen "ophaaldienst" zoals Inspecteur van Dam bepaalde wijzen van mondeling examineren placht te noemen. De verleiding in ongepaste mildheid te vervallen is groot nu het gemiddelde eindexamen kan beslissen over directe toelating of loting voor bepaalde studierichtingen.

De strenge leraar die de normen van het centrale werk veel te slap vindt moet niet het schoolonderzoek misbruiken om de duimschroeven nog eens extra aan te draaien.

Het schoolonderzoek is immers niet bedoeld om een cijfer te leveren dat à priori een ontkenning of à priori een bevestiging van het centraal schriftelijk cijfer inhoudt. Het schoolonderzoek is bedoeld om samen met het centraal schriftelijk examen een beter gefundeerd cijfer te leveren dan met één van beide afzonderlijk mogelijk is.

De opgaven voor het schoolonderzoek zullen enerzijds noodzakelijkerwijze veel overeenkomst vertonen met de opgaven voor het centraal schriftelijk werk omdat beide bedoeld zijn om natuurkundige kennis en kunde, begrip en inzicht te toetsen op het niveau van het verlangde diploma. Even noodzakelijkerwijze zullen ze verschil moeten vertonen omdat ze bedoeld zijn om de zojuist genoemde fysische kwaliteiten uit een ander gezichtspunt en/of bij andere belichting te beoordelen.

Hierna volgen een paar voorbeelden van verschillen die m.i. zullen optreden.

1. In de nabije toekomst zullen de leerlingen op het centraal schriftelijk examen veel formules in het tabellenboek kunnen opzoeken. Het paraat zijn van de formules levert wel tijdwinst maar is niet persé nodig. Wel is nodig het herkennen van het geleerde in de vraag, het kunnen relateren van het probleem aan het juiste gebied uit de vele gebieden die aangewezen zijn. Wie een goed overzicht heeft over de hele stof is ongetwijfeld in het voordeel, ondanks eventuele manco's in detailkennis.

Bij een schoolonderzoek over een min of meer beperkt stuk stof b.v. trillingen en golven weet de leerling dat de vragen gaan over trillingen en golven.

Als hij dat deel van z'n sporensysteem gereactiveerd heeft, en dat heeft hij bij het repeteren de avond tevoren ongetwijfeld gedaan, zal hij direct beschikken over het systeem dat hij nodig heeft en niet gehinderd worden door het gebrek aan overzicht over alle stof. Maar gebrek aan kennis van definitie's en formules zal nu wel hinderlijk zijn omdat de leraar nu wèl eist dat deze paraat zijn.

2. Een leraar kan niet zoveel tijd en moeite aan de formulering en de redactie van de opgaven besteden als de samenstellers van de opgaven voor het centraal schriftelijk examen. Dat is gelukkig ook niet nodig omdat leraar en klas elkaar kennen, waardoor code's en gewoonten, al of niet expliciet geformuleerd, zijn ontstaan, b.v. touwtjes zijn massaloos en verbindingsdraden weerstandloos.

Hierdoor wordt de stijl van de opgaven wat lossier, huiselijk en vertrouwd, terwijl de opgaven voor het schriftelijk werk noodgedwongen veel scherper moeten worden geformuleerd.

Als tijdens een schoolonderzoek blijkt dat een gegeven of vraag voor tweeërlei uitleg vatbaar is dan kan dit makkelijk rechtgezet worden. Bij het schriftelijk examen ontbreekt die mogelijkheid.

3. De leraar, die precies weet hoe uitgebreid en hoe diepgaand hij een onderwerp heeft behandeld, kan als regel vrij goed vooraf weten of een vraag makkelijk is te beantwoorden omdat het antwoord op reproductie berust of dat de vraag moeilijk is omdat het probleem in die vorm nooit is behandeld. Bij het centrale werk zal men niet altijd kunnen vermijden dat het gebruik van een bepaald boek een voor- of nadeel blijkt te zijn.

Op de eerste Woudschoten conferentie in 1966 hebben Zandstra en steller dezes al betoogd dat het examen niet alleen dient te zijn aangepast aan de behandelde stof, maar ook aan de wijze van presentatie van de leerstof. Uit het voorgaande blijkt dat het schoolonderzoek makkelijker zal kunnen voldoen aan de eis van aanpassing aan de behandelde leerstof dan het centrale werk.

Het belangrijkste voordeel van het schoolonderzoek ligt in het voldoen aan de tweede eis, de aanpassing aan de wijze van presentatie.

Wie meent dat het practicum niet meer is dan één van de vele didactische hulpmiddelen die ons ten dienste staan, zal aan de relatie 'examen - presentatie van de leerstof' niet zwaar tillen. "Als op het examen maar blijkt dat ze het kennen en kunnen en begrepen hebben dan is het niet belangrijk hoe ze het geleerd hebben". Dat lijkt een sluitende en afdoende redenering".... als blijkt dat ze het kennen en kunnen en begrepen hebben". Het levensgrote gat in deze redenering zit in de vaagheid van het woordje "het", dit woordje "het" kan niet slaan op de fysica als geheel. Er zijn aspecten aan het fysische kennen, kunnen, begrijpen en toepassen die op een schriftelijk examen niet kunnen blijken. Wie met mij van mening is dat het practicum (naast een nuttig hulpmiddel) een integrerend deel is van de fysica en van het onderwijs in de fysica, die zal het met mij een hoogst onbevredigende toestand vinden dat bij de beoordeling van de leerling zijn prestaties op dit essentiële onderdeel niet eens worden onderzocht.

Als het grote voordeel van het schoolonderzoek zie ik het scheppen van de mogelijkheid, voor iedere leraar individueel, het practicum in het oordeel over de leerling te betrekken. Enige jaren geleden heb ik in een artikel in Faraday (39e jaargang, No. 5, pag. 134) een aantal voorbeelden gegeven voor practicumproefwerken. In Doen en Denken deel I en II staan er nog meer. En als u er eenmaal mee begint vindt u ze zelf aan de lopende band.

Nu wil ik nog uw aandacht vestigen op de vele facetten van het fysisch bezigzijn die besloten liggen in een eenvoudige proef met de daarbij behorende vragen.

- Lezen en uitvoeren van een eenvoudige opdracht, daarbij aandacht schenkende aan de aangegeven "negatieve" precisie.
- Waarnemen.
- Verklaren aan de hand van een zelf getekende figuur.
- Aanbrengen en beschrijven van een zelf te vinden verandering in de proefopstelling met een voorgeschreven gevolg.
- Formuleren, in woorden en in formulevorm, van een vermoedde, maar niet eerder geleerde, relatie tussen een aantal grootheden in de proefopstelling.

- Lezen en uitvoeren van een nadere opdracht.
- Waarnemen en beschrijven.
- Aanbrengen van een voorgeschreven verandering.
- Waarnemen en beschrijven van het gevolg.
- Verklaren aan de hand van een zelf getekende figuur (2 maal).
- Lezen van de beschrijving van een analoge proef met afwijkende uitkomst.
- Verklaren van een beschreven afwijkende uitkomst (N.B. De leerling heeft geen andere lens; hij kan het dus niet proberen).
- Juist interpreteren van numerieke gegevens.
- Een bekende formule hanteren in een nieuwe fysische situatie.
- Denken aan andere oplossing.

Terloops zij opgemerkt dat het goed kunnen lezen en uitvoeren van een opdracht ook voor niet-fysici een nuttige verworvenheid is. Denkt U maar aan het in gebruik stellen van een nieuw (huishoudelijk) apparaat na het lezen van de gebruiksaanwijzing.

Wellicht vraagt U zich af of inderdaad de zojuist gepresenteerde waslijst in z'n geheel kan voorkomen bij een eenvoudige proef. Daarom wil ik U nu de proef tonen en aan de hand van proef en opdracht de lijst nog eens doornemen.

Iedere leerling vindt op zijn tafel de volgende benodigdheden.

- 1 fietslampje in fitting met batterij of aansluiting op laagspanningscontactdoos
- 1 pos. lens (zonder montuur) met kleefwas bevestigd op een blokje hout
- 1 blikken plaatje, ongeveer even groot als de lens, met kleefwas bevestigd op blokje hout (lampje, lens en blikken plaatje bevinden zich ongeveer even hoog boven de tafel)
- 1 mat-glazen scherm, verticaal bevestigd in blok hout met zaagsnede

Ik vestig Uw aandacht er op hoeveel tijd het kost om de beschrijving te lezen van een fysische situatie die in werkelijkheid gepresenteerd in één oogopslag is te overzien.

De leerling krijgt een papier met opdrachten.

Opdracht 1:

- a. Zet het brandende fietslampje ongeveer 60 cm voor het scherm. Zet het blikken plaatje ongeveer midden tussen lampje en scherm.

Je ziet nu op het scherm een donkere vlek. Verklaar het ontstaan van die vlek aan de hand van een tekening van de stralengang.

Opmerking: Het begin is zeer eenvoudig, met opzet, om hem op zijn gemak te stellen. Toch biedt de opdracht gelegenheid te constateren of de leerling een opdracht kan lezen en uitvoeren met inachtneming van de aangegeven "negatieve precisie".

Of men één of meer strafpunten wil geven aan de leerling die heel nauwkeurig 60,0 cm afmeet en dan nog eens 30,0cm, valt te overwegen. Men kan ook verdedigen dat hij al genoeg gestraft is door het tijdverlies die deze overbodige precisie levert.

"A thing worth doing, is worth doing well enough for the purpose at hand. Doing it any better is probably silly and surely a waste of time" (Maxwell).

Hij moet tevens een eenvoudig waargenomen verschijnsel verklaren.

- b. Hoe kun je met hetzelfde plaatje een grotere donkere vlek op het scherm krijgen zonder het scherm te verplaatsen? (2 manieren).

Hij moet een voorgeschreven gevolg kunnen veroorzaken en kwalitatief beschrijven.

Het antwoord "lampje en/of plaatje verschuiven" is niet voldoende want er staat "grotere vlek".

- c. Welk verband bestaat er volgens jou tussen de middellijn van het plaatje, de middellijn van de vlek en andere afmetingen in de proefopstelling?

Je hoeft dat verband niet af te leiden of te bewijzen. Schrijf het alleen maar op in woorden en in formule-vorm. (Kies zelf de letters voor de diverse grootheden).

Het in 1b geconstateerde kwalitatieve verband moet nu in woorden en in formules gekwantificeerd worden.

Hier liggen voetangels en klemmen.

De proefopstelling roept wellicht associaties op met een formule uit de geometrische optica  $\frac{BB_1}{LL_1} = \frac{b}{v}$  wat vertaald kan worden in

$$\frac{R_{\text{schaduw}}}{R_{\text{plaatje}}} = \frac{b}{v}$$

Of dit antwoord goed is hangt er van af of hij de juiste afstanden met b en v benoemt.

Als hij scherm met beeld en lampje met voorwerp associeert dan is het fout.

Het in de opdracht tussen haakjes geplaatste kan men ook weg laten. Een eindexamenkandidaat die niet zelf op dat idee komt verdient niet anders dan te verdrinken in de vele woorden waarmee hij de afstanden gaat omschrijven.

Opdracht 2:

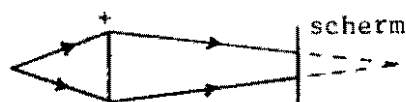
- a. Zet het brandende lampje ongeveer 60 cm van het scherm en plaats de lens ongeveer midden tussen lampje en scherm. Beschrijf wat je op het scherm ziet.

Afmetingen en lenssterkte zijn zo gekozen dat er geen beeld van de gloeidraad op het scherm ontstaat maar een lichtvlek met donkere rand.

- b. Schuif nu, terwijl lampje en lens op hun plaats blijven, het scherm langzaam naar de lens toe. Beschrijf de veranderingen op het scherm. Afmetingen en lenssterkte zijn zo gekozen dat de lichtvlek eerst kleiner wordt en daarna weer groter. Of men tevreden zal zijn met de beschrijving van deze twee veranderingen (eerst kleiner, dan groter) of dat men ook verlangt dat het kleiner worden van de buitendiameter van de donkere rand wordt opgemerkt en beschreven, verdient nadere overweging.

- c. Verklaar wat je in 2a) gezien hebt aan de hand van een schets van de stralengang.

Wellicht hebt u zich bij het lezen van de beschrijving der behoeften afgevraagd of de lens zonder montuur alleen diende om het "houtje-touwtje-karakter" te accentueren. Welnu, dat is het niet. Bij een lens met montuur verklaart de leerling het optreden van de donkere rand als schaduw van het montuur. Wilt u 2c) en 2b) van plaats verwisselen? Dat lijkt logisch, verklaren direct na het constateren, maar het in 2a) waargenomene kan ook zo



- verklaard worden, en dan zal de leerling, na het verschuiven z'n verklaring gaan herzien. M.a.w. 2c) is met opzet na 2b) geplaatst.
- d) Verklaar aan de hand van dezelfde schets wat je in 2b) hebt opgemerkt. (Geen nader commentaar).

Opdracht 3:

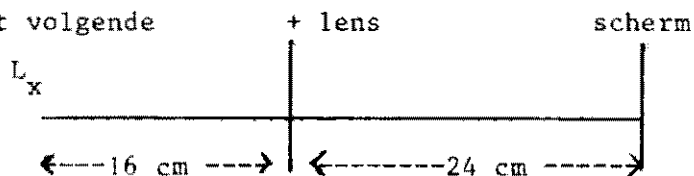
- a) Iemand doet dezelfde proef als jij net in opdracht 2 gedaan hebt, maar met een andere lens. Hij zet weer het scherm ongeveer 60 cm van het lampje en de lens ongeveer midden tussen lampje en scherm. Hij beweert, dat de lichtvlek steeds groter wordt als hij het scherm uit die stand naar de lens toeschuift, terwijl lampje en lens op hun plaats blijven.

Maak duidelijk hoe dat kan.

De leerling is nu door waarnemingen aan zeer eenvoudige proeven (zelfs de onhandigste kan ze uitvoeren) ingeleid in een voor hem nieuwe fysische situatie (bij de geometrisch optica werkt men met scherpe beelden, niet met lichtvlekken).

Hij moet nu tonen een analoge situatie met beschreven afwijkende uitkomst te kunnen verklaren. Hij kan niet gaan proberen, want hij heeft maar één lens.

- b) Iemand doet proeven zoals jij net gedaan hebt, maar met een andere lens en vindt het volgende



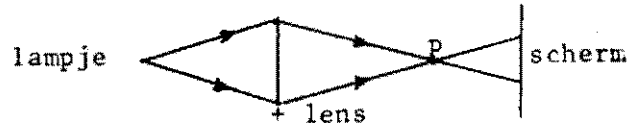
De middellijn van de lichtvlek is half zo groot als de middellijn van de lens.

Bereken de brandpuntsafstand van de lens.

Ook deze opgave zit vol voetangels en klemmen. De kans is groot dat de leerling zich eindelijk op bekend gebied waant. De brandpuntsafstand wordt gevraagd en hij schrijft meteen op:

$$\frac{1}{f} = \frac{1}{b} + \frac{1}{v} \quad b = 24 \text{ cm} \\ v = 16 \text{ cm}$$

Een ander die zich bewust is dat een lichtvlek geen beeld is, kijkt naar de figuren die hij bij 2c) gemaakt heeft.



Maar hij heeft geen formule nodig om  $f$  uit te rekenen want hij noemt P het brandpunt.

Slechts de allerbesten lossen dit goed op en slechts een deel daarvan vermeldt dat er twee oplossingen zijn.



3. Voordracht Broekman: Psychometrische aspecten van de meerkeuzetoets

Doel van vrijwel alle onderwijs is, dat een docent tracht bepaalde veranderingen teweeg te brengen in het gedrag van een leerling. Een leerling die bijvoorbeeld niet kan vermenigvuldigen kan deze vaardigheid leren door middel van onderwijs.

Bloom c.s. onderscheiden drie gebieden waarop deze veranderingen zich kunnen voltrekken:

- a. cognitief: b.v. het leren rekenen,
- b. affectief: belangstelling krijgen voor b.v. natuurkunde,
- c. psychomotorisch: b.v. het leren opbouwen van een proefopstelling; het leren plakken van een fietsband.

Heeft onderwijs succes gehad, dan is het onderwijsresultaat, dat de leerling een vaardigheid beheerst en in staat is een nieuw soort gedrag te vertonen.

Prof. Ebel, hoogleraar aan de Michigan State University, een autoriteit op het gebied van studietoetsen, verdedigt in zijn boek "Measuring Educational Achievement" de opvatting dat een opvoedkundige dus ook een onderwijskundige beïnvloeding op welk gebied dan ook, moet betekenen, dat er een verandering in het gedrag gaat optreden en hij zegt: "Overal waar die verandering kan worden geconstateerd kan zij ook gemeten worden".

Daarmee komen we voor een uiterst belangrijke vraag te staan namelijk: "met welke procedure kunnen onderwijsresultaten geëvalueerd worden," d.w.z.

volgens welke methode kan worden vastgesteld of een leerling kennis, inzicht en vaardigheden op een of ander vakgebied heeft verworven en zo ja, in welke mate?

Om dit te kunnen achterhalen gaan docenten over tot toetsing.

Het woord toets mag misschien nog wat vreemd klinken, niet echter het begrip toetsen want dit is zo oud als het leren.

Men kan het begrip toets zeer algemeen omschrijven (Solberg) en wel als volgt:

"een toets is iedere procedure die leidt tot een uitspraak over kennis, begrip en vaardigheid van iemand op grond van door hem behaalde resultaten naar aanleiding van een aantal gegeven opdrachten.

Onder deze omschrijving kunnen worden begrepen:

proefwerken, overhoringen, scripties, tentamens, meerkeuze-toetsen, praktijkexamens etc.

In deze definitie is de term toets dus niet gereserveerd voor een bepaalde vorm van toetsen.

De functie van de toets is het meten van onderwijsresultaten.

Men wil nagaan welk effect de onderwijs- en leerprocessen hebben gehad.

Er is een engere opvatting van de term toets mogelijk.

In zijn boek studietoetsen, construeren, afnemen, analyseren (A.D. de Groot en R.F. van Naessen) definieert Prof. de Groot een studietoets, en wel als volgt:

"Ieder proefwerk, examen, tentamen, ieder hulpmiddel voor schriftelijke toetsing van door onderwijs en studie verworven kennis, inzicht of vaardigheid op een of ander vakgebied, mits de bepaling van de score (het aantal goede antwoorden), die een (proef-)persoon behaald heeft, geheel objectief kan geschieden".

Ook de studietoets is een instrument waarmee men de stand van iemands kennis, inzicht en vaardigheden op een bepaald gebied kan meten.

Het is duidelijk dat men lang niet alle effecten van onderwijs kan bepalen met behulp van studietoetsen. Ten aanzien van kennis en inzicht zijn er weinig beperkingen met betrekking tot het gebruik van studietoetsen.

Maar van de vele soorten vaardigheden die men kan onderscheiden, lenen zich alleen bepaalde intellectuele vaardigheden tot meting met behulp van studietoetsen, zoals b.v. het kunnen toepassen van een wet of een geleerd principe, de analyse van een probleem, het kunnen interpreteren van een stuk tekst etc.

Creatieve, sociale en praktische vaardigheden zoals het schrijven van een opstel, het kunnen leiden van een groep personen, het repareren van een versterker, vallen grotendeels buiten het "meetbereik" van studietoetsen.

#### Score.

Er wordt bij deze definitie van studietoets gesproken van score en niet van cijfer om te doen uitkomen dat de vaststelling van de grens voldoende/onvoldoende (caesuur) nog een apart probleem is. Bij de bepaling van de scores is nog geen uitspraak gedaan over de vraag hoe aan bepaalde scores cijfers of beoordelingen moeten worden toegevoegd.

#### Objectiviteit.

De term objectiviteit in de Groot's definitie van de studietoets heeft niets te maken met de gedachte als zou de inhoud van de studietoets, "vrij van subjectief bepaalde voorkeuren" zijn. In dit opzicht kan de inhoud van de studietoets even subjectief bepaald zijn als de inhoud van een proefwerk bestaande uit open vragen. De term objectief heeft alleen de betekenis: zō, dat de persoon van de beoordelaar, mits hij zich aan de scoringsvoorschriften houdt, géén rol meer kan spelen.

De scoringstaak kan bijvoorbeeld worden overgenomen door iemand die niets van het vak afweet, desgewenst door een machine, mits de regels voor goed- of fout-rekenen en voor bepaling van de score beslist eenduidig zijn.

Objectief geregeld is alleen datgene wat nã de constructie van de studietoets plaatsvindt namelijk de afname, de correctie en de bepaling van de scores van de leerlingen.

Men kan algemeen stellen, dat iedere beoordelingsprocedure aan bevredigende objectiviteitseisen moet voldoen.

De enige factoren die de toetsresultaten mogen bepalen zijn iemands kennis inzicht en vaardigheden. Storende factoren die het proces van beoordelen kunnen beïnvloeden zoals:

- minder fraai handschrift van de leerling,
- verbeteringen en doorstrepingen in het proefwerk,
- vooroordelen in de trant van "Oh, dit is het werk van Piet, dat zal wel weer niet veel zaaks zijn,
- stijl,
- beïnvloeding door nawerking van voorafgaande beoordelingen,
- optredende vermoeidheid doordat men een groot aantal proefwerken moet corrigeren,,
- een al te persoonlijke opvatting van de onderwijsdoelstellingen etc.,

moeten voorkomen worden.

### De studietoets van het meerkeuze-type.

De meeste studietoetsen bestaan uit enige tientallen kleine opgaven (items) van het meerkeuze type.

Ieder goed beantwoord item levert gewoonlijk één punt op waarbij de totaalscore tot stand komt door de som van de itemscores te bepalen. De term "meerkeuze" is de vernederlandsing van "multiple choice" en uit deze laatste naam blijkt dat deze methode uit de Angelsaksische landen afkomstig is. Al vóór de tweede wereldoorlog werd er in Amerika met multiple choice toetsen gewerkt.

Het gaat bij deze vraagvorm om het kiezen uit reeds gegeven antwoorden (alternatieven).

Een alternatief is een goed of fout antwoord; een afleider is een fout antwoord.

In elke opgave wordt de leerling een vast aantal alternatieven aangeboden. Hierbij bevat één alternatief het juiste antwoord. De overige alternatieven zijn afleiders.

De alternatieven kunnen bijvoorbeeld zijn: verschillende antwoorden op een vraag, aanvullingen op een zin, een aantal aparte beweringen of een aantal diagrammen, schakelingen of tekeningen.

De leerling moet trachten het goede antwoord te vinden, d.w.z. hij moet uit de aangeboden alternatieven kiezen.

Zijn keuze legt hij vast op een antwoordblad.

Het "meerkeuze-type" wordt zeer veel toegepast want deze vorm is voor vele doeleinden zeer bruikbaar gebleken, dit omdat men bij meerkeuze-items een aantal belangrijke voordelen kan onderscheiden:

- meerkeuze-items zijn geprecodeerd, het goede antwoord staat van te voren vast, is voor iedereen zichtbaar en maakt een objectieve vergelijkbaarheid mogelijk,
- met behulp van itemvormen kan men tot onaanvechtbare relatieve prestatie-beoordelingen komen,
- meerkeuze-items zijn flexibel, men kan zeer veel problemen zonder veel verlies aan inhoud en efficiëntie in deze vraagvormen omzetten.

Een meerkeuze-item dat zeer veel voorkomt is het vierkeuze-item.

Het komt wel voor dat men met 5 of meer alternatieven scherpere vragen kan stellen. Men kan bovendien opmerken, hoe meer alternatieven des te beter, immers des te kleiner wordt de kans op "goed" raden.

Bij 6 of meer alternatieven wordt een item echter al spoedig onoverzichtelijk. De ervaring leert dat itemschrijvers veelal niet meer dan vier acceptabele alternatieven weten te bedenken. Het kost vaak veel tijd om nog meer onjuiste alternatieven te vinden.

Bij items met slechts twee alternatieven wordt de kans op "goed" raden zó groot, dat men zeer veel vragen zou moeten stellen, wil men zeker zijn van een betrouwbaar resultaat.

Een variant van een tweekeuze-item is het waar/onwaar type.

Deze vraagvorm is niet geschikt voor elke stof. Een bewering moet bij deze vorm absoluut waar of onwaar zijn. Uit een onderzoek van L.J.Cronbach (1950) is gebleken dat er bij het beantwoorden van dit type vragen antwoordtendenties optreden. Veel leerlingen blijken geneigd meestal met "waar" te antwoorden, anderen daarentegen vertonen de neiging op alle vragen bijna altijd met "onwaar" te antwoorden.

Voordat ik er toe overga u een aantal vierkeuze-items voor natuurkunde te tonen is het noodzakelijk u eerst enige informatie te geven over itemanalyse en de daarbij gebruikte termen.

Deze items die op pagina 42 en volgende staan afgedrukt kwamen voor in proef-toetsen die via proefafnames aan ongeveer 600 leerlingen (de zogenaamde steekproef) zijn voorgelegd.

De bedoeling van een proefafname is het verkrijgen van informatie over de toets en de items door middel van een toets- en itemanalyse.

De proefafname geeft informatie over:

- de moeilijkheid van ieder item,
- de mate waarin het item onderscheidt tussen de hoogst-scorende en de laagst-scorende leerlingen,
- de wijze waarop de afleiders hebben gefunctioneerd,
- de mate van overeenstemming tussen de geschatte en de feitelijk benodigde werktijd.

De bij de proefafname behorende toets- en itemanalyse leveren een belangrijke bijdrage om te kunnen nagaan in hoeverre de toets verbeterd dient te worden. Ook van de herziene versie van de toets is een nieuwe proefafname noodzakelijk om te kunnen vaststellen in hoeverre de reconstructie ook een verbetering is.

Itemanalyse is een analyse van de item-resultaten door middel van p-, a- en  $r_{it}$ -waarden.

De p-waarde van een item geeft aan hoeveel % van de leerlingen het goede antwoord heeft gekozen. Hoe hoger de p-waarde, hoe gemakkelijker het item. De p-waarde is herkenbaar door een getal met een sterretje (\*).

De a-waarde (afleider-waarde) is het percentage leerlingen dat deze afleider als het goede antwoord beschouwde.

Voorbeeld: gegeven een vierkeuze-item met de alternatieven A, B, C en D.

Laat A het goede antwoord zijn; B, C en D zijn dan afleiders.

Heeft nu b.v. 70% van de leerlingen A gekozen en 12, 8 en 10% de respectievelijke afleiders B, C en D dan heeft dit item:

een p-waarde van 70 en

drie a-waarden van respectievelijk 12, 8 en 10.

#### Normen voor p- en a-waarden.

Allerlei toevallige factoren hebben invloed op de toetsscore.

Minder goede lichtval in het lokaal, storende geluiden van buiten, het in goede conditie zijn of juist op de dag verkouden. Voor iedere leerling zijn deze invloeden verschillend.

Men splitst nu in de testleer de ruwe toetsscore X in een ware score W (waar het in feite om gaat) en een toevallige fout e (error).

De ruwe score  $X = W + e$ , en onder ideale omstandigheden zou  $X = W$  zijn.

Door de verschillen die er bij de leerlingen bestaan t.a.v. kennis, inzicht en vaardigheden, ontstaat er een spreiding in de toetsscores.

Naarmate de scores op de toets meer spreiden, meer variëren is de "toetsvariantie"  $V_x$  groter.

In het algemeen stelt men nu dat de toetsvariantie  $V_x$  gelijk is aan de som van een ware variantie  $V_w$  en een foutenvariantie  $V_e$ . ( $V_x = V_w + V_e$ ).

(De betrouwbaarheid (r) van een toets geeft aan de mate waarin de toetsvariantie  $V_x$  "ware variantie" is en géén "foutenvariantie".

$$r = \frac{V_w}{V_x}$$

De betrouwbaarheid van een toets wordt op pag.40 nog nader toegelicht).

Ook een item draagt met zijn ware variantie bij tot de ware variantie van de toets en met zijn foutenvariantie tot de foutenvariantie van de toets. Hoe groter de bijdrage tot de ware variantie des te groter is de bijdrage tot de toetsbetrouwbaarheid.

Items die door alle leerlingen worden goed gemaakt hebben een variantie gelijk aan nul. Deze items hebben geen invloed op de rangorde van de leerlingen of op de toetsbetrouwbaarheid. Om deze reden streeft men in het algemeen naar items met noch te hoge noch te lage p-waarden (Van Naerssen pag. 248).

De itemvariantie is maximaal bij een p-waarde van 50.

De p- en a-waarden zeggen zonder verdere aanduiding niet zoveel. Deze waarden worden pas zinvol wanneer zij tegen een criterium worden afgezet. Dit criterium is echter afhankelijk van het doel waarvoor de toets gebruikt wordt. De items die tijdens de lezing en in de discussiegroepen zijn getoond waren alle opgenomen in toetsen die vaardigheden over een bepaald onderwerp bedoelden te toetsen op het niveau van de respectievelijke eindexamens voor MAVO-3, MAVO-4 en het L.T.O.-T.

Aan de p- en a-waarden zijn daarom dezelfde normen gesteld als die welke gesteld worden bij eindexametoetsen. Net als bij examentoetsen lag het ook bij deze proeftoetsen in de bedoeling de toets goed te laten differentiëren over het gehele scoregebied, d.w.z. zoveel mogelijk onderscheid te laten maken tussen de scores van de leerlingen.

Voor dat doel zijn items met een p-waarde tussen de 49 en 75 het beste.

Voor de toetsen die bestemd waren voor einde 3e klas HAVO of einde 3e klas VWO (breekpunten) zijn voor de p- en a-waarden dezelfde normen gehanteerd.

Voor examentoetsen is m.b.t. normen voor p- en a-waarden het volgende compromis gevonden:

<u>p-waarden:</u> minder dan 49	: te laag	<u>a-waarde</u> niet minder dan 5.
50 - 74	: goed	
75 en hoger	: redelijk	

Hoe extremer de p-waarde, hoe minder het item in het algemeen bijdraagt. Dergelijke items verhogen de gemiddelde score, de toetsvariantie en de toetsbetrouwbaarheid worden door deze items echter niet beïnvloed.

Te moeilijke items bederven de betrouwbaarheid van de studietoets.

Zij kosten de leerlingen meer tijd dan de overige items maar bovendien heeft de factor raden vooral een ongunstig effect bij de moeilijke items, zie R.F. van Naerssen<sup>1</sup>), pag. 249.

Hij merkt hierover op: "Indien niemand bij een bepaald vierkeuze-item het antwoord weet, dan zal door raden alleen toch nog 25% het juiste antwoord aangestreept hebben, zoals we weten. Maar alle variantie van dit item is foutenvariantie. Het gaat bij items vooral om de verhouding tussen wat zij bijdragen aan ware variantie en wat aan foutenvariantie. Deze verhouding is bij studietoetsen door het raden des te ongunstiger naarmate het item moeilijker is. Te gemakkelijke items 'wegen zichzelf weg' door hun geringe variantie; zij zijn daarom inefficiënt. Maar te moeilijke items bederven door hun foutenvariantie de betrouwbaarheid van de studietoets."

Een uitzondering hierop vormt het geval dat een lage p-waarde toch vergezeld gaat van een  $r_{it}$ -waarde van b.v. 0.40 of meer. In dergelijke gevallen kunnen items met lage p-waarde toch nog de betrouwbaarheid van de toets verhogen. Zie voor begrip  $r_{it}$  pag. 37.

Itemanalyse beperkt zich niet tot een beschouwing van alleen p- en a-waarden. Een goed item moet als het differentieert - dus als de p-waarde niet of bijna 100 is, - zó differentieëren dat het onderscheid maakt tussen "betere" en "minder goede" leerlingen (of tussen betere en minder goed voorbereide).

Heeft men een criterium voor "betere" en "minder goede" leerlingen, b.v. een rapportcijfer dan mag men veronderstellen dat onder degenen die het item goed hebben, relatief meer "betere" leerlingen zullen voorkomen dan onder hen die het item fout hebben.

Men kan vervolgens vragen hoe sterk het item discrimineert; men kan vragen naar de grootte van de correlatie tussen itemscore en gekozen criterium (in dit geval rapportcijfer).

De grootte van die correlatie (en die is te berekenen) noemt men itemvaliditeit t.o.v. het gekozen criterium. Er moeten positieve itemvaliditeiten worden verwacht t.o.v. juist gekozen criteria.

Hiermee zijn we gekomen bij het begrip validiteit van de toets.

Het begrip validiteit is een verzamelbegrip.

De vraag naar de validiteit van een instrument heeft altijd betrekking op "de mate waarin dat instrument beantwoordt aan het doel waarvoor het wordt gebruikt".

Bij de validiteitsanalyse gaat het altijd om de vraag of wat we meten wel hetgene is dat we willen toetsen.

De validiteit van een toets kent twee facetten:

de inhoudsvaliditeit: vormen de verschillende items tezamen een toets die als geheel de hoofdzaken van de leerstof redelijk dekt en

de begripsvaliditeit: in hoeverre meet de toets kennis, inzicht t.o.v. bijvoorbeeld natuurkunde en niet tekstbegrip.

Beantwoordt de toets aan de doelstellingen van het onderwijs? De Groot formuleert het aldus: "in hoeverre representeren de opgaven in een toets getrouw en evenwichtig de soorten opgaven die tezamen een goede operationele definitie zouden vormen van de kennis, het inzicht en/of de vaardigheid die men heeft willen aankweken en die men nu wil meten"?

Onderzoeken we dit bij de afzonderlijke items dan kunnen itemvaliditeiten vooral als zij vergelijkend worden bekeken zeer nuttige gegevens verschaffen.

Zij geven zelden beslissende informatie over de waarde van de items.

Itemgegevens - en dus ook p- en a-waarden - zeggen nooit alléén iets over de kwaliteiten van een item. Itemgegevens zijn afhankelijk van het item zelf maar ook van de groep leerlingen die het beantwoorden. Tevens spelen daarbij een rol het gegeven onderwijs, de omstandigheden waaronder de toets wordt afgenomen en de mogelijk storende factoren die wij "toeval" noemen.

Richten we de toepassing van studietoetsen op meten - evalueren - van onderwijsresultaten, dan is het vaak moeilijk voor bepaling van itemvaliditeiten een criterium te vinden dat houvast biedt.

Men kan echter altijd een empirisch gegeven verkrijgen over de items van de afgenomen toets, namelijk de correlaties tussen de afzonderlijke itemscores en de totaalscore op de toets zelf. Deze correlaties worden genoemd de item-totaal-correlaties ( $r_{it}$ -waarden).

De  $r_{it}$  (item-totaal-correlatie) is de correlatie tussen itemscore (fout = 0; goed = 1) en de totaalscore op de toets.

De  $r_{it}$  van een item geeft aan in hoeverre een item goed is beantwoord door de goede (d.w.z. hoog scorende) leerlingen en fout is beantwoord door de minder goede (d.w.z. laag-scorende) leerlingen.

Hoe hoger de  $r_{it}$ , hoe beter het item onderscheid maakt tussen goede en minder goede leerlingen.

Van de toets berekent men eerst de gemiddelde score  $\bar{Y}$  en de standaarddeviatie  $S_y$ . De standaarddeviatie is de wortel uit de variantie\*).

Van elk item afzonderlijk berekent men het aantal personen P, dat het item goed heeft en hun gemiddelde score  $\bar{Y}_1$ .

(De p-waarde =  $\frac{P}{N} \times 100\%$ ).

De  $r_{it}$  is gelijk aan:  $\frac{\bar{Y}_1 - \bar{Y}}{S_y} \sqrt{\frac{p\text{-waarde}}{(1-p \text{ waarde})}}$

Voor een examentoets wordt voor de p-, en  $r_{it}$ -waarden de volgende normen gehanteerd:

minder dan 49	: te laag	$r_{it}$ van 0,40 en hoger	: goed tot zeer goed
50-74	: goed	$r_{it}$ van 0,30-0,39	: redelijk goed
75 en hoger:	redelijk	$r_{it}$ van 0,20-0,29	: twijfelachtig
		$r_{it}$ van 0,19 en lager	: slecht

Een gemakkelijk item (p-waarde is bijvoorbeeld 96) dat een  $r_{it}$  heeft van 0,40 heeft met succes de 4% "slechte (d.i. laag-scorende) leerlingen weten te onderscheiden.

Betekenis van p-, a- en  $r_{it}$ -waarden bij de itemanalyse.

p-waarde.

Wat betekent nu een te lage p-waarde? Hierop is meer dan één antwoord te geven:

- a. Men moet zich bij de itemanalyse in het geval van een te lage p-waarde afvragen of zo'n item qua moeilijkheid wel goed op de groep leerlingen en het onderwijsniveau is afgestemd.
- b. Ook leert de ervaring dat ondanks overeenstemming van deskundigen er soms een fout of zwakheid in de formulering van het item is gekomen die maakt dat het als goed bedoelde alternatief niet zo onaanvechtbaar is als men aanvankelijk meende. Als een of meer alternatieven als het juiste verdedigd kunnen worden zullen die alternatieven ook door vele leerlingen worden gekozen (er zullen hoge a-waarden volgen).
- c. Het item kan een strikvraag bevatten waardoor de leerlingen op een dwaalspoor komen.
- d. Het kan zijn dat het item qua taalgebruik en lay-out onoverzichtelijk is. Misschien was er een tekening bij nodig geweest of is de gegeven tekening onduidelijk of fout.
- e. Het kan ook zijn dat bij het onderwijs het desbetreffende onderdeel van de stof dit jaar wat minder aandacht heeft gekregen. Het onderwijs over dit gedeelte van de stof kan niet goed zijn overgekomen. Leerboeken of stencils zijn misschien her en der onduidelijk.

\*) Wanneer men van de ruwe toetsscore Y de gemiddelde toetsscore  $\bar{Y}$  aftrekt krijgt men de deviatiescore y. ( $y = Y - \bar{Y}$ ). Men definieert de variantie

van een toets als  $\sum \frac{y^2}{N}$  (N is de totale populatie)  
 De standaarddeviatie  $S_y = \sqrt{\sum \frac{y^2}{N}} = \frac{1}{N} \sqrt{N\sum Y^2 - (\sum Y)^2}$

De oorzaken b en c veroorzaken een slechte  $r_{it}$ . Zowel 'goede' als 'slechte' leerlingen gaan dit item fout maken. Het item onderscheidt dus niet goed, d.w.z. niet in overeenstemming met de rest van de toets.

De gegevens van de itemanalyse kunnen bij ieder item aanleiding geven tot vermoedens over de oorzaak van de resultaten.

De docent krijgt informatie over zijn onderwijs en de onderwijsresultaten en via de analyse van de itemgegevens ook aanwijzingen aan welk onderdeel van de leerstof extra aandacht moet worden besteed.

De itemgegevens kunnen opvattingen en veronderstellingen t.a.v. de inhoud van de probleemstelling in het item weerleggen en steunen.

De opmerking "dit item is voor dit schooltype te moeilijk" kan worden weerlegd door een hoge p-waarde. De opmerking "dit item is alleen goed te maken door raden" kan worden weerlegd door een hogere p-waarde dan 25 en een  $r_{it}$  van b.v. 0,40 of hoger.

Is de p-waarde gelijk aan 100 (alle leerlingen hebben het item goed beantwoord) dan kan men twee dingen zeggen:

- het onderwijsdoel dat door dit item wordt getoetst is door alle leerlingen in de groep bereikt,
  - dit item differentieert niet in deze groep, het is te gemakkelijk.
- Er is een kans dat het triviaal is of dat het kan worden opgelost zonder vakkennis.

a-waarde. (het percentage leerlingen dat een afleider als het goede antwoord beschouwt).

Men hoopt in het algemeen bij niet te hoge p-waarden dat de a-waarden niet gelijk aan nul zullen zijn. Een afleider met een a-waarde gelijk aan 0 heeft in deze groep leerlingen niet overeenkomstig de bedoeling gewerkt, namelijk als een gangbare fout of een voor de hand liggende misvatting. Het kan ook zijn dat het alternatief waarop deze a-waarde betrekking heeft zo weinig plausibel is, dat het ook voor slechte leerlingen niet voor keuze in aanmerking komt.

Is een a-waarde groter dan of bijna even groot als de p-waarde, dan kan dit wijzen op

- een fout in het item - het foute alternatief waarop die a-waarde betrekking heeft is niet veel slechter dan het juiste antwoord,
- er is een kans dat er in het onderwijs onvoldoende tegen deze misvatting of foutenbron is gewaarschuwd, aangenomen dat het goede antwoord bij nader onderzoek als onaanvechtbaar juist stand houdt.

$r_{it}$ -waarde.

Er is voorzichtigheid geboden bij de interpretatie van  $r_{it}$ -waarden. Niet altijd geldt hoe hoger de  $r_{it}$ -waarde hoe beter het item. Heeft men in het onderwijs getracht de leerlingen zoveel mogelijk feitenkennis bij te brengen dan zal een toets die deze kennis wil toetsen best een grote spreiding in de scoreresultaten kunnen opleveren.

De items kunnen naar psychometrische maatstaten goed voldoen maar vanuit onderwijskundig standpunt verwerpelijk zijn.

Een item met een  $r_{it}$  kleiner dan 0,20 hoeft niet alleen op grond hiervan te worden gediskwalificeerd. Het item kan gaan over een stofonderdeel dat enigszins los staat van de rest. Het kan een beroep doen op andere capaciteiten terwijl het item op zich goed kan zijn.

Ook kan het item op zich juist zijn, maar betrekking hebben op een onderdeel van de stof, dat onvoldoende of verkeerd is onderwezen, zodat de goede leerlingen het systematisch fout doen.



Een lage  $r_{it}$ -waarde kan ook worden veroorzaakt doordat het item een te aantrekkelijke afleider bevat, die eigenlijk niet zo erg fout is of niet zoveel slechter dan het beste antwoord (gevolg is ook een hoge  $a$ -waarde).

Het item kan veel te moeilijk zijn zodat bijna alle leerlingen maar eens een gokje wagen. Het item kan dan niet onderscheid maken tussen degenen die het probleem wel konden oplossen en zij die het niet konden.

Een positieve  $r_{it}$ -waarde betekent dat leerlingen die het desbetreffende item goed beantwoord hebben gemiddeld een hogere score hebben dan zij die het niet goed beantwoord hebben. Een negatieve  $r_{it}$ -waarde betekent dat de leerlingen die het antwoord goed beantwoord hebben, gemiddeld een lagere score op de toets hebben dan de anderen.

#### Betrouwbaarheid van de toets.

Op pagina 35 is het woord betrouwbaarheid genoemd in verband met het begrip toetsvariantie.

De betrouwbaarheidscoëfficiënt  $r$  of kortweg betrouwbaarheid wordt gedefinieerd als de proportie ware variantie (waar het om gaat) in de totale toetsvariantie (ware variantie + foutenvariantie).

De betrouwbaarheidscoëfficiënt van een toets is een getal tussen 0 en 1.

Bij de vraag naar betrouwbaarheid gaat het om de nauwkeurigheid van het meetinstrument als meetinstrument.

Betrouwbaarheid van een toets geeft aan, hoe constant de toets de vaardigheden van leerlingen meet, niet welke vaardigheden de toets meet want deze vraag heeft betrekking op het begrip validiteit van de toets.

Betrouwbaarheid geeft aan in hoeverre de leerling bij herhaling van bijvoorbeeld het examen (zonder dat hij de inhoud van de examentoets kent) dezelfde score zou behalen.

De leerlingen moeten door de toets op een eerlijke wijze op volgorde van de door de toets gemeten vaardigheid worden gezet, d.w.z. op dezelfde volgorde nu als bij een eventuele herhaling van het examen.

Hoe groter de betrouwbaarheid, hoe minder "toevalsfactoren" invloed hebben op de score van de leerling. Een toets heeft geen betrouwbaarheid zonder de populatie te noemen waarop de toets is afgenomen. De lengte van de toets is mede bepalend voor de betrouwbaarheid.

Een hoge betrouwbaarheid is zeker in het geval van een examen een dwingende eis. Een hoge betrouwbaarheid wordt bereikt wanneer de toets aanleiding geeft tot een grote spreiding in de toetsscores van de leerlingen. Deze grote spreiding wordt weer bereikt door items met een  $p$ -waarde tussen de 50 en 74.

Dit betekent, dat veel leerlingen aan het einde van een onderwijs-type voor een tamelijk groot deel de toets niet kunnen oplossen.

Dit is in strijd met de gedachte dat een goede toets in de betekenis van "bereikte doelstellingen" juist hoge  $p$ -waarden zou moeten bevatten.

Dit laatste heeft echter weer tot gevolg een geringe spreiding in scores waardoor de betrouwbaarheid van de toets laag wordt.

In dat geval zijn de beslissingen die met een niet-betrouwbare toets worden genomen niet "eerlijk", want "geslaagd" zou dan ook "gezakt" kunnen zijn en omgekeerd.

Dit dilemma is het beste op te lossen door het kiezen van een zo betrouwbaar mogelijke toets.

Het doel van iedere toets is de fouten-marge zo klein mogelijk te maken, zeker bij die toetsen waarmee beslissingen worden genomen. Met de standaardmeetfout kan de fouten-marge bij iedere score worden aangegeven.

Enkele vierkeuze-items voor natuurkunde.

Bij deze items kunnen de volgende aanduidingen staan:

M3 = examenklas Mavo-3.

M4 = examenklas Mavo-4.

LTO-T = examenklas Lager Technisch Onderwijs T-stroom (4e klas).

H3 = 3e klas Havo.

V3 = 3e klas V.W.O.

Een sterretje (\*) bij een getal geeft de p-waarde aan.

Komen bij een item b.v. de aanduidingen M3, LTO-T en H3 voor en onder deze letters de p-, a- en  $r_{it}$ -waarden dan betekent dit, dat het item zowel aan leerlingen uit de examenklassen van de Mavo-3 respectievelijk het LTO-T is voorgelegd als aan leerlingen van de 3e klas van het Havo.

De resultaten van de itemanalyse kunnen onderling worden vergeleken.

Men legt voorzichtig een zware steen op een ijslaag.  
We beschouwen drie factoren.

I de dikte van de ijslaag.

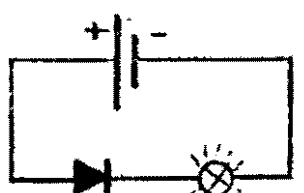
II het gewicht van de steen.

III de grootte van het contactoppervlak.

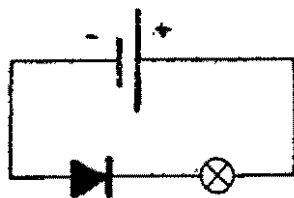
Of de steen door het ijs zal zakken hangt af:

	M3	M4	H3	LTO-T
A alleen van I en II.	15	12	19	12
B alleen van I en III.	5	6	5	4
C alleen van II en III.	10	9	5	4
D van I, II en III.	69*	73*	70*	80*
$r_{it}$	0.35	0.30	0.33	0.27

Een diode (  $\rightarrow|$  ) is een apparaat dat de stroom slechts in één richting geleidt, zie de schakelingen I en II.

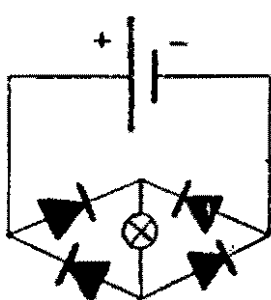


I. lamp brandt

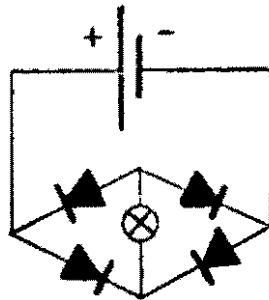


II. lamp brandt niet.

We beschouwen nu onderstaande schakelingen 1 en 2.



schakeling 1



schakeling 2

Het lampje  $\otimes$  zal branden.

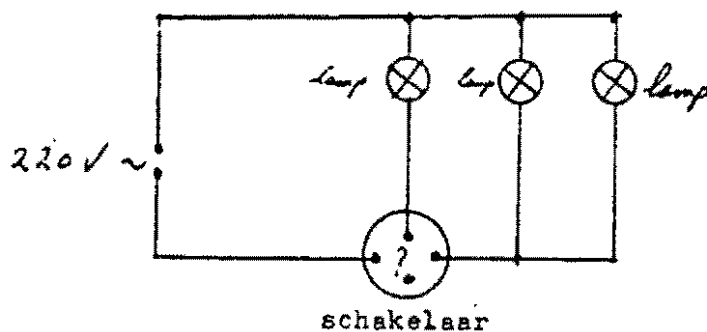
A uitsluitend in schakeling 1.

B uitsluitend in schakeling 2.

C zowel in schakeling 1 als in 2.

D niet in schakeling 1 en ook niet in schakeling 2.

M4	LTO-T	H3
7	6	10
5	4	6
58*	68*	49*
30	23	33
Zet: 0.41    0.46    0.39		



schakelaar

In bovenstaande schakeling is een schakelaar die in vier standen gesteld kan worden, zie onderstaande figuren I t/m IV.



figuur I

figuur II.

figuur III

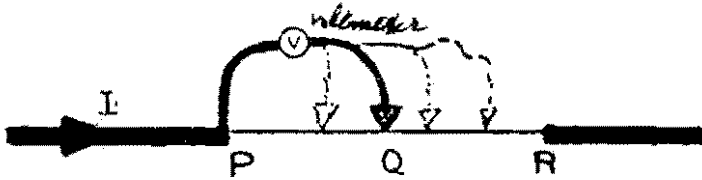
figuur IV

Als men slechts 2 van de 3 lampen tegelijk wil laten branden dan moet men de schakelaar plaatsen in de stand van

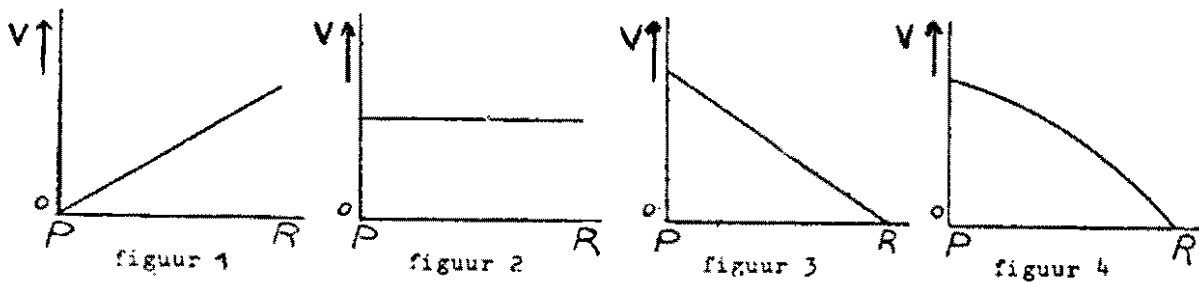
	M3	LTO-T	M4
A figuur I	16	5	5
B figuur II	8	3	7
C figuur III	72*	88*	85*
D figuur IV	3	4	2

Zet: 0.42    0.27    0.35

Door een dun draadstuk P R loopt een stroom I. Draadstuk P R is overal even dik. Men verbindt P met een voltmeter en meet de spanning op tussen P en verschillende punten gelegen tussen P en R, zie onderstaande figuur.

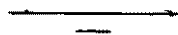


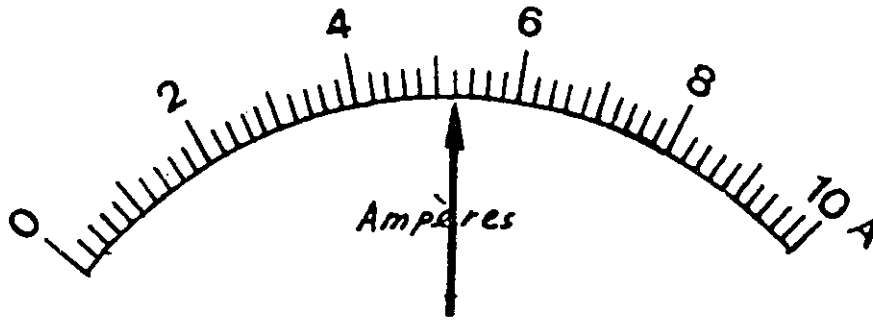
Hierna zet men in een diagram verticaal de spanning uit die de voltmeter aangeeft en horizontaal de afstand tot P.



Het verband tussen de spanning en de afstand tot P is juist weergegeven in het diagram van

	M3	LTO-T	H3
A figuur 1	40*	31*	54*
B figuur 2	37	47	27
C figuur 3	14	11	10
D figuur 4	8	8	9
$\Sigma$	0.18	0.38	0.41





De wijzer van bovenstaande ampèremeter geeft een stroomsterkte aan van

	M3	LTO-T	M4	H3
A 4,52 A	1	0	0	0
B 4,6 A	12	18	11	7
C 5,1 A	14	4	4	15
D 5,2 A	73*	77*	84*	76*
	$\Sigma$ : 0,35	0,21	0,16	0,23

Een radiator van een centrale verwarmingsinstallatie heeft een groot oppervlak omdat

	M3	M4	LTO-T	H3
A het warme water er dan gemakkelijker doorstroomt.	3	1	1	3
B er dan per $\text{cm}^2$ meer warmte wordt uitgestraald.	33	28	21	32
C er dan veel warmte aan de lucht kan worden afgestaan.	52*	63*	72*	52*
D anders de temperatuur van de radiator te veel stijgt.	10	7	6	13
	$\Sigma$ : 0,35	0,31	0,34	0,21

Men beschikt over vier genummerde koperdraden I t/m IV, waarvan enkele gegevens in de onderstaande tabel vermeld zijn.

draadnummer	massa	lengte
I	9 g	80 cm
II	9 g	160 cm
III	18 g	80 cm
IV	18 g	160 cm

De koperdraad met de grootste weerstand heeft nummer

	M3	LTO-T	M4	H3
A I	13	12	11	14
B II	47*	56*	55*	34*
C III	15	10	8	16
D IV	24	22	24	33
	$\Sigma$ : 0,41	0,47	0,40	0,31



Door een draadstuk P Q R, zie bovenstaande figuur, loopt een elektrische stroom. P Q is dikker dan Q R. De lengten van P Q en Q R zijn gelijk.

Beschouw de volgende beweringen:

- I De stroomsterkte in P Q is gelijk aan de stroomsterkte in Q R.
- II De warmteontwikkeling in P Q is gelijk aan de warmteontwikkeling in Q R.

	LTO-T	M3	H3
A I en II zijn beide juist.	2	5	7
B alleen I is juist.	57*	43*	46*
C alleen II is juist.	3	4	6
D I en II zijn beide onjuist.	37	47	39
	$\Sigma$ : 0.25	0.32	0.36

De voorwaarde, dat correctie en scoring bij een studietoets objectief moeten zijn betekent, dat verschillende beoordelaars bij hun beoordeling van het werk tot dezelfde beoordelingsresultaten moeten komen.

Deze beoordelingsresultaten moeten uiteraard te verantwoorden zijn tegenover de leerlingen.

De ervaringen leren dat dit moeilijk bereikbaar is bij grote opgaventypen als vertalingen, opstellen, samengestelde vraagstukken, waarbij het niet alleen om het antwoord gaat maar vooral om de gevolgde oplossingsmethode. Zulke opgaven vragen om een genuanceerde beoordeling om iedere leerling zoveel mogelijk recht te doen. Hoe genuanceerder de beoordelingen, des te kleiner de kans dat verschillende beoordelaars tot dezelfde beoordelingsresultaten komen.

Het is dus nodig om in studietoetsen met betrekkelijk kleine opgaven (items) te werken, opgaven die gespreid zijn over het gehele gebied waarop de studietoets betrekking heeft, en die zoals De Groot het uitdrukt "klein genoeg zijn om qua betekenis en qua niveau, onderwijskundig te kunnen worden doordacht en doorzien".

Hij merkt daarbij op, dat dit laatste bij open vragen zelden of nooit lukt, zeker niet als men met de gangbare grote opgaventypen werkt. Open vragen kan men immers praktisch altijd op verschillende wijzen en op verschillende niveaus beantwoorden met als gevolg dat de vraag wat "het kunnen oplossen van zo'n opgave" eigenlijk voorstelt, niet scherp te beantwoorden is. Een open opgave is niet altijd een duidelijk omschreven taakeenheid: een item is dat wel.

Ook de moeilijkheidsgraad (het niveau) van een open opgave is van vele factoren afhankelijk. Zo bepalen de normen en opvattingen van de beoordelaar in belangrijke mate waarin een opgave moeilijk is.

Maar ook de wijze waarop de leerling de vraag leest, zijn taak opvat en aanpakt en hoe het hem daarbij vergaat, spelen een belangrijke en veelal ondoorzichtige rol.

Hoe "moeilijk" is het onderwijskundig gezien om een tekst te verklaren of om een complexe natuur- of wiskunde-opgave op te lossen? Men kan die moeilijkheid, dat is "het niveau van de eisen" die zo'n opgave aan de leerling stelt veelal niet exact omschrijven en meestal slecht schatten.

Het is zeker niet waar, dat het schatten van de moeilijkheidsgraad van items probleemloos is. De ervaringen wijzen wel anders uit.

Toch is het bij een item veel beter mogelijk inzicht te verkrijgen in het niveau van de eisen die de opgave aan de leerlingen stelt. Immers, dat wat gevraagd wordt - het doen van de juiste keuze - is duidelijk, en welke kennis en inzicht een leerling moet hebben om het juiste antwoord te kunnen kiezen en niet in één van de veel gemaakte fouten te vervallen die in de verkeerde alternatieven zijn vervat, is door onderwijskundige kwalitatieve itemanalyse achteraf, voor een groot deel te achterhalen.

Omdat een item een kleine wel omschreven taakeenheid is, is het eerder mogelijk om het eens te worden over de eisen die het stelt en de daarmee samenhangende moeilijkheidsgraad.

De beoordeling van het goede antwoord is bovendien niet voorzien van "beoordelingsruis".

Deze overwegingen leiden tot de gedachte dat men onderwijsdoelstellingen met bijbehorende minimeisen alleen met redelijke duidelijkheid en scherpte kan definiëren indien men dat doet in termen van items en itemverzamelingen (De Groot, pag. 28).

Nu hoort men vaak de opmerking: "ja maar bij een meerkeuzevraag gaat het alleen maar om het kiezen of herkennen van het goede antwoord; open vragen toetsen toch heel wat meer dan meerkeuze-vragen. Met open vragen kunnen creativiteit en inventiviteit van de leerlingen worden gemeten.

Het gaat bij een meerkeuze-vraag wel om wat meer dan het even kiezen of herkennen van het juiste antwoord, dat leren de resultaten wel. Ook bij een goede meerkeuze-vraag moet de leerling een beroep doen op zijn creativiteit en inventiviteit om het aangeboden probleem tot oplossing te brengen.

Daarbij heeft de leerling bij de beantwoording van meerkeuze-vragen nog een extra moeilijkheid te overwinnen. Hij moet een aantal alternatieven en beschouwingen lezen en deze op hun al dan niet juist zijn beoordelen. Deze alternatieven zijn argumenten, stellingen of beweringen die hem van buiten af worden toegereikt. Zij komen veelal niet in de gedachten van de leerling op als hij zelf een antwoord moet formuleren op een in alle talen zwijgend blank stuk papier. Nu moet hij ze beoordelen, ze kunnen hem in twijfel brengen, hoe zeker is hij van zijn zaken, de leerling moet sterk in zijn schoenen staan.

Niemand zal de betekenis van de "open" samengestelde opgaven tijdens het onderwijs in twijfel willen trekken. Het is van groot belang dat men leert een antwoord correct te formuleren en zelf een exposé op schrift te stellen.

Maar het is veelal niet gemakkelijk daarbij duidelijke criteria aan te geven op grond waarvan de creativiteit valt te meten bij de beoordeling van een geschreven formulering als antwoord op een open vraag. Die creativiteit wordt vaak overtrokken en meestal zeer verschillend beoordeeld!

Wat de ene beoordelaar zeer belangrijk vindt, spreekt de ander veel minder aan.

Ik wil een en ander toelichten aan de hand van enige voorbeelden ontleend aan het boekje "een en ander over examineren en de Multiple Choice methode"\* waarin een aantal gegevens zijn opgenomen uit een artikel van Drs. C.Souren (aprilnummer, 2e jaargang: van maandblad "Richting" (1948)):

"Onderzoeken naar de mate van subjectiviteit bij de beoordeling van schriftelijk werk werden al omstreeks 1925 in het buitenland verricht. De resultaten waren niet minder dan verbazingwekkend, om niet te zeggen overdonderend.

In de Verenigde Staten (Starch en Elliott) werden aan 144 leerkrachten twee taalproefwerken ter beoordeling gegeven. Verzocht werd punten te geven tussen 0 en 100. Het ging dus met de "goudschaal". Uit het resultaat bleek, dat het proefwerk van leerling A. als laagste beoordeling 68 punten kreeg; als hoogste beoordeling 98 punten. Zou men aan 100 punten het cijfer 10 toekennen dan betekent dit, dat de laagste beoordeling 7-, en de hoogste beoordeling 10- zou zijn geweest. Frappanter was de beoordeling van leerling B. Voor deze liep de beoordeling uiteen van 50 punten tot 95 punten. Deze leerling zou volgens onze normen van sommige examinatoren een 5 (dus onvoldoende) hebben gekregen, van andere een 9½ (ongeveer uitmuntend...).

In België vonden wat later dergelijke experimenten plaats.

Zo werd aan 17 examinatoren een meetkunde-proefwerk ter beoordeling gegeven. De te gebruiken beoordelingsschaal liep van 0 tot 30 punten. Het onderzoek betrof het werk van twee leerlingen.

Voor leerling I was het laagste oordeel 16, het hoogste 27 punten (omgerekend in de 1-10 schaal: laagste 5+, hoogste 9).

Leerling II kreeg als laagste oordeel 7 punten; als hoogste 24½ punt (in de 1-10 schaal het verschil tussen 2+ en een 8+).

Aan leerling I werden 10 verschillende cijfers gegeven; aan leerling II zelfs 12...

Tijdens een internationaal pedagogisch congres dat in 1935 in Engeland (Folkestone) werd gehouden, experimenteerde men op hetzelfde gebied. Men stelde twee "examencommissies" van elk zes examinatoren en een voorzitter samen. Aan de beide commissies werden 30 proefwerken scheikunde voorgelegd.

Voordat de commissies tot beoordelen overgingen, stelde elk der groepen een aantal richtlijnen vast, die bij het beoordelen dienden te worden gevolgd. De commissieleden waren dus niet meer geheel vrij uitsluitend hun eigen maatstaven toe te passen. De te beoordelen werken moesten in vier categorieën worden ondergebracht:  
uitstekend - goed - voldoende - onvoldoende.



Ondanks afspraken en richtlijnen deden zich nog zeer opmerkelijke verschillen in beoordeling voor.

Tot de categorie "uitstekend" behoorden volgens enige beoordelaars geen enkel werk; volgens anderen echter 8 van de 30 werken; voor de categorie "goed" waren deze cijfers 9 en 16; "voldoende" waren volgens sommigen 2 werken; volgens anderen 11; het predicaat "onvoldoende" moest volgens sommige examinatoren aan 5 worden toegekend; anderen waren van mening dat 10 werken niet meer verdienen.

Hoewel men dus de zeer juiste maatregel had genomen om vooraf enige afspraken te maken met betrekking tot de criteria op grond waarvan men de kwaliteit van het werk zou beoordelen, liepen de meningen van deze deskundigen toch nog opvallend uiteen. Een niet te gewaagde conclusie is, dat er te weinig-duidelijke criteria werden vastgesteld.

Een proefneming van enigszins andere aard werd in België door Eels genomen. Hij legde aan 61 examinatoren een stel aardrijkskunde- en geschiedenis-proefwerken voor. Ook deze werken moesten worden geclassificeerd. De onderzoeker liet het niet bij het constateren van de verschillen in beoordeling. Hij legde namelijk elf weken later aan dezelfde examinatoren nog eens dezelfde werken ter beoordeling voor. En toen bleek ...

dat bij alle examinatoren belangrijke verschillen tussen hun eerste en hun tweede beoordeling moesten worden geconstateerd."

Om tot een goede beoordeling te kunnen komen moet men verschillende toetstechnieken naast elkaar gebruiken.

De vraag "open- of meerkeuze vragen" moet worden voorafgegaan door de vraag: "welke concrete onderwijsdoelstellingen heb ik voor ogen en welke prestaties moet een leerling verrichten zodat ik weet dat die doelstellingen bereikt zijn?"

Dan moet het voorafgaande onderwijs en de vereiste vaardigheden worden geanalyseerd en aan de hand van die analyse moet worden uitgemaakt welke toetsvorm de meest geschikte is.

Of een studietoets voor examen- of evaluatiedoeleinden wordt gebruikt, altijd zal die toets relevant geachte doelstellingen van het onderwijs moeten toetsen.

Weten wat men met het onderwijs wil, is dus een voorwaarde voor het kunnen maken van studietoetsen.

Omgekeerd is het maken van studietoetsen een uitstekend middel om meer inzicht te verkrijgen over wat men wil met het onderwijs. Het maken van geschikte items en het daaruit samenstellen van studietoetsen dwingt tot bezinning over doelstellingen van het onderwijs bij constructeurs en gebruikers. Items en studietoetsen kunnen aan discussies over doelstellingen een meer concrete basis geven. Via concretisering en operationalisatie van onderwijsdoelstellingen waartoe de constructie van de studietoetsen, constructeurs en itembeoordelaars dwingt, komen ook vanzelf andere onderwijsproblemen aan de orde.

Het is zoals De Groot en vele experts met hem zeggen:

"de ontwikkeling van studietoetsen en de vernieuwingsactiviteiten in het onderwijs, kunnen elkaar wederzijds bevorderen en hierin ligt misschien wel de belangrijkste waarde van de studietoetsmethode".

4. Voordracht Ten Brinke:

Ontwikkelingen in het evaluatiedenken en wat we daar als leraar mee zouden kunnen doen.

(samenvatting)

4.0. Ik zal eerst kort met u nagaan wat zich op het terrein van het evaluatiedenken heeft afgespeeld, en daarna wat we daar als leraar aan zouden kunnen hebben.

4.1. Evaluatie wordt nu door velen als een zeer belangrijke soort van menselijk handelen herkend en erkend. Individuele onderwijskundigen -vooral in de Verenigde Staten- zijn hun daarin voorgegaan. Bij Mursell (1946) b.v. is evaluatie één van de 6 basisprincipes waarop goed onderwijs gebaseerd moet zijn en hij wijdt er 2 hele hoofdstukken aan.

Dat evaluatie nu zo'n "rage" geworden is, moet diepere oorzaken hebben. Ik vermoed dat een grotere drang tot het afleggen van verantwoording de, of een, diepere oorzaak is. In de twee gebieden waar de "rage" is ontstaan, dat van de onderwijsleerprogramma-ontwikkeling en van de groepsdynamica, wordt het afleggen van verantwoording inderdaad als een noodzaak gezien. In het eerste gebied wordt dit veroorzaakt door het vele geld dat ermee gemoeid is en de macht die de ontwikkelaars bezitten. In het tweede gebied is het de coöperatieve houding die het wederzijds afleggen van verantwoording noodzakelijk maakt.

Er heeft inmiddels, door de jaren heen, een diepgaande discussie plaatsgevonden over wat evalueren eigenlijk is. Van huis uit betekent het "de waarde van iets bepalen" (dus in het geval van onderwijs: de waarde van onderwijs bepalen). Maar we zien dat dit begrip in de literatuur op allerlei manieren is beperkt.

De eerste beperking die genoemd kan worden, is:

evalueren=meten. Deze beperking is natuurlijk afkomstig uit de natuurwetenschappelijk-georiënteerde hoek van de sociale wetenschappen. Het is een zeer schadelijke beperking: immers bij veel belangrijke beslissingen wordt geëvalueerd (d.w.z. de waarde van iets wordt bepaald) op grond van heel andere dingen dan metingen. Deze beperking heeft, voor zover ik weet, ook niet veel aanhangers meer.

Een tweede besprekenswaardige beperking is: evalueren=nagaan of je je onderwijsdoel bereikt hebt.

Door deze beperking kun je alleen maar aan het eind van een of ander proces evalueren, niet voorafgaand aan of tijdens.

Hieraan kan worden tegemoetgekomen door de definitie aldus uit te breiden: evalueren = nagaan of je je doel bereikt hebt (summatieve evaluatie) of aan het bereiken bent (formatieve evaluatie). Zelfs in deze uitgebreide vorm is de definitie echter nog veel te beperkt (en deze beperktheid maakt hem wederom tot een "schadelijke" definitie). Hij heeft de volgende nadelen:

- 1e. hij omvat niet een zeer essentiële vorm van evaluatie, n.l. de evaluatie van het onderwijsdoel, of de onderwijsdoelen;
- 2e. bij onderwijs kunnen onvoorspelde en, sterker, onvoorspelbare resultaten bereikt worden (de rol die dit verschijnsel speelt, kan van vak tot vak verschillen);
- 3e. bij het beoordelen van onderwijs zal men ook verschijnselen die niets met het bereiken van onderwijsdoelstellingen te maken hebben, willen evalueren, b.v. de kosten die iets meebrengt, de stress die iets bij de leraar veroorzaakt.

Deze bezwaren worden weggenomen door de ruime definitie opgesteld door het team rondom Stufflebeam: (in verkorte vorm) evalueren is het bijbrengen van informatie ten behoeve van beslissingen. Mijn persoonlijke ervaringen met deze definitie in de praktijk zijn gunstig. Toch hebben critici aangetoond dat zelfs deze ruime definitie nog een ongeoorloofde beperking inhoudt. Inderdaad zijn er gevallen waarin "waardebepaling" optreedt zonder dat dit gekoppeld is aan een beslissing. Het doel van de evaluatie kan dan zijn iets als bewustworden. In kringen van de juistgenoemde critici is voorgesteld, evalueren simpelweg te definiëren als: de waarde van iets bepalen. Daarmee zijn we terug bij de betekenis die evalueren al in het woordenboek van het Engels had sinds het daar vóór het eerst in werd opgenomen!

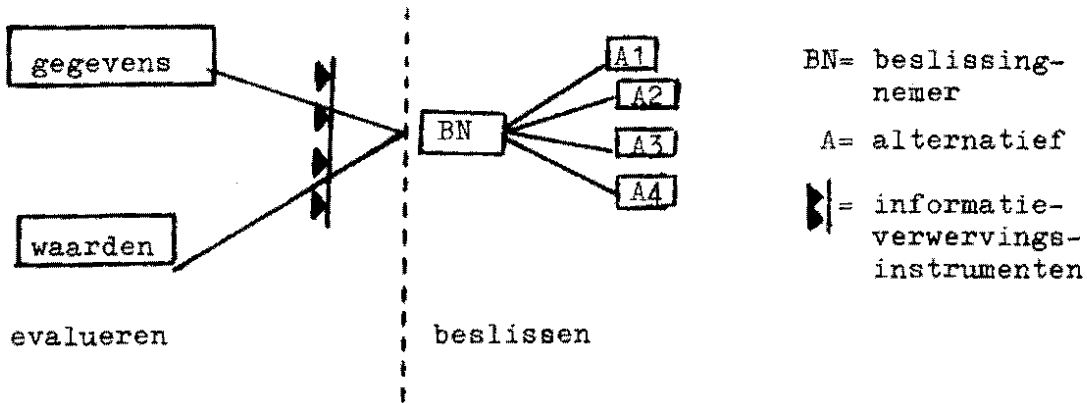
Als wij goed luisteren kunnen wij Erasmus horen grijnslachen.

Alvorens over te gaan tot het tweede deel van mijn inleiding wil ik nog een beperkte definitie signaleren, die niet in de literatuur maar wel in het spraakgebruik in Nederland voorkomt. Volgens deze houdt evalueren (aan het eind van een vergadering) in:

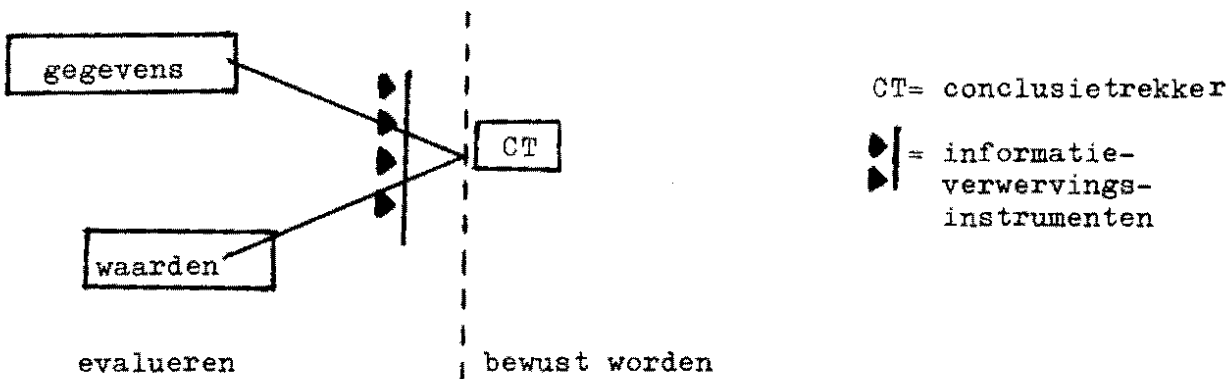
"zeggen hoe je je voelt." Deze inperking moet wel het gevolg zijn van een eenzijdige toepassing van het Feelings Are Facts principe dat in het gebied van de groepsdynamica een voorname, maar uiteraard niet de enige, rol speelt.

- 4.2. Het gewroet in de twee genoemde terreinen (onderwijsleerprogramma-ontwikkeling en groepsdynamica) heeft natuurlijk méér dan alleen maar een woordenboek-definitie van evalueren opgeleverd. Er zijn een aantal denkmodellen en procedures ontwikkeld die ook voor het dagelijks onderwijs van groot belang kunnen zijn. Ik zal trachten dit exemplarisch aan te tonen door een klein aantal bruikbare ideeën te bespreken (volledigheid is bij deze korte inleiding uiteraard onmogelijk).

Een ware schatkamer is het juistgenoemde werk van de groep van Stufflebeam. Een beslissingsmodel dat zij hanteren ziet er-in door mij enigszins aangepaste vorm-zo uit:



voor een situatie waarin geen sprake is van een beslissing, is het volgende gereduceerde model bruikbaar:



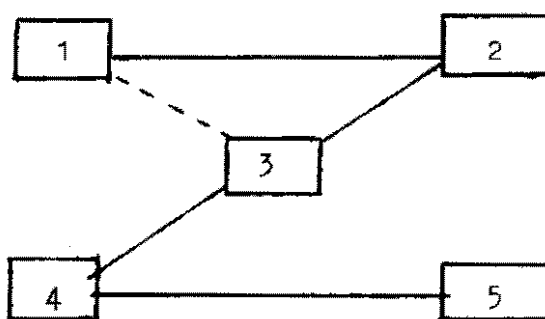
Als dit model in de praktijk wordt toegepast, ontstaan allerlei speciale problematieken. Voor vele aspecten daarvan zijn door de groep-Stafflebeam, en door anderen, speciale denkmodellen en procedures ontwikkeld.

Ik beperk me tot één aspect: het feit dat beslissingnemers, en conclusietrekkers die in het onderwijs verschillende rollen vervullen (b.v. leraren en beleidsvormers) verschillende informatie nodig hebben. Daarom is de onderstaande matrix van belang (waarin ik een paar categorieën BN's en CT's heb ingevuld waarmee ik me dadelijk wat nader zal bezig houden):

WAAR- OVER?	WIE?	beleids- vormers	leraren	leer- lingen	ouders	enz.

Laten we nu eens nagaan wat we met deze twee grondgedachten in onze onderwijspraktijk zouden kunnen doen. Daarvoor is nodig dat we in de matrix sub waarover een bepaalde vraag invullen. Ik kies daarvoor de vraag: "Hebben de leerlingen voldoende relevante dingen geleerd?" En om aan te sluiten op wat tijdens het voorafgaande deel van deze conferentie is behandeld, stel ik het probleem: in hoeverre zijn nu centrale eindexamens en schoolonderzoek bruikbare informatie-verwerkings-instrumenten om deze vraag te beantwoorden?  
Ik kom tot de volgende punten.

1. Het beslissingsmodel leert ons dat zonder een diepgaande discussie over onderliggende waarden geen serieuze evaluatie mogelijk is. Ook in het natuurkundeonderwijs is in die zin pluriformiteit aanwezig dat verschillende, eventueel tegengestelde, opvattingen mogelijk zijn over het belang van streng-exact denken, creativiteit, met-je-handen-denken, enz. In tegenstelling met wat algemeen gedacht en gedaan wordt is het dus onmogelijk de vraag of de leerlingen voldoende relevante dingen geleerd hebben, regelrecht uit de resultaten van centraal examen en schoolonderzoek af te leiden.
2. Een bruikbare procedure waarin gegevens en waarden t.a.v. iemands onderwijs betrokken kunnen worden, lijkt de zg. Z-procedure te zijn (Nederlands Pedagogisch Instituut, Zeist). In voor dit doel aangepaste vorm ziet deze procedure er aldus uit:



- 1= Van welke principes probeer(de) ik uit te gaan?
- 2= Wat doe ik in werkelijkheid?
- 3= Van welke principes ga ik dus blijkbaar echt uit?
- 4= Van welke principes zou ik echt uit willen gaan?
- 5= Wat moet ik dan in werkelijkheid doen?

Mijn persoonlijke ervaring met deze procedure is, dat hij veel tijd neemt maar ook zeer belangrijke resultaten oplevert (het vormgeven waarvan jaren vergt!). Deze procedure moet m.i. dan ook slechts eens-per-zoveel jaar worden toegepast.

3. Centraal examen en schoolonderzoek zijn des te relevantere informatie-verwervings-instrumenten als onderwijsresultaten, die hoog in Taxonomy I liggen en die in Taxonomy II genoemd worden, mede worden "gevangen".

T.a.v. pogingen om de eerstgenoemde onderwijsresultaten in het schoolonderzoek te betrekken heeft Stellen~~en~~ met een en ander opgemerkt.

Het betrekken van de tweede groep, de zg. affectieve onderwijsresultaten, in het examen: het feit dat de beslissing die op de evaluatie volgt, een selectieve is maakt, dat het moeilijk zal zijn op dit terrein betrouwbare gegevens te krijgen. Eén van die problemen is b.v.: als een leerling weet dat zijn leraar "het plezier hebben in experimenteren" erg belangrijk vindt, dan zal hij zorgen dat hij zich als zodanig gedraagt, of hij er nu echt plezier in heeft of niet. En neem hem dat eens kwalijk!

4. Laten we nog eens naar de matrix kijken. Tot dusver hebben we het uitsluitend-en dit geldt ook voor de conferentie als geheel-over de leraar als BN of CT gehad. Maar de matrix brengt als zodanig ook de leerling en de ouders binnen ons gezichtsveld. Het is tekenend voor het onderwijsklimaat waarin we leven dat met name de leerling als BN(CT) nog steeds zo'n ondergeschikte rol inneemt. Persoonlijk heeft tijdens mijn eigen middelbaar onderwijs geen leraar mij gevraagd eens rustig en eerlijk op te schrijven hoe ik tegenover zijn vak en tegenover zijn onderwijs stond, om daar dan met mij eens diepgaand over te praten. Ik heb het zelf als leraar a.v.o. ook niet gedaan(maar als leraar w.o. wel), en ook binnen de leraarsopleiding speelt het, voor zover ik weet, nog nauwelijks een rol.

Naar mijn eigen ervaring (zoals gezegd in het w.o.) draagt serieuze evaluatie door de leerling op spectaculaire wijze bij aan de kwaliteit van het onderwijs dat men geeft (en natuurlijk ook tot de verhouding leraar-leerling). Daar de ervaringen aanvankelijk schokkend kunnen zijn, zou men de juistgenoemde werkwijze m.i. op een zeer geleidelijke manier moeten invoeren.

N.B. Voor literair-geïnteresseerden onder u zullen de ervaringen van de leerling Miles bij zijn scheikundeleraar belangwekkend kunnen zijn, zoals Snow die beschrijft in "The Search".

5. En de ouders als BN en CT? Laten we nog eenmaal naar het beslissingsmodel kijken, en nu op de plaats van BN of CT "ouder" invullen. Volgens het model heeft hij recht op informatie, betrekking hebbende op gegevens en waarden. En met wat voor informatie moet hij het nu gewoonlijk doen: een cijfer! Primitiever kan het niet. U zou dit zelf als particulier in welke beslissingssituatie dan ook niet accepteren. B.v.: u vraagt een reisbureau wat voor aardige reizen ze hebben, en men antwoordt: "We hebben er één waar we een 9 voor gaven; neemt u die maar!". Het evaluatiedenken leert ons dus dat wij, zo handelende, ouders in een absurde situatie brengen. Wat meer is, we versterken daarvoor een denkwijze volgens welke het enige waar het op aankomt een hoog cijfer is. Als we daar niets aan doen, kunnen we beter ophouden naar waardevoller onderwijs te streven! Er zijn altijd al scholen geweest die méér dan cijfer-rapporten aanboden. Hun voorbeeld verdient dringend navolging. Omdat in veel scholen de te verstrekken informatie geschikt moet zijn voor administratieve verwerking door b.v. bepaalde coördinatoren, zal een zekere structurering dikwijls wel noodzakelijk zijn. Men kan dan b.v. per vak een aantal belangrijke categorieën van onderwijsresultaten aangeven (vaardigheid in X, belangstelling voor Y, basis-kennis, enz.), en daar simpele schaalverdelingen in aanbrengen. Aan dit geheel kan dan nog een categorie opmerkingen worden toegevoegd.
- 4.3. Volgens de afspraak heb ik slechts een zeer klein deel van de binnen het evaluatiedenken ontwikkelde inzichten besproken. Ik hoop te hebben duidelijk gemaakt dat er in de dagelijkse onderwijspraktijk veel mee te doen is.

#### Genoemde Literatuur

Educational evaluation and decision making. Written bij PDK  
National Study Committee on Evaluation (chairman: Daniel L. Stufflebeam)  
Itasca, Illinois, 1971

Mursell, James L. Successful teaching. Its psychological principles  
New York etc. 1946

Snow, C.P. The Search. New York, 1934 (ook in Signet)

Taxonomy of educational objectives  
etc. Handbook I : Cognitive domain  
Handbook II : Affective domain  
New York, 1956, 1964

INHOUD

	blz.
1. Analyse C.M.L.N.-doelstellingen	55
1.1. Vraagstelling	55
1.2. Materiaal	55
1.3. Verslagen:	
1.3.1. Groep Van Genderen	57
1.3.2. Groep Van Ark	58
2. Analyse van vragen met behulp van de categorieën van Klopfer	59
2.1. Vraagstelling	59
2.2. Materiaal:	
2.2.1. Methoden ter analyse van examenvragen	59
2.2.2. Proefwerk	65
2.2.3. Het proefwerk en de categorieën van Klopfer	68
2.3. Verslagen:	
2.3.1. Groep Jetses	69
2.3.2. Groep Van der Rijst	69
2.3.3. Voorgroep De Mink	70
3. De categorieën van Klopfer concreet in de lespraktijk	71
3.1. Vraagstelling	71
3.2. Materiaal	71
3.3. Verslag groep Veth	71
4. Het maken van vragen met behulp van de categorieën van Klopfer	72
4.1. Vraagstelling	72
4.2. Materiaal	72
4.3. Verslag groep Heij	72
5. Cijfer geven	74
5.1. Vraagstelling:	
5.1.1. Voorgeschiedenis	74
5.1.2. Gebruik van de meerkeuzevragen	74
5.1.3. Enige achtergrondinformatie bij de antwoorden op de vragen	74
5.2. Materiaal: 9 vragen over cijfergeving	78



5.3. Verslagen:	
5.3.1. Groep De Jong	80
5.3.2. Groep De Mink	80
5.3.3. Groep Van der Hilst	81
5.3.4. Groep Van Ark	81
6. Praktikumbeoordeling	82
6.1. Vraagstelling	82
6.2. Materiaal	82
6.3. Verslag groep Verhagen	82
7. Beoordeling van leerlingen	83
7.1. Vraagstelling	83
7.2. Materiaal	84
7.3. Verslag groep Seller	87
8. Itemanalyse	88
8.1. Vraagstelling	88
8.2. Materiaal:	
8.2.1. Begrippenlijst itemanalyse	89
8.2.2. De te beoordelen items	92
8.3. Verslagen:	
8.3.1. Groep Schröder	99
8.3.2. Groep Vervoort	99
8.3.3. Groep Van Genderen	99
8.3.4. Groep Jetses	103
8.3.5. Groep De Jong	103
9. Evalueren in een breed kader	104
9.1. Vraagstelling	104
9.2. Materiaal	105
9.3. Verslag groep Verhagen	105
10. Evalueren van complexe natuurwetenschappelijke doelen	105
10.1. Vraagstelling	105
10.2. Materiaal	106
10.3. Verslag groep De Mink	106
11. Problemen bij beoordelingslijst van leerlingen	107
11.1. Vraagstelling	107
11.2. Materiaal	107
11.3. Verslag groep Veth	107

### DEEL 3: DE WERKZAAMHEDEN IN DE DISKUSSIEGROEPEN

#### 1. Analyse C.M.L.N.-doelstellingen

##### 1.1. Vraagstelling

1.1.1. Geef voor de in de bijlage geformuleerde doelstellingen (uit het C.M.L.N.-rapport) aan:

- bij welke doelen het aksent ligt op de voorbereiding voor verdere studie en
- bij welke doelen het aksent ligt op algemene vorming met betrekking tot de natuurwetenschappen.

1.1.2. Aan welke van deze doelen moet volgens U de meeste aandacht worden besteed?

1.1.3. Vindt U dat er in deze lijst nog belangrijke doelen ontbreken?

##### 1.2. Materiaal

Uit: Commissie Modernisering Leerplan Natuurkunde, rapport 1974, deel VII

#### VOORSTELLEN BETREFFENDE DE EXAMENPROGRAMMA'S VOOR MAVO-3, MAVO-4, HAVO EN VWO

##### Voorwoord

De in dit deelrapport vermelde onderwerpen (feiten, terminologieën, begrippen, afspraken, wetten, theorieën en toepassingen) dienen te behoren tot het kennispakket van de examenkandidaten. Voor nadere detaillering, diepgang en motivering verwijzen wij naar de betreffende deelrapporten voor de onderscheiden schooltypen.

In de examenopgaven zal in het algemeen het hanteren van eenheden uit het praktische stelsel (S.I.-eenheden) geeist worden.

Eventueel afwijkende eenheden zullen in de opgaven verklaard moeten worden. Herhaaldelijk is achter een onderwerp tussen haakjes de toevoeging 'kwalitatief' geplaatst. Het is daarbij de bedoeling niet een volledige wiskundige beschrijving te geven van de verschijnselen, maar wel om de onderlinge verbanden duidelijk te laten uitkomen - dus of een grootheid groter/kleiner of meer/minder wordt bij een verandering van een ermee samenhangende grootheid. Als voorbeeld kan hier genoemd worden de huiging van licht aan een spleet, waar wel het verband: kleinere spleetwijdte → grotere afbuighoek gegeven moet worden (liefst uit experimenten), maar geen berekening van het kwantitatieve verband. Vaak zal bij de uitleg gebruik gemaakt kunnen worden van een schets of van een grafiek ter verduidelijking van de bespreking.

Aan de hand van het kennispakket zal in het examen worden nagegaan in hoeverre de kandidaten vaardigheid vertonen in het gebruik van deze kennis.

van MAVO-kandidaten wordt verwacht, dat zij:

- A.1. bij een beschrijving van een fysisch verschijnsel kunnen herkennen met welke begrippen of wetten het verschijnsel in verband gebracht kan worden;

- B.1. indien de uitkomsten van een experiment in tabelvorm zijn gegeven, met behulp hiervan een grafiek kunnen tekenen;
- 2. uit grafieken gegevens kunnen aflezen;
- 3. de waarde van een grootheid kunnen berekenen met behulp van een aan de leerlingen bekende relatie;
- C.1. de afloop van een variatie op een bekend experiment kunnen voorspellen met behulp van bekende begrippen en wetten;
- 2. kunnen nagaan of de uitkomsten van een variatie op een bekend experiment in overeenstemming zijn met bekende wetten;
- 3. een bekend verschijnsel of experiment met behulp van een bekende theorie kunnen verklaren;
- D.1. aan de hand van bekende wetten kunnen nagaan welke van enkele gegeven voorspellingen over de afloop van een tevoren niet bekend experiment in overeenstemming zal zijn met de werkelijke afloop ervan.

Van HAVO-kandidaten wordt bovendien verwacht, dat zij:

- B.4. gegevens uit enkele grafieken met elkaar in verband kunnen brengen;
- 5. een in wiskundevorm gegeven relatie kunnen overbrengen in een grafiek;
- 6. in niet te gecompliceerde gevallen een door een grafiek voorgestelde relatie kunnen vertalen in formulevorm;
- 7. gegevens die voor het oplossen van een probleem nodig zijn kunnen selecteren en, zo nodig door berekening, een probleem ermee kunnen oplossen;
- 8. kunnen interpoleren en extrapoleren en in het laatste geval inzien, wanneer zij daarbij extra voorzichtig moeten zijn;
- 9. bij berekeningen waarbij gebruik wordt gemaakt van meetwaarden, de grootte-orde van de onzekerheid in de uitkomst kunnen schatten;
- D.2. de afloop van een nieuw experiment kunnen voorspellen, gebruik makend van bekende begrippen en wetten;
- 3. kunnen aangeven hoe een gegeven veronderstelling op een binnen de leerstof gelegen terrein experimenteel getoetst zou kunnen worden;
- 4. een veronderstelling op een binnen de leerstof gelegen terrein experimenteel kunnen verifiëren (deze eis moet in het schoolonderzoek worden getoetst);
- 5. aan de hand van bekende wetten nagaan of een gegeven voorspelling over de afloop van een tevoren niet bekend experiment al dan niet in overeenstemming is met de werkelijke afloop ervan.

Van VWO-kandidaten wordt voorts verwacht, dat zij:

- E.1. de juistheid van onbekende formules kunnen toetsen met behulp van een dimensieberekening;
- 2. kunnen nagaan of door een meetexperiment op grond van de daarbij bereikte nauwkeurigheid een gegeven relatie wordt bevestigd of weersproken;
- 3. kunnen nagaan of een veronderstelling logisch uit een bekende theorie voortvloeit;
- 4. de afloop van een nieuw experiment met behulp van een bekende theorie kunnen voorspellen;
- 5. kunnen nagaan of de uitkomst van een toetsingsexperiment een theorie ondersteunt of (gedeeltelijk) weerlegt;
- 6. kleine veranderingen in een bekende theorie kunnen formuleren die de uitkomst van een toetsingsexperiment verklaren;
- 7. een eenvoudige redenering kunnen opbouwen op een onbekend terrein der natuurkunde indien de structuur daarvan analoog is aan die van een bekend terrein, onder uitdrukkelijke verwijzing naar die analogie.

### 1.3. Verslagen

#### 1.3.1. *Groep van Genderen*

Over de genoemde doelen werden de volgende opmerkingen gemaakt:

- Deze doelstellingen moeten nog operationeel gemaakt worden voor de leraar. Er moeten bijvoorbeeld vraagstukken gemaakt worden, die duidelijk maken wat de verschillende doelen in de praktijk van een leerling eisen. Dit is een zeer moeilijke opgave.
- De doelstelling B7: het selecteren van gegevens is voor verschillende uitleg vatbaar. Is het alleen het kunnen gebruiken van het tabellenboek? Of is het ook het selecteren van aanwezig veronderstelde kennis, of het selecteren van fysische aspecten uit het geheel van het gebeuren?

Het valt op dat deze doelstelling niet bij de MAVO voorkomt. In een eenvoudige kontekst zou dit bij het MAVO-examen getest moeten worden.

In de lijst van doelstellingen komen de volgende punten niet of niet voldoende aan bod:

- Het praktikum op de MAVO. Dit blijkt op dit moment om praktische en organisatorische redenen niet mogelijk te zijn.
- Het creatieve element op de MAVO ontbreekt. We denken b.v. aan het verbeteren van een experiment, het zelf bedenken van een experiment.
- Op HAVO en VWO ontbreekt ook het uit experimentele gegevens kunnen opbouwen van een model. Hiervoor moeten nieuwe toetsingsprocedures ontwikkeld worden.

Er zijn twee wegen om de nadruk op de verschillende onderwijsdoelen te verschuiven.

1. De examendoelen veranderen. Dan moet het onderwijs zich wel aanpassen.
2. Methoden voor het bereiken van en toetsen van nieuwe onder-

Want de docenten willen het wel maar kunnen het niet alleen.

Conclusie: De huidige toetsingsmethoden zijn niet geschikt om de eigenlijke onderwijsdoelen te testen. De proef met schaduwvlek en lichtvlek van Steller is in feite een in beeld gebracht, streng begrensd vraagstuk, van het soort dat wij al kennen.

Het is wel de vraag of alle onderwijsdoelen in het eindexamen getest moeten worden.

1.3.2. *Groep van Ark*

1. In de groep was geen duidelijkheid over de betekenis van de begrippen, algemeen vormend (a) en voorbereidend (v).

Besloten werd individueel een voorkeur, a of v, uit te spreken ten aanzien van de doelen (zie bijl.1, hieronder).

Uit de hierop volgende discussie bleek het in kort bestek onmogelijk a en v te beschrijven. (Genoemd: Algemeen vormende doelen zijn doelen die voor alle leerlingen van belang zijn, onafhankelijk van de vervolgopleiding)

2. Toevoegingen

Gemist worden doelen met betrekking tot de plaats van de natuurkunde in de samenleving en doelen met betrekking tot experimentele en praktische vaardigheden.

3. Opmerkingen

- a) De genoemde doelen en de huidige examens dekken elkaar niet.
- b) De doelen moeten worden geïnstrumentaliseerd.

Bijlage 1

	a	v		a	v		a	v		a	v			
A <sub>1</sub>	8	3	C <sub>2</sub>	1	10	B <sub>6</sub>	2	9	D <sub>3</sub>	0	11	E <sub>3</sub>	7	4
B <sub>1</sub>	6	5	C <sub>3</sub>	7	4	B <sub>7</sub>	3	8	D <sub>4</sub>	0	11	E <sub>4</sub>	2	9
B <sub>2</sub>	7	4	D <sub>1</sub>	1	10	B <sub>8</sub>	7	4	D <sub>5</sub>	0	11	E <sub>5</sub>	5	6
B <sub>3</sub>	3	8	B <sub>4</sub>	8	3	B <sub>9</sub>	3	8	E <sub>1</sub>	1	10	E <sub>6</sub>	2	9
C <sub>1</sub>	0	11	B <sub>5</sub>	3	8	D <sub>2</sub>	0	11	E <sub>2</sub>	1	10	E <sub>7</sub>	3	8

De doelstellingen A<sub>1</sub> enz. vindt U onder 1.2. (ter conferentie uitgereikt materiaal), op blz. 55/56.

## 2. Analyse van vragen met behulp van de categorieën van Klopfer

### 2.1. Vraagstelling

Analyseer het reeds toegezonden schoolexamen en de praktikum-opdracht uit Steller's voordracht met behulp van de categorieën van Klopfer.

### 2.2. Materiaal

#### 2.2.1. Methoden ter analyse van examenvragen

Uit: "Scheidende wegen?" Dr. S. Auer

Analyse van tot dusverre samengestelde examenopgaven kan twee doeleinden dienen.

- 1 De analyse kan bijdragen tot de bepaling van het niveau van de examens en daarmee een discussie over de met deze onderwijsvorm te bereiken doeleinden vruchtbaarder maken.
- 2 De analyse kan aanknopingspunten bieden voor heroriëntatie van het gegeven onderwijs.

Examenvragen kunnen op verschillende wijzen worden geanalyseerd. Voor een in hoofdzaak in termen van leerstof omschreven examenprogramma is een inhoudsanalyse een logische zaak. Hiermee kan immers worden vastgesteld of de in de examens gevraagde leerstof dezelfde is als die welke in het programma omschreven werd.

Deze controle is niet in beide richtingen even belangrijk. Er wordt bij de samenstelling vrij rigoreus op toegezien, dat de examenvragen binnen de examenstof blijven, al komen af en toe kleine grensoverschrijdingen voor. Niet ongevaarlijk is trouwens ook de omgekeerde situatie: het niet in recente examens voorkomen van een aanzienlijk gedeelte der examenstof. Zolang deze handelwijze impliciet blijft, kunnen leerlingen (en docenten?) hun examenvoorbereiding er op af gaan stemmen, terwijl anderzijds de samenstellers van nieuwe opgaven plotseling in ruime mate uit die tot dan toe verwaarloosde stof kunnen gaan putten. Een inhoudsanalyse van een aantal opeenvolgende examens kan zulke verschillen aan het licht brengen. In het onderhavige geval kan bovendien worden nagegaan of eventuele afwijkingen in de richting gaan van het door de subcommissie bovenbouw-havo van C.M.L.N. uitgebrachte interimrapport (9). Deze vergelijking is te interessanter daar bij de discussies over dit rapport tijdens regionale Velnes-vergaderingen op de leerstofflijst geen fundamentele kritiek is geleverd, maar bij de commissie alleen (!) is aangedrongen op verdergaande matiging in de omvang van het programma. Omdat nog niet bekend is in hoever-

re bovengenoemde subcommissie deze kritiek zal verwerken, is een dergelijke vergelijking hier achterwege gelaten.

Vooraf voor de vormgeving van het onderwijs is een anderssoortige analyse belangrijk. Daarbij staat niet langer de leerstof zelf centraal, maar de wijze waarop de examenkandidaat met deze leerstof kan werken.

Zo is in het examenprogramma onder het hoofd *roosterstructuur* opgenomen:

geleiders en isolatoren; wet van Ohm; afhankelijkheid van de weerstand van materiaal, lengte, doorsnede en temperatuur; substitutieweerstand; meting van weerstand en spanning; verandering van het meetgebied van een stroommeter en een spanningsmeter; spanningsdeler/potentimeter; warmteontwikkeling door een stroom; wet van Joule.

Een vraag als: 'Hoe luidt de wet van Ohm?' of 'Hoe kan het meetgebied van een stroommeter worden veranderd?' doet slechts een beroep op reproductie van (een deel van) de geciteerde leerstof.

De beantwoording van de vraag: 'een stroommeter met een weerstand van  $0,10\Omega$  heeft een meetbereik van 10 m A.

a kan aan deze meter een meetbereik van 1 A gegeven worden?

b kan aan deze meter een meetbereik van 1 m A gegeven worden?

Zo ja, hoe, zo nee, waarom niet?'

vraagt van de leerling in de eerste plaats inzicht in het probleem 'verandering van meetgebied', daarnaast toepassing van de wet van Ohm.

Weer zwaardere eisen stelt beantwoording van de vraag: 'Op de voorkant van een geheel gesloten kistje zitten twee elektrische aansluitingen. Iemand neemt het kistje op in de getekende schakeling (fig.) en meet bij verschillende standen van de spanningsdeler de spanning over de aansluitpunten en de stroomsterkte. Hij verkrijgt de volgende meetwaarden (tabel) en concludeert dat tussen de aansluitpunten

een metaaldraad is gespannen. Is deze conclusie volgens jou:

- a beslist onjuist
- b waarschijnlijk onjuist
- c misschien juist, misschien onjuist
- d waarschijnlijk juist
- e zeker juist?

Motiveer je keuze.'

In feite zou de kandidaten niet alleen bekend moeten zijn welke stof op het examen kan worden gevraagd, maar ook hoe zij deze leerstof moeten kunnen gebruiken. Daartoe zouden de doelstellingen van het onderwijs moeten zijn geoperationaliseerd, d.w.z. op een zodanige wijze omschreven dat de kandidaten daar direct uit kunnen afleiden welke eisen het examen stelt.

Indien een dergelijke reeks doelstellingen voor het nederlandse natuurkundeonderwijs zou zijn geformuleerd, dan zou daaruit vermoedelijk blijken dat ze niet alle dezelfde mate van compleetheid en/of abstractie zouden vereisen. Onderwijsniveaus zouden wellicht mede kunnen worden onderscheiden door de complexiteit van het van de leerlingen verwachte eindgedrag. Om dergelijk werk vruchtbaar te kunnen verrichten bestaat er behoefte aan een ordeningssysteem waarmee men doelstellingen kan categoriseren.

Een bekend ordeningssysteem is dat van Bloom *e.s.*, dat in Nederland bij de analyse van examenvragen meermalen werd gebruikt. In de cognitieve taxonomie van Bloom (10) worden onderwijsdoelen in zes hoofdcategorieën gerangschikt, te weten:

- 1.00 kennis (knowledge)
- 2.00 begrip (comprehension)
- 3.00 toepassing (application)
- 4.00 analyse (analysis)
- 5.00 synthese (synthesis)
- 6.00 evaluatie (evaluation)

Deze categorieën zijn dan onderverdeeld in subcategorieën, die vooral bij de laagste categorieën het abstractieniveau van de in de doelstellingen omschreven of in toetsvragen verwerkte leerstof aanduiden. Minder ver ging bij de analyse van het hbs-examen scheikunde Kollaard (11). Hij gebruikte drie subcategorieën, A, B en C, die hij omschreef als:

- A: kennis van feiten, concrete begrippen en konventies
- B: kennis van relaties, wetten, complexe begrippen en voorstelbare modellen
- C: kennis van theorieën, de meest complexe begrippen en algemene methodologische principes.

Zoals de omschrijving al suggereert is deze in-

deling eigenlijk slechts hanteerbaar bij de eerste categorie (kennis) van Bloom. Zeker niveau A leent zich nauwelijks voor toepassing, analyse of synthese.

Overigens vestigt Klopfer (12) de aandacht op een taxonomie voor natuurwetenschappen die, hoewel in grote lijnen overeenkomsten vertonend met die van Bloom, bruikbaar is in verband met zijn op de natuurwetenschappen toegespitst karakter, vooral wanneer men deze wetenschappen niet in de eerste plaats onderwijst als een kennisverzameling, maar als een manier om leerlingen met methoden van kennisverwerving vertrouwd te maken.

Hoewel te verwachten valt dat het resultaat teleur zal stellen, zullen deze categorieën naast die van Bloom bij de analyse van de havo-examens worden gehanteerd omdat in vele beschouwingen over de doelstellingen van onderwijs in de natuurwetenschappen de hier vermelde categorieën steeds weer meer of minder expliciet voorkomen. Ook hier geldt immers weer dat een kloof tussen wat we doen en wat we willen gesignaleerd moet worden. In de eerste plaats om degenen die zich op de examens voorbereiden en hun begeleiders hulp te bieden bij hun taak. Maar daarnaast ook om wegen te zoeken waardoor onze bedoelingen en de uitkomsten van de op grond daarvan verrichte arbeid meer in overeenstemming met elkaar gebracht worden.

#### Literatuur:

- (9) Commissie Modernisering Leerplan Natuurkunde, deel III Havo bovenbouw, 1969
- (10) B.S. Bloom (ed) Taxonomie van Opvoedings- en Instructiedoelen, I: cognitief gebied Rotterdam 1971 (vertaling)
- (11) U.H. Kollaard: Wat meten we met het eindexamen scheikunde? Faraday, 41 (210) 1972
- (12) L.E. Klopfer: Evaluation of Learning in Science, in: B.S. Bloom, J.Th. Hastings, G. Nadaus: Handbook of Formative and Summative Evaluation of Student Learning, New York 1971

**SPECIFICATIE VAN GEDRAGSCATEGORIEËN TEN BEHOEVE VAN ONDERWIJS IN DE NATUURWETENSCHAPPEN (COGNITIEF GEDEELTE)**

**TAXONOMIE VAN KLOPFER**

- A Kennis en begrip
  - A 1 feitenkennis
  - A 2 kennis van wetenschappelijke terminologie
  - A 3 kennis van begrippen
  - A 4 kennis van conventies
  - A 5 kennis van trends en sequenties
  - A 6 kennis van classificaties, categorieën en criteria
  - A 7 kennis van wetenschappelijke technieken en procedures
  - A 8 kennis van wetenschappelijke regels en wetten
  - A 9 kennis van theorieën of belangrijke netwerken van begrippen
  - A 10 herkenning van kennis in een nieuwe samenhang
  - A 11 vertaling van kennis van de ene symbolische vorm in de andere.
- B Methoden van wetenschappelijk onderzoek: I waarnemen en meten
  - B 1 waarnemen van voorwerpen en verschijnselen
  - B 2 beschrijven van waarnemingen in geschikte taal
  - B 3 meten van objecten en veranderingen
  - B 4 keuze van geschikte meetinstrumenten
  - B 5 schatting van metingen en herkenning van beperkingen in meetnauwkeurigheid
- C Methoden van wetenschappelijk onderzoek: II een probleem zien en wegen zoeken om het op te lossen
  - C 1 herkennen van een probleem
  - C 2 formuleren van een werkhypothese
  - C 3 middelen kiezen om een hypothese te toetsen
  - C 4 geschikte wegen uitstippelen om proeven uit te voeren
- D Methoden van wetenschappelijk onderzoek: III gegevens interpreteren en generalisaties formuleren
  - D 1 experimentele gegevens verzamelen
  - D 2 de gegevens presenteren in de vorm van functionele relaties
  - D 3 interpreteren van experimentele gegevens en waarnemingen
  - D 4 extrapoleren en interpoleren
  - D 5 evaluatie van een hypothese met behulp van verkregen gegevens
  - D 6 formulering van generalisaties op grond van de verkregen relaties
- E Methoden van wetenschappelijk onderzoek: IV constructie, toetsing en herziening van een model
  - E 1 herkennen van de noodzaak van een model
  - E 2 formulering van het model
  - E 3 specificeren van de relaties die door het model verklaard worden
  - E 4 afleiden van nieuwe hypothesen uit een model
  - E 5 interpreteren en evalueren van toetsingsexperimenten
  - E 6 formuleren van een herzien, verfijnd of uitgebreid model
- F Toepassen van wetenschappelijke kennis en methoden
  - F 1 toepassen op nieuwe problemen in hetzelfde wetenschapsgebied
  - F 2 toepassen op nieuwe problemen in een ander wetenschapsgebied
  - F 3 toepassen op nieuwe problemen buiten de natuurwetenschap
- G Handvaardigheid
  - G 1 ontwikkelen van vaardigheid in het gebruik van laboratoriuminstrumentarium
  - G 2 veilig en zorgvuldig uitvoeren van normale laboratoriumtechnieken



# EINDEXAMEN VAN DE SCHOLEN VOOR HOGER ALGEMEEN VOORTGEZET ONDERWIJS IN 1969

1. In de figuren 1 en 2 zijn met de punten  $A_1$  en  $A_2$  de spanning, het volume en de temperatuur van een zekere hoeveelheid ideaal gas in een bepaalde toestand aangegeven; de punten  $B_2$  en  $C_2$  hebben betrekking op twee andere toestanden van dezelfde hoeveelheid gas.

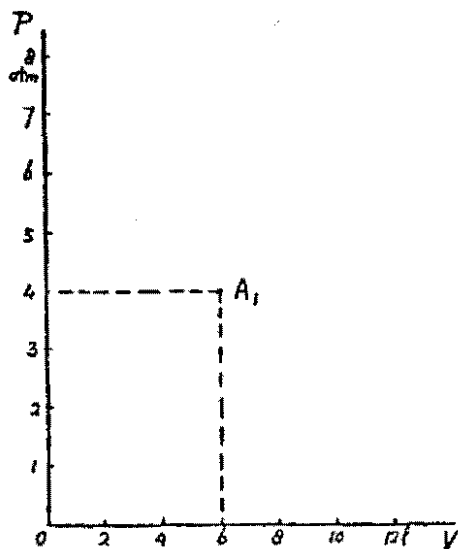


fig. 1

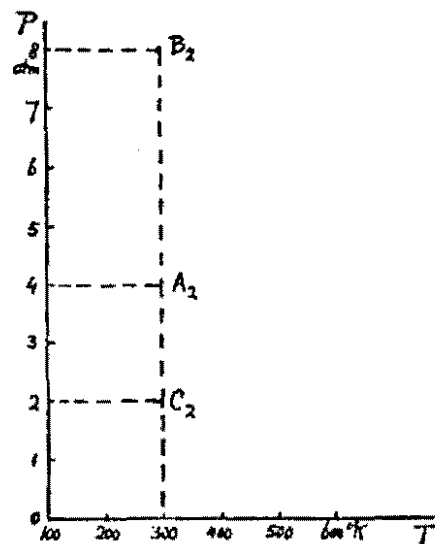
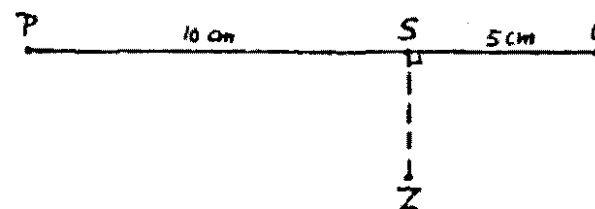


fig. 2

- Schrijf de waarden van de drie genoemde grootheden op voor toestand  $A_1$ .
- Hoe heet de in fig. 2 getekende verticale lijn  $B_2C_2$ ?
- Teken fig. 1 over (neem voor 1 liter 0,5 cm met voor 1 atm 1 cm), geef in deze figuur de bij  $B_2$  en  $C_2$  behorende punten aan en zet er de letters  $B_1$  en  $C_1$  bij. Licht toe hoe je de plaats van deze punten bepaald hebt.
- Teken in deze figuur de met  $B_2C_2$  overeenkomende lijn.

2. Een hefboom met een massa van 50 gram is draaibaar om een horizontale as  $S$ . Het vlak van tekening moet verticaal gedacht worden zó dat  $S$  loodrecht op dit vlak staat. Als de onhelaste hefboom in evenwicht is, bevindt het zwaartepunt  $Z$  zich onder  $S$ ; de lijn  $PSQ$  is dan horizontaal.  $P$  en  $Q$  zijn punten van de hefboom, respectievelijk 10 cm en 5 cm van de as verwijderd.



Als men aan  $P$  een voorwerp  $A$  met een massa van 1 gram ophangt, blijkt de hefboom een stand in te nemen, waarbij  $PQ$  een hoek van  $45^\circ$  met het horizontale vlak maakt.

De doorbuiging van de hefboom mag verwaarloosd worden.

- Maak een duidelijke tekening van de situatie en bereken de afstand van  $Z$  tot  $S$ .

Men plaatst nu een glas water onder  $A$  zodat bij horizontale stand van  $PQ$  het voorwerp  $A$  geheel is ondergedompeld. In die stand is de hefboom in evenwicht indien men aan  $Q$  eveneens een voorwerp met een massa van 1 gram ophangt.

- Welke waarde voor de soortelijke massa van  $A$  volgt hieruit?
- Hoe groot is de kracht die nu door de hefboom op de as wordt uitgeoefend?

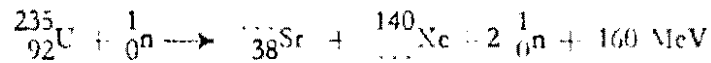
3. Iemand heeft de beschikking over een spanningsbron (24 V, 1 ohm) en over twee lampjes  $A$  en  $B$ , respectievelijk voor 12 V, 2 A en voor 6 V, 3 A. Bovendien zijn aanwezig een aantal verbindingsdraden met verwaarloosbare weerstand en een onbeperkt aantal weerstanden van verschillende waarden. Teken een aantal schakelingen (drie is voldoende), waarbij aan de volgende eisen moet worden voldaan:

- beide lampjes moeten normaal branden;
- in elke schakeling moeten twee weerstanden opgenomen worden.

Bereken voor elke schakeling de grootte van deze twee weerstanden.

4. Beantwoord naar keuze *een* van de beide volgende vragen 4A en 4B.

4A. De isotoop  ${}_{92}^{235}\text{U}$  wordt gespleten volgens de reactievergelijking

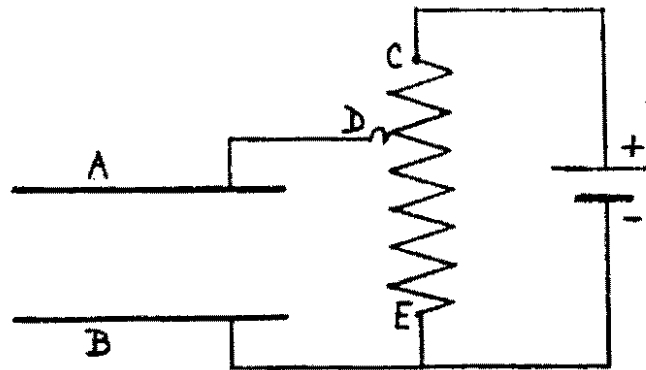


- Schrijf deze reactievergelijking volledig op (vul dus de ontbrekende getallen in) en geef aan wat de betekenis is van deze getallen.
- Leg uit hoe uit deze reactie een kettingreactie kan ontstaan.
- Waarvoor dient in een kernreactor de moderator?

d. Met een kernreactor als neutronenbron kan de fosfor-isotoop  ${}_{15}^{32}\text{P}$  gemaakt worden. Deze isotoop vervalt onder uitzending van betastraling tot zwavel (S) en wordt gebruikt als „tracer“.

- Schrijf de reactievergelijking voor dit radioactieve verval op.
- Bespreek het principe van de tracer-methode.

4B. A en B zijn de horizontale platen van een vlakke condensator. Via een zogenaamde potentiometerschakeling (zie de figuur) wordt de condensator door een spanningsbron geladen. Bij een bepaalde stand van het schuifcontact D brengt men tussen A en B een oliedruppeltje. Het druppeltje blijkt elektrisch geladen te zijn, want het beweegt in de richting van de bovenste plaat (A).



- Welke plaat heeft de hoogste potentiaal? Wat is dus het ladingsteken van het druppeltje? Verklaar je antwoorden.
- Moet men D naar C of naar E verschuiven om te bereiken dat het druppeltje tussen de platen blijft zweven? Verklaar je antwoord. Het druppeltje zweeft bij een potentiaalverschil tussen de platen van 490 V. De afstand tussen de platen is 1 cm.
- Bereken voor die situatie de elektrische veldsterkte tussen A en B.
- Bereken in eenheden van het m.k.s.A.-stelsel de verhouding van de lading en de massa van het druppeltje. De valversnelling is  $9,8 \text{ m/s}^2$ .

5. Beantwoord naar keuze *een* van de beide volgende vragen 5A en 5B.

5A. Met behulp van twee coherente, lijnvormige lichtbronnen ontwerpt men een interferentiefiguur op een scherm en neemt daarop lichte en donkere lijnen waar.

- Wat zijn coherente lichtbronnen?
- Bespreek een methode om coherente lichtbronnen te verkrijgen.
- Verklaar het ontstaan van de donkere lijnen.
- Bij gebruik van monochromatisch licht met een golflengte van  $4000 \text{ \AA}$  meet men de afstand tussen de middens van twee donkere lijnen en vindt daarvoor  $0,8 \text{ mm}$ . Hoe groot wordt deze afstand als men monochromatisch licht gebruikt met een golflengte van  $6000 \text{ \AA}$ ?
- Wat zal men op het scherm waarnemen als er wit licht gebruikt wordt? Geef hierbij een korte verklaring.

5B. Een grote vlakke spiegel staat verticaal, loodrecht op de hoofdas van een positieve lens, die op  $10 \text{ cm}$  links van de spiegel staat opgesteld. De brandpuntsafstand van de lens is  $5 \text{ cm}$ . Midden tussen de lens en de spiegel bevindt zich een lichtpunt, op  $2 \text{ cm}$  onder de hoofdas. Het lichtpunt is naar de kant van de lens afgeschermd, zodat het licht alleen maar na terugkaatsing door de spiegel de lens kan bereiken.

- Construeer het beeld dat uiteindelijk door de lens van het lichtpunt gevormd wordt.
- Bereken ook de plaats van het beeld, d.w.z. zowel de afstand van het beeldpunt tot de lens als de afstand van het beeldpunt tot de hoofdas.

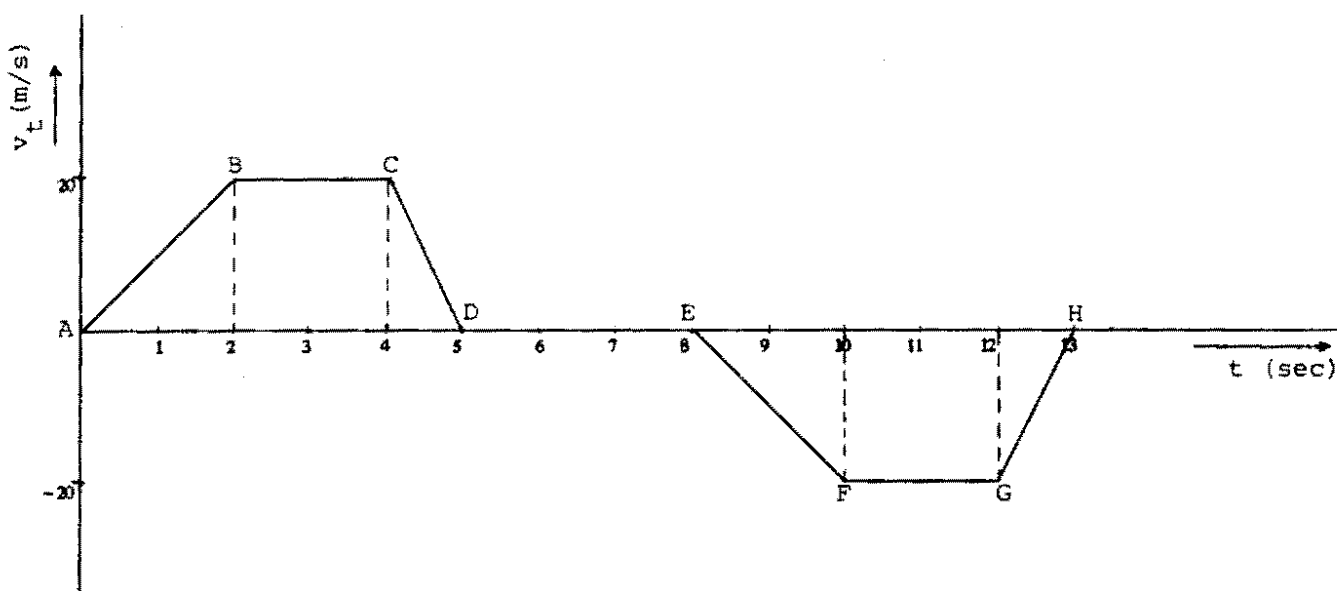
EINDEXAMEN HAVO 1969

In categorieën ingedeeld volgens Bloom en Klopfer

	B	K		B	K
1 a. Aflezen waarden van $p$ , $V$ en $T$ uit diagrammen	2.10	A11	b. Lezen van een schakelschema met spanningsdeler, toepassen van relatie tussen potentiaal en kracht	2.10	A11
b. Reproductie van term	1.11	A2	c. Toepassen van dezelfde relatie als bij b	2.20	A10
c. Toepassing van wet van Boyle in eenvoudige situatie en vertalen van de uitkomst in meetpunten	2.20	A11	d. Toepassen van $\Sigma F = 0 + qE = mg$ in eenvoudige situatie	2.20	F1
d. Reproductie van kennis $PV$ -diagram (kromme lijn)	1.30	A8	5A a. Reproductie van het begrip coherente lichtbronnen	1.11	A9
2 a. Vertaling gegevens in een tekening	2.10	A11	b. Reproductie van de beschrijving van de proef van Fresnel	1.20	A7
b. Toepassing van hefboomwet met gebruikmaking van begrip zwaartepunt	3.00	F1	c. Reproductie van een redenering	1.20	A5
c. Toepassing van hefboomwet en wet van Archimedes	3.00	F1	d. Toepassing van een formule in een eenvoudige situatie (als de formule gereproduceerd kan worden)	2.20	A10
d. Toepassing van $\Sigma F = 0$	3.00	F1	e. Reproductie van een verschijnsel	1.22	A5
3 Toepassing van wet van Ohm in vrij complexe niet vertrouwde situatie	3.00	F1	5B a. Constructie van stralengang	2.10	A11
4A a. Toepassing van wetten van behoud van lading en aantal kerndeeltjes (kennis wetten is voldoende)	2.20	A9	b. Toepassing van lensformule en beeldvorming door vlakke spiegel	2.20	F1
b. Reproductie kennis van een redenering	1.22	A5			
c. Reproductie van een begrip	1.11	A3			
d. 1 Als a	2.20	A9			
2 Reproductie van een methode	1.20	A7			
4B a. Kennis van de afspraak betreffende ladingstekens, kennis van de richting van krachten tussen ladingen. Eigen formulering hiervan	2.20	A5			

2.2.2. Proefwerk

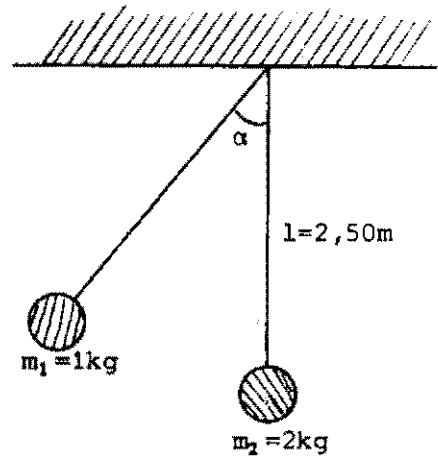
NORM	
<u>totaal:</u>	<u>dijfer:</u>
28 - 32	3
33 - 37	3½
38 - 42	4
43 - 47	4½
48 - 52	5
53 - 57	5½
58 - 62	6
63 - 67	6½
68 - 72	7



Figuur 1

1. De grafiek van figuur 1 geeft het verband aan tussen de snelheid en de tijd van een motor (massa 300 kg).  
De motor begint te rijden op  $t = 0$  sec.
- 4 a. Welke soort beweging voert de motor uit tussen AB; DE; FG ?
- b. Geeft een formule voor at en vt:
  - 2 1. tussen A en B,
  - 2 2. tussen B en C,
  - 2 3. tussen D en E.
- c. Hoe groot is de afgelegde weg in het tijdsinterval
  - 3 1. van  $t = 2$  sec tot  $t = 4$  sec,
  - 3 2. van  $t = 0$  sec tot  $t = 2$  sec.
- d. Hoe groot is de kracht die op de motor werkt tussen A en B; respectievelijk tussen B en C ?

2. Een slinger van 2,5 m lengte en een massa  $m_1 = 1$  kg krijgt een uitwijking van 10 cm in horizontale richting en botst tegen een zelfde slinger van gelijke lengte en een massa  $m_2 = 2$  kg. De massa's zijn van stopverf en de botsing is dus onelastisch.



Figuur 2

- 4 a. Bereken de slingertijd van de eerste slinger.  
4 b. Bereken de snelheid van slinger 1 ( $m_1$ ) vlak voor de botsing.  
4 c. Bereken de snelheid van het geheel na de botsing.  
4 d. Welke slingertijd heeft het systeem nu?  
4 e. Bereken de maximale hoogte na de botsing.

4  
20

3. Objektief en oculair van een astronomische kijker hebben brandpuntsafstanden van respectievelijk 200 cm en 10 cm. De as van de kijker is gericht op de onderkant van de maan. De schijnbare middellijn van de maan is 1 cm. De diameter van het objektief is 50 cm. De diameter van de ooglenzen is 5 mm.
- 2 a. Hoe groot is de afstand van het objektief tot het oculair bij waarneming door een normaal oog zonder akkommodatie?  
4 b. Onder welke hoek ziet men het beeld?  
4 c. Hoe verhouden zich de hoeveelheden licht die de astronomische kijker van de maan ontvangt tot de hoeveelheid licht die ons oog van de maan ontvangt?  
4 d. Schets een figuur van de opstelling (niet op schaal). Teken daarin drie stralen afkomstig van de bovenkant van de maan.

We plaatsen nu een fotografische plaat achter het oculair op een afstand van 240 cm van het objektief.

- 6 e. Hoever moeten we het oculair verplaatsen om een scherp beeld op de fotografische plaat te ontwerpen?

20

4. In een cilinder bevindt zich onder een zuiger 56 gram stikstof van 2,0 atmosfeer bij een temperatuur van  $27^{\circ}\text{C}$ . Aangenomen mag worden dat het stikstofgas zich in dit geval gedraagt als een ideaal gas.

3 a. Toon aan dat het volume van de stikstof ongeveer  $25\text{ dm}^3$  bedraagt.

We gaan de stikstof bij konstant volume verwarmen tot de temperatuur  $227^{\circ}\text{C}$  bedraagt (toestandsverandering A).

2 b. Bereken de nieuwe waarde van de druk.

4 c. Geef deze toestandsverandering weer in een p-V diagram én in een p-T diagram.

Hierna gaan we bij konstante druk de temperatuur verhogen tot  $427^{\circ}\text{C}$  (toestandsverandering B).

2 d. Bereken de nieuwe waarde van het volume.

4 e. Geef ook deze toestandsverandering weer in een p-V diagram én in een p-T diagram.

5 f. Beredeneer waarvoor meer warmte toegevoerd moet worden: voor toestandsverandering A of voor toestandsverandering B.  
20

5. Een dunne draad is horizontaal gespannen en wordt aan zijn linkeruiteinde transversaal en harmonisch aan het trillen gebracht. De frekwentie bedraagt 100 Hz. De draad is 1,5 m lang en het rechteruiteinde is vast. In de draad ontstaan knopen en buiken. De afstand tussen de eerste en de zesde knoop bedraagt 1,0 m.

4 a. Bereken de golflengte en de voortplantingssnelheid in de draad.

6 b. In welke mate zouden de golflengte en de voortplantingssnelheid veranderd zijn als:

1. de spankracht in de draad verdubbeld was? Geef een toelichting.

2. de frekwentie verdubbeld was? Geef een toelichting.

3. de massa van de draad verdubbeld was? Geef een toelichting.

5 c. De amplitude van de golfbeweging blijkt 4,0 cm te zijn in de buiken. Bereken de amplitude in een punt dat precies midden tussen een knoop en een buik ligt.

5 d. Hoe zal een foto van de trillende draad eruit zien als de belichtingstijd van deze foto  $1/50$  seconde bedraagt?  
20

2.2.3. Het proefwerk en de categorieën van Klopfer volgens 'zo maar vier kollega's' (ter beschikking gesteld door J.Ph. Steller, achteraf aan alle konferentiegangers uitgereikt)

Proefwerknr. I    Categorie van  
van  
Opgavenr.        Klopfer

I	a	6, 7, D2	7, (B1, D3)	3, 11	3,
	b	11 D3	11	3, 8	3, 11
	c	3, 8	3 (D3)	3, 4, 8	3, 8
	d	8	3 (F1)		3, 8
	e				
	f				
II	a	8	3	1, 3, 4	3, 8
	b	8	8	3, 4, 8	3, 8, 9
	c	8, 3	8 (9)	3, 4, 8	3, 8, 9
	d	8	3	1 of 7?	3, 8
	e	8	8	3, 4, 8	3, 8, 9
	f				
III	a	3	3, 2, 8	3, 8	3, 8
	b	8	7	3, 4	3, 8, 9
	c	3, 8	3	8	3
	d	11	(7)	7	4
	e	8	F1	3, 8	4, 10
	f				
IV	a	8	8	1, 3, 4	3, 8
	b	8	3	3, 4, 8	3, 8
	c	11	11	7	D2
	d	7, 8	3	3, 4, 8	3, 8
	e	11	11	7	D2
	f	8, 9	8 (D5)	6, 8	9
V	a	3, 8	3, 8	3, 4, 1	3, D3
	b	3, 6, 8	8	1, 3, 4, 6	3, 8, 9
	c	8	8	3, 4, 7	3, 10
	d	10, 11	(D5)	3, 9	3, 10
	e				
	f				

### 2.3. Verlagen

#### 2.3.1. *Groep Jetses*

Het toegestuurd schoolonderzoek is "geklopfed":

1. a A6	2. a A1	3. a A8	5. a A7( $\lambda$ ), A8(v)
b A11	b A8	b A7	b A8
c A7	c A8	c A9	c A8
d A8	d A10	d A11	d F1
	e A8	e F1	

Daarbij werden de volgende opmerkingen gemaakt:

- a) In de A categorie zijn kennis en begrip beide opgevoerd.  
Dat gaf nogal wat moeilijkheden. Bij Bloom zijn die wel onderscheiden, zodat die taxonomie voor onze schoolsituatie waarschijnlijk gemakkelijker toe te passen is.
- b) Bij vele vragen wisten we niet of het eerder behandelde stof betreft of een heel nieuwe situatie. Dat zou de keuze voor een bepaalde categorie beïnvloeden.
- c) A1 en A8 vonden we moeilijk te onderscheiden. Behoort  $v = \lambda f$  tot feitenkennis of tot kennis van een wetenschappelijke wet? Evenzo voor  $T = 2\pi \sqrt{\frac{l}{g}}$ .
- d) Het werd als een gemis ervaren, dat we de taxonomie zonder voorbeelden hadden.

Vraag: zou Auer een lijstje willen bijvoegen met de indeling, zoals hij die gemaakt zou hebben voor deze vraagstukken.

#### 2.3.2. *Groep van der Rijst*

- Eerst werd 20 minuten besteed aan het indelen aan de hand van Klopfer van de vragen van proefwerk I

Vraag 1. a. A11 b. A11 c. A7 of A8 d. A7 of A8

Vraag 2. a. A7 of A8 b. A7 c. A7 d. A10 of F1 e. A7

Opmerkingen: Kategorie A van Klopfer bevat erg veel.

Vraag 2b is slecht: veel rekenwerk.

Klopfer sluit slecht aan bij dit niveau, liever een instrument met 7 à 8 criteria.



- Gekeken is welke categorieën van Klopfer voorkomen in de les  
(dus niet bij toetsen)

HAVO: B1 t/m 3, C (?), D1 t/m 3, (F1), G1, 2

VWO : B1 t/m 4, C (?), D1 t/m 4, F1, G1, 2

Opmerkingen: Wat wordt bedoeld met C1?

Bij categorie E discussie of modelbegrip in die vorm bij de leerling aan de orde komt.

Conclusie: Veel categorieën komen wel aan de orde tijdens de lessen, maar worden niet getoetst.

Mening: Een (aangepast) instrument om toetsen te "toetsen" is wel gewenst en bruikbaar.

### 2.3.3. Voorgroep de Mink

(deze groep heeft zich voor de conferentie reeds bezig gehouden met deze opdracht)

vraagstuk	kategorie	vraagstuk	kategorie
1a	A6 (A11)	4a	A8
b	A11(A3)	b	A8
c	A7	c	A11
d	A8	d	A8
		e	A11
2a	A8	f	A8 (A9)
b	A8		
c	A8 (A6)	5a	A8
d	A8	b	A8
e	A8	c	A8
		d	A8(A11)
3a	A9		
b	A11		
c	A8(A11)		
d	A11		
e	A8		

Opmerkingen:

- 2a was een rekenvraagje
- 3 was een zeer antiek vraagstuk, te veel standaard
- 5d was moeilijk te beoordelen met de categorieën van Klopfer. Wel goede vragen; doet een goed beroep op het voorstellingsvermogen.

3. De categorieën van Klopfer concreet in de lespraktijk

3.1. Vraagstelling

Wat kan ik concreet met de categorieën van Klopfer doen binnen mijn lespraktijk?

3.2. Materiaal: zie 2.2.1, blz. 59 .

3.3. Verslag

*Groep Veth*

Allereerst werd kritiek geleverd op het proefwerk.

In vraagstuk 2 zitten de bollen van stofverf aan het begin van de proef al vrijwel tegen elkaar.

In vraagstuk 3 is "schijnbare middellijn" niet duidelijk gedefinieerd.

Vervolgens werden vraagstuk 1 en 2 van hetzelfde proefwerk volgens de "taxonomie" van Klopfer ingedeeld.

Naar aanleiding hiervan werd gediscussieerd over:

"Wat kan ik concreet met de categorieën van Klopfer doen binnen mijn lespraktijk?"

De klassificatie is afhankelijk van het niveau van de leerlingen en van de lessen die de leerlingen gehad hebben - wat volgens ons bijvoorbeeld A8 t/m 11 is, is voor de leerlingen dikwijls C1, etc. Het is moeilijk, zo niet onmogelijk, objectief te "klopferen". Via een zijspoor kwamen we op de vraag: Hoe leren we de leerlingen nieuwe (nieuw in de zin van: met zeer kleine wijzigingen) problemen oplossen?

Algemeen gesteld: Hoe kan men creativiteit bij de leerlingen behouden, respectievelijk ontwikkelen?

Misschien helpen "open-boek" toetsen hierbij?

Slotopmerking: "klopferen" kan leiden tot beter doordachte repetities.

4. Het maken van vragen met behulp van de categorieën van Klopfer

4.1. Vraagstelling

4.1.1. Maak zelf een opgave die valt in één van de categorieën B, C, D, E of F.

4.1.2. Wat kan ik concreet met de categorieën van Klopfer doen binnen mijn lespraktijk?

4.2. Materiaal: zie 2.2.1, blz. 59.

4.3. Verslag

*Groep Hey*

De groep besluit uiteen te gaan in vier subgroepjes van drie mensen, die elk zullen proberen een opgave te maken binnen twee aaneengrenzende Klopferse categorieën. Om een samenvattende discussie mogelijk te maken zullen de opgaven betrekking hebben op de wet van Ohm, Joule en dergelijke. Uitdrukkelijk wordt gesteld dat het proces belangrijker is dan de geproduceerde fysica.

Bij opgaven die in de B/C-kategorie kunnen worden ingedeeld is de voorgeschiedenis van de leerlingen erg belangrijk, wil de inschaling juist zijn. Wat is en wanneer en hoe uitgebreid behandeld, is van meer belang dan de vraagstelling zelf. Het tweede vertroebelende aspect is dat (jonge) leerlingen anders waarnemen dan wij vaak veronderstellen dat ze doen. Aan een waarneming wordt onmiddellijk een interpretatie verbonden.

De subgroep die zich met de C/D-kategorie heeft bezig gehouden vindt de volgende opgave een goed voorbeeld van C4:

Je hebt één enkele lamp van bekende lichtsterkte (die op een vaste spanningsbron is aangesloten).

Beschrijf nu eens hoe je hiermee het verband tussen verlichtingssterkte en de weerstand van een LDR kunt bepalen.

Aan welke eisen moet de lamp voldoen?

In de D-kategorie valt een opgave van de volgende soort:

Sluit een 12 V lampje in serie met een bekerglas met water aan op een wisselspanning die kan variëren van 0 tot 24 Volt.

Meet de stroomsterkte door het lampje bij toenemende spanning.

Maak van de resultaten een grafiek.

Verklaar de waarnemingen.

Wat zul je voor resultaten vinden als je de proef herhaalt met gedestilleerd water?

Een opgave die in D6/D4 past is de volgende:

Bepaal met een wisselspanningsbron het verband tussen de weerstand en de concentratie van een zoutoplossing.  
Leid uit de resultaten een relatie af tussen deze twee.  
Geldt deze relatie voor elke concentratie?

In de D/E-kategorie vallen in het algemeen de opgaven, waarbij uit grafieken of stellen grafieken ofwel nieuwe grafieken moeten worden gemaakt, ofwel algemenere conclusies getrokken.

Voorbeeld:

Doel van het vraagstuk is leerlingen die de warmteontwikkeling van een draad nog niet in formule kennen, deze te laten ontdekken. Gegeven worden drie T-t-diagrammen van een dompelaar in een bepaalde hoeveelheid vloeistof. Er worden dan eerst een aantal vragen gesteld uit de Klopfer A-kategorie om aanwezige kennis te reactiveren. Dan pas worden vragen gesteld om tot het bovenomschreven doel te komen.

In de E/F-schaal valt te denken aan problemen die analogieën behandelen tussen de waterstroommodellen en de wet van Ohm, serie- en parallelschakeling. Een tweede mogelijkheid: vanuit het molekuulmodel voorspellingen laten doen over de temperatuurafhankelijkheid van de weerstand van een metaaldraad, een halfgeleider en een zoutoplossing. Vervolgens constateren dat de resultaten met elkaar in tegenspraak zijn zolang het model niet verfijnd is.

De randvoorwaarde bij het construeren van dit soort problemen is dat het te gebruiken model tevoren goed omschreven moet zijn (of nauwkeurig bekend) inclusief alle daarmee samenhangende als juist aanvaarde conclusies.

Samenvattende conclusie van alle opgavenmakers was wel, dat het zonder voorbeelden erg moeilijk is opgaven te construeren binnen de Klopfer-schalen, zeker in de E en F categorie. In de lespraktijk is de manier van werken immers vaak tegengesteld.

## 5. Cijfer geven

### 5.1. Vraagstelling

#### 5.1.1. Voorgeschiedenis

De meerkeuzevragen zijn opgesteld door een aantal leraren, die daarmee bij hun kollega's op school het denken over cijfergeving op gang wilden brengen. Als alternatieven van de meerkeuzevragen hebben zij uitspraken genomen, die door leraren op die school op rapportvergaderingen gemaakt zijn. Dit ter rechtvaardiging van een door hen gegeven cijfer of als argumentatie in een meer algemene discussie over beoordeling van leerlingen.

#### 5.1.2. Gebruik van de meerkeuzevragen

Ofschoon de antwoorden naar zaken verwijzen, die op de betreffende school gespeeld hebben, lijkt ons de problematiek algemeen genoeg om ze ter discussie te stellen.

Technisch mankeert er nogal het een en ander aan de konstruktie van de meerkeuzevragen

- de antwoorden die bij de vragen horen zijn veelal onvergelijkbaar
- het ene antwoord sluit het andere vaak in
- het ene antwoord is zeer algemeen geformuleerd, het andere zeer specifiek

Als gevolg van genoemde punten is het goed mogelijk dat men het met alle antwoorden eens, c.q. oneens is.

De diverse antwoorden zijn dan ook meer bedoeld als een 'praten over' dan als een 'kiezen tussen'. In plaats van individueel naar het meest juiste alternatief te gaan zoeken, kunnen de meerkeuzevragen daarom beter gebruikt worden als gespreksingang voor een discussiegroep. Vooral de vragen 1, 6 en 7 achten wij daarbij belangrijk.

Bij een eventuele rapportage lijkt het ons zinvoller met de groep te komen tot een aantal stellingen en/of vragen, dan een opsomming van de meest juiste alternatieven te geven.

#### 5.1.3. Enige achtergrondinformatie bij de antwoorden op de vragen

Voor iemand met weinig of geen schoolervaring zal de aangesneden problematiek door de antwoorden op de vragen alléén, te weinig uit de verf komen. Daarom zullen we voor een verdere beeldvorming een aantal van deze probleemsituaties nader uitwerken en wel, zoals ze op de betreffende school gespeeld hebben. Aan het woord laten we een van de opstellers van de vragen.

1. *Overgangsnormen moeten al dan niet gebaseerd zijn op alle rapportcijfers (1b: 1e, 2c)*

We beginnen het schooljaar altijd met een algemene docentenvergadering. De eindexamenzorgen zijn dan wat weggeëbd. Iedereen is met goede voornemens bezield en heeft zijn idealen weer opgepoetst.

Toen op zo'n bijeenkomst de tekenleraar het belang van de algemeen vormende vakken aansneed, waren we ook direkt bereid om hem te steunen: "Als we zelf die vakken niet serieus nemen, hoe kunnen we dan van onze leerlingen verlangen dat zij het wel doen".

Om het belang van deze vakken te onderstrepen en om het deze keer nu eens niet bij praten te laten, werd besloten de betreffende vakken in de overgangsnormen op te nemen. Hun gewichtsfaktor zou wel is waar klein zijn, maar ze telden mee. Iedereen was tevreden. Aan een kommissie werd opgedragen het een en ander nader uit te werken.

Op het eind van het zelfde schooljaar werden we gekonfronteerd met de konsekwenties van de nieuwe regeling. Een van de leerlingen had zoveel onvoldoendes op zijn rapport, dat hij juist buiten de bespreekzone viel. Zonder zijn onvoldoende voor muziek was hij wel een bespreekgeval geweest. Het bleek zelfs dat in dat geval een meerderheid van de leraren voor bevordering van de betreffende leerling zou zijn geweest. Geen van de docenten, die deze leerling een onvoldoende hadden gegeven, bleek bereid zijn cijfer te verhogen. De leerling bleef zitten.

Achteraf hadden de meesten onder ons een sterk gevoel van onvrede met deze afloop. We hadden er toch wel moeite mee dat als deze leerling 'beter op de maat had kunnen lopen' - een van de argumenten waarom hij een onvoldoende had - hij niet had hoeven te doubleren.

2. *Moeten er wel overgangsnormen zijn? (1c, 1d)*

Het hiervoor beschreven geval had zeker niet zo'n dramatische afloop gehad bij een veel ruimere bespreekzone of totale afwezigheid van overgangsnormen. Vroeger, toen de school nog een stuk kleiner was, waren er ook nauwelijks overgangsnormen op schrift vastgelegd. Wel was er een door traditie gegroeide norm, die door ons allen bewaakt werd.

Een van de konrektoren was bij alle rapportvergaderingen aanwezig.

Wanneer wij het moeilijk hadden met een beslissing over een leerling bladerde hij een paar klassen terug en wees ons op genomen beslissingen bij vergelijkbare leerlingen.

De enige faktor die nog iets aan deze norm kon veranderen was het aantal geslaagden en gezakten bij de eindexamens van het afgelopen schooljaar. Bij goede resultaten werd de norm wat soepeler gehanteerd; bij slechte wat strenger.

Toen de school groter werd, werd de behoefte aan vastgelegde normen steeds groter. Door de grote aantallen leerlingen was het praktisch ondoenlijk ze nog met elkaar te vergelijken. Bovendien hadden de verschillende klassen een totaal andere konstellatie van leraren. De gemeenschappelijke norm vervaagde en werd steeds minder consistent toegepast. Zeker voor ouders en leerlingen was het onakseptabel dat een leerling met een beter rapport dan een medeleerling, die wel werd bevorderd, kon worden afgewezen.

Op dit moment zijn er uitvoerig vastgelegde normen met een zeer kleine bespreekzone. De vergaderingen worden door de staf goed voorbereid. Ondanks het feit dat het aantal klassen is toegenomen, duren de vergaderingen veel korter. Was ik vroeger erg bij deze vergaderingen betrokken, nu lijken ze langs me heen te gaan. Soms vraag ik me wel eens af of deze vergaderingen nog wel enige zin hebben.

### *3. Verschillende leraren cijferen verschillend (1a, 1d, 2a, 2b, 2c, 2d, 4a, 4b, 5c, 7a, 7b, 7c, 7d, 7e)*

We hebben op school een franse lerares, die op een klas van 30 leerlingen meestal geen of een onvoldoende voor haar vak heeft. Haar duitse kollega geeft echter in diezelfde klassen vaak 2/3 deel van de leerlingen onvoldoende.

Vroeger, toen er nog nauwelijks overgangsnormen waren, hield je daar rekening mee. Een onvoldoende voor frans telde je heel zwaar, terwijl je bij duits een vijf al voldoende vond. Niemand sprak dat uit, maar iedereen rekende ermee, met inbegrip van de betreffende leraar en lerares.

Nu is dat een groter probleem. Je kunt voor de verschillende docenten toch moeilijk verschillende normen maken (1a)? Of zou je van de docenten moeten verlangen dat zij een bepaald gemiddelde (4a) en een voorgeschreven aantal onvoldoendes in een klas hebben (4b)?

Overigens praat ik wel over andere vakken, maar in onze eigen sektie speelt hetzelfde probleem. De konrektor beweert dat hij aan de cijferlijst voor ons vak in een klas kan zien wie er van onze sektie in die klas lesgeeft, met andere woorden dat die lijst meer karakteristiek is voor de leraar dan voor de leerlingen.

Misschien is het wat overdreven dat hij op die manier alle leden van onze sekte - zes in getal - van elkaar kan onderscheiden, maar ik moet wel toegeven dat de verschillen erg groot zijn.

Centrale proefwerken en vaderlijke gesprekken met de rektor kunnen daar weinig aan veranderen. Het gekke is dat hierdoor de onderlinge afstemming eerder is verslechterd dan verbeterd. De meningen hebben zich verhard: het onderlinge wantrouwen is groter geworden.

Sommigen vinden dat er nog strengere maatregelen getroffen moeten worden, zoals: alleen nog maar centrale proefwerken en elkaars werk nakijken.

Ik ben bang dat dit het laatste stukje kommunikatie, dat er nu nog bestaat, nog verder zou afbreken.

Vanuit deze problematiek zijn de alternatieven van vraag 7 geboren. We hebben ze nog niet allemaal geprobeerd, dus er is nog hoop.

#### *4. De inhoud van een cijfer (3a, 3b, 3c, 3d, 5a, 5d, 6)*

Waar ik zelf erg mee zit is: 'wat zou de inhoud van mijn cijfer in het ideale geval moeten weergeven?'

De bij de keuzevraag vermelde alternatieven zijn voor mij even zovele vragen.

Daarbuiten zijn nog een groot aantal vragen die niet in de keuzevragen zijn opgenomen, zoals

- hoe verhouden zich kennis en begrip in een cijfer
- beloon of bestraf ik ijver ("hij krijgt een zes, omdat hij zo zijn best heeft gedaan" of "ik heb hem een zes gegeven, omdat ik er van overtuigd ben dat hij beter kan")



5.2. Materiaal

9 vragen over cijfergeving

1. Welke van de volgende beweringen is het meest juist?
  - a. Overgangsnormen moeten gebaseerd zijn op alle rapportcijfers, waarbij rekening gehouden moet worden met de cijferende docenten.
  - b. Overgangsnormen moeten alleen gebaseerd zijn op de belangrijke vakken (pakket)
  - c. Er moeten geen overgangsnormen zijn
  - d. Zolang er geen normen zijn voor het geven van rapportcijfers missen overgangsnormen (alleen gebaseerd op cijfers) elke grond
  - e. Overgangsnormen moeten gebaseerd zijn op alle rapportcijfers.
2. Voor welke vakken denk je, dat over het algemeen te hoge of te lage rapportcijfers worden gegeven
  - a. Voor de exacte vakken, omdat deze zo moeilijk zijn
  - b. Voor de leervakken, zoals aardrijkskunde en geschiedenis
  - c. Voor de expressievakken, omdat die niet te becijferen zijn
  - d. Ligt geheel aan de docent
3. Als een leerling een onvoldoende op zijn eindrapport heeft, behoort daarmee aangegeven te worden:
  - a. dat de leerling het gehele jaar zwak is geweest
  - b. dat de betreffende leraar vindt, dat deze leerling moet blijven zitten.
  - c. de leerling een van de slechtsten van de klas is.
  - d. dat de leerling het betreffende vak het volgend jaar niet meer goed zal kunnen volgen
  - e. dat de leerling op dat moment onvoldoende staat
4. Bij het maken van rapportcijfers is het van belang ervoor te zorgen:
  - a. dat het gemiddelde in de klas ongeveer  $6\frac{1}{2}$  is
  - b. dat er per klas voldoende onvoldoendes zijn
  - c. dat het een en ander statistisch zuiver gebeurt
  - d. dat het verschil in nivo tussen HAVO en Athneum duidelijk blijkt
5. Een rapportcijfer behoort tot stand te komen
  - a. door de repetitiecijfers, eventueel met gewichtsfactoren, te middelen.
  - b. na een gesprek met de betreffende leerling
  - c. in de rapportvergadering, nadat de leraar een cijfer heeft voorgesteld
  - d. door er een heleboel factoren in te betrekken zoals: indruk, prognose sympathie, pedagogische overwegingen etc.

6. Geef een rangorde aan van de volgende factoren, die bij het cijfergeven een rol spelen: prestatie, indruk en prognose
7. Het volgende voorstel zou tot een verbetering in de cijfergeving kunnen leiden:
  - a. de leraar stelt het cijfer "voor", de vergadering "vast"
  - b. er moet meer overleg komen over rapportcijfers tussen gelijksoortige vakken
  - c. de overgangsnormen moeten betrekking hebben op groepscijfers (zoals bij gymnasium oude stijl) en niet op losse cijfers
  - d. teveel afwijkende leraren moeten zich aanpassen.
  - e. bezinning op het fenomeen "cijfer"
8. Bij het afronden van cijfers geven externe factoren de doorslag, zoals:
  - a. bij lastige leerlingen naar beneden afronden
  - b. als pa kennis, familie of kollega is naar boven afronden
  - c. de uitspraak "bij mij krijg je nooit meer dan een 4" konsekwent volhouden
  - d. de leerling moet overgaan, omdat hij goed is voor de sfeer in de klas
9. Bij het afronden van cijfers behoren de volgende factoren de doorslag te geven
  - a. altijd naar boven afronden
  - b. naar boven of beneden afronden om pedagogische redenen
  - c. het gemiddelde in tienden uitrekenen
  - d. de prognose de doorslag laten geven
  - e. bij een positieve houding tov het vak naar boven afronden
  - f. uitschieters in de repetitiecijfers moeten geelimineerd worden

### 5.3. Verlagen

#### 5.3.1. *Groep De Jong*

Belangrijke punten uit de discussie:

- 1) Overgangsnormen uitsluitende gebasseerd op cijfers, missen elke grond, omdat er geen absolute normen zijn voor het geven van cijfers.
- 2) De normen op grond waarvan een rapportcijfer wordt gegeven hangen af van de docent, van de klas, van de school.
- 3) De cijfercode (1 t/m 10) moet door alle docenten op dezelfde wijze worden gebruikt. Duidelijk moet zijn welke cijfers een voldoende of onvoldoende beoordeling betekenen.

Voorgesteld is om het cijfer 5 als onvoldoende en het cijfer 6 als voldoende te tellen.

Een twijfelgeval (tussen 5 en 6) wordt op het overgangsrapport met + aangegeven. De leerling wordt hierdoor automatisch bespreekgeval. Hierdoor wordt voorkomen dat twijfelgevallen vóór de rapportvergadering op voldoende of onvoldoende worden afgerond.

#### 5.3.2. *Groep De Mink*

Men richt zich in de discussie op vraag 2 en formuleert de vraag aldus: Exacte vakken krijgen lagere cijfers dan andere vakken.

Men geeft kort geformuleerd de volgende oorzaken aan:

Het vak wordt door de leerlingen als (te) moeilijk ervaren.

Dit wordt ook gedeeltelijk veroorzaakt door het imago van het vak. De leerlingen die van de basisschool komen, hebben te weinig natuurwetenschappelijke bagage. Dit wordt geweten aan het feit dat de opleiding aan de Pedagogische Akademie teveel gebreken vertoont in deze.

De hogere cijfers onder andere bij de talen zou toegeschreven kunnen worden aan het feit dat talendocenten meerdere malen kleinere leerstofeenheden "overhoren" waarbij kennis (dus weetjes) afgevraagd worden.

De onvoldoenden bij het eindexamen in de exacte vakken (tot 40%) schrijft men onder andere toe aan het aanwezig zijn van vooraf vastgestelde normen, terwijl dit bij de "andere vakken" niet het geval is. Een oplossing zou te vinden zijn in het diagnostisch gebruik van door een centraal lichaam gemaakte toetsvragen.

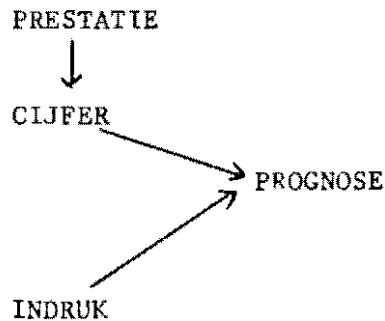
Er blijven nog wel een aantal vragen liggen.

Waarvan wij noemen:

1. Stellen "wij" onze eisen te hoog?
2. Is de opleiding van  $\beta$ -leraren wel zo goed?  
(van de andere vakken wel?)
3. Is natuurkunde wel aan te leren of is het een natuurlijke  
begaafdheid?
4. Hoe moeten wij cijfers geven, absoluut of relatief?
5. Wat is eigenlijk onvoldoende?

### 5.3.3. Groep Van der Hilst

Het drietal prestatie, indruk en prognose kan als volgt worden  
samengevat.



Er is (een) inflatie gaande ten aanzien van het aantal leerlingen,  
dat naar het AVO gaat. (Liefst) 40% van de 12-jarigen blijkt  
naar HAVO en VWO te gaan.

De becijfering van leerlingen (zowel bij schriftelijke werken als  
bij rapporten) moet met meer openheid geschieden (ook ten aanzien  
van de leerlingen). Ook is hierbij meer uniformiteit gewenst (zeker  
binnen een vaksectie).

### 5.3.4. Groep Van Ark (zaterdag)

I Van de verstrekte vragenlijst werden de punten 1,2,3 en 8 besproken:

ad 1b: Redactie van de vraag verandert in: "Overgangsnormen  
moeten alleen gebaseerd zijn op examenvakken".

Opmerkingen: Het lijkt onjuist iemand te laten zitten op een

slecht cijfer voor muziek, tekenen enz. omdat:

- a) dit geen examenvakken zijn en dus niet be-  
palend voor de vraag of de leerling het  
volgend leerjaar in het algemeen met vrucht  
zal kunnen volgen.
- b) docenten in deze vakken vaak slechte didactici  
zijn waarvan de leerlingen niet de dupe  
mogen worden.

- ad 2. ad 2a geschrapt: "omdat deze zo moeilijk zijn". Cijfers zijn hier onzes inziens vaak laag omdat deze vakken "anders" zijn: een hoger abstractie nivo eisen.  
ad 2e: cijfers zijn hier onzes inziens vaak te hoog.  
ad 2d: veranderd in "ligt (gedeeltelijk) aan de docent".
- ad 3. 3a: dit antwoord lijkt het meest juiste (gehele jaar zwak geweest).
- ad 8: Elke leraar zal proberen alle hier genoemde factoren zoveel mogelijk uit te sluiten en eventuele afrondingen vast te stellen op grond van zijn algemene indruk.

## II Waardering Praktikum

Er kunnen 2 soorten praktika worden onderscheiden:

A) VORMENDE PRAKTIKA: vrije opdrachten, binnen een langere periode uit te voeren.

Hierbij is veel inbreng van buiten mogelijk; daardoor moeilijk objectief te becijferen.

B) TEST-PRAKTIKA: (opdracht met direct aansluitende rapportage)

Hierbij is - bij uitsplitsing in op elkaar aansluitende doel-  
vragen - een redelijk objectieve becijfering mogelijk.

### 6. Praktikumbeoordeling

6.1. Vraagstelling: niet nader uitgewerkt

6.2. Materiaal: geen

6.3. Verslag

#### *Groep Verhagen*

Als inleiding werd gevraagd naar de motivatie van de keuze voor deze discussiegroep. Hierbij kwam de vraag naar voren: "Wat wil je met praktikum?" Het antwoord hierop bleek veelledig te zijn.

- a) de leerlingen een beter inzicht geven
- b) laten zien dat natuurkunde een experimentele wetenschap is
- c) leren samenwerken
- d) problemen leren aanpakken

De manier waarop deze doelen werden nagestreefd bleek wel van de wijze van praktikum geven af te hangen: Open praktikum, eerst theorie dan proef, eerst proef daarna theorie etc.

Daarna bespraken we: hoe evalueren we het praktikum.

Hiervan was het resultaat:

- a) verslag
- b) verslag met aansluitend gesprek
- c) gesprek achteraf
- d) vraagstukken achteraf

Hierbij rees wel de vraag of het überhaupt wel zinnig was het praktikum te beoordelen en zo ja hoe doen we dit dan? Hierop wisten we geen duidelijk antwoord te geven.

## 7. Beoordeling van leerlingen

### 7.1. Vraagstelling:

Door de leraren van de Rijksscholengemeenschap Breukelen is een lijst opgesteld met enkele aspecten van leerlingengedrag, waarvan zij de indruk hadden, dat die bij de beoordeling van leerlingen een rol zouden kunnen spelen. De lijst bevat alleen algemene aspecten, die niet speciaal aan een schoolvak gebonden zijn. Daarnaast zou iedere vaksektie nog een lijst opstellen, gericht op de beoordeling van leerlingen met betrekking tot dat betreffende schoolvak. De bijgevoegde lijst is in een brainstormbijeenkomst ontstaan. De bedoeling is dat hieruit een beter hanteerbare lijst ontstaat, die door het gehele docentencorps gebruikt zou kunnen worden. De beschrijvingen zouden dan het cijfersysteem moeten vervangen.

Een aantal vragen die men zich hierbij kan stellen zijn:

- Waarom wil ik een cijfer door beschrijvingen vervangen? Wie heeft daar wat aan? De beoordelende leraar zelf? Zijn kollega's? De ouders of de leerlingen? En wat hebben zij eraan?
- Afhankelijk van het hierboven gestelde: wat moet ik beoordelen van leerlingen?
- Aan wie mag/moet ik de beoordelingen meedelen?
- Op welke manier kan ik dat het beste doen?

BEOORDELING VAN LEERLINGEN  
=====Kenmerken van de leerling:Hoe komt dat tot uitdrukking in lessituaties:

plezier in het vak

- hij doet actief en enthousiast mee in de les
- hij stelt veel vragen
- hij komt met kranteknipsels
- hij komt buiten de les nasprenken of navragen
- .....(zelf invullen)

zorg

- zijn schriften/boeken zien er netjes uit
- .....

concentratie

- hij is in staat de hele les erbij te blijven
- hij is de hele les erg druk
- .....

doorzettingsvermogen

- laat zich niet ontmoedigen door een slecht cijfer
- .....

vlijt/ijver

- heeft altijd zijn huiswerk af
- .....

kreativiteit

- komt vaak met een oorspronkelijk idee
- .....

nauwgezetheid

- werkt zijn opdracht altijd heel precies uit
- .....

zekerheid

- is snel uit het veld geslagen/gespannen bij onduidelijkheid in de les
- wil alles opschrijven
- .....

prestatie

- hecht erg veel belang aan cijfers, ook als ze van weinig belang zijn
- .....

Kenmerken van de leerling:

Hoe komt dat tot uitdrukking in lessituaties:

zelfstandigheid

- zelf kunnen beginnen zonder leraar
- zelf materiaal kunnen pakken
- eigen werk kunnen beoordelen
- zelf kunnen bepalen of je er genoeg van weet/  
klaar bent
- .....

inzicht in eigen kunnen

- inzien dat fouten maken belangrijk is
- .....

impulsiviteit

- komt snel met zijn bijdrage
- .....

meer prakties dan teoreties

- doet liever praktikum dan sommen
- .....

agressiviteit

- wordt gauw kwaad
- .....

verlegenheid

- alleen met de leraar
- in de groep
- .....

kritische instelling

- kan opbouwende kritiek ontwikkelen op lesinhoud
- lesvorm
- houding docent
- kan relativeren

houding en vaardigheid binnen  
de groep

a. voor zichzelf

.....

1. wil en kan luisteren
2. wil en kan formuleren
3. durft en kan initiatief nemen
4. durft en kan leiding geven
5. kan eigen belang erkennen en naar  
voren brengen.



Kenmerken van de leerling:

Hoe komt dat tot uitdrukking in lessituaties:

b. voor de ander

- 1. wil en kan de ander res- .....  
pekteren in zijn werk
- 2. wil en kan de ander accep- -is niet jaloers  
teren
- 3. wil en kan zich inleven -eist alle aandacht voor zichzelf op  
in de ander
- 4. is behulpzaam -samen opruimen
- 5. durft en kan zich aan-  
passen
- 6. voelt zich verantwoorde- -aksepteert spelregels  
lijk voor groep -kan groepsontevredenheid manifesteren  
-kan anderen solidair corrigeren

Mogelijke wijzen van gebruik:

1. Maak een verslag van de leerling en maak daarbij gebruik van de links en rechts genoemde categorieën.

b.v. de leerling is erg geïnteresseerd in dit vak, doet altijd mee en vertelt er thuis over. Zijn tempo ligt wat laag, maar door zijn enthousiasme raakt hij niet achter.

2. Achter iedere categorie komt een schaal

b.v.	goed		matig		slecht	geen oordeel
concentratie	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tempo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
kreativiteit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

3. Beschrijf de ontwikkeling van de leerling aan de hand van categorieën

b.v. hij is de laatste tijd veel geconcentreerder en heeft veel meer kritiek op de lesinhoud

7.3. Verslag

*Groep Seller*

Bij een beoordeling volgens een schema met kenmerken zoals het ons werd voorgelegd zijn vooraf de volgende vragen te stellen:

1. Is het mogelijk en nodig al deze facetten zinvol te beoordelen?
2. Is zo'n beoordeling voor alle leerlingen te realiseren?
3. Mag en/of moet een aldus gevormde indruk betrokken worden in de beoordeling van en beslissing over de leerling?

Daarnaast is een belangrijke vraag: Voor wie is het aldus gevormde oordeel van belang? In welke situatie? Aan wie en in welke vorm wordt dat oordeel meegedeeld?

Algemeen werd wel aanvaard dat de beoordelingsaspecten voor de evaluatie van het leerproces en bij de diagnose van de moeilijkheden van de individuele leerling van belang zijn. In het laatste geval zowel voor de leraar als b.v. voor de mentor.

Deze beoordelingsaspecten zouden wel meetellen bij beslissingen omtrent verandering van schooltype, pakketkeuze en dergelijke, maar niet bij bevordering. (Hoe zou dat moeten?)

Het leek de groep ondoenlijk alle leerlingen op alle facetten te beoordelen. De beoordeling zou beperkt moeten zijn tot die leerlingen voor wie het nuttig is; op verzoek van de mentor b.v.

In de natuurkundeles zijn een aantal facetten goed te beoordelen. In een praktikumles komen houding en vaardigheid van de leerling zeer duidelijk naar voren. Zelfstandigheid, zorg en nauwgezetheid hangen daar nauw mee samen.

Zekerheid, plezier in het vak, meer praktisch dan theoretisch zijn in mindere mate te beoordelen, maar toch nog altijd beter dan impulsiviteit, verlegenheid, kritische instelling, prestatie (drift), vlijt/ijver en doorzettingsvermogen.

Sommige van de laatste aspecten vallen pas op zodra ze in het negatieve merkbaar zijn.

Een beoordeling waarin al deze facetten aan de orde komen en vooral het zinvol werken met deze beoordeling zal ontzettend veel tijd gaan kosten. Teveel?

## 8. Itemanalyse

### 8.1. Vraagstelling

Beoordeel de items uit 8.2.2. (zie blz. 92) aan de hand van onderstaand beoordelings-protocol. Zie voor de hieronder gebruikte begrippen blz. 90 en 91.

1. Geef zelf eerst eens een globale schatting van de grootte van de p- en de a-waarden (in procenten).
2. Bekijk hierna eens de resultaten die zijn verkregen bij de proefopname. Deze p- en a-waarden (+  $r_{it}$ -waarden) vindt U op de bladzijden 97 en 98.  
Bij de analyse van de itemresultaten kunnen de volgende punten naar voren komen:
3. a. De p-waarde is  $< 50$ , dat wil zeggen het item is voor deze populatie van leerlingen te moeilijk. Waarom is dit item nu te moeilijk?  
b. Een a-waarde is groter of gelijk aan de p-waarde, dat wil zeggen één der afleiders wordt veel te vaak door de leerlingen als het goede antwoord aangewezen. Wat zou hiervan de oorzaak kunnen zijn?  
c. Er zijn één of meer a-waarden  $< 5\%$ , dat wil zeggen deze afleider functioneert niet. Waar kan dat aan liggen?  
d. Dit item onderscheidt niet ( $r_{it} < 0,20$ ) of onvoldoende ( $0,20 \leq r_{it} < 0,30$ ) tussen goede en minder goede leerlingen. Waarom onderscheidt dit item niet?

Bij het zoeken naar antwoorden op de vragen 3a t/m 3d kunt U onderstaande punten in overweging nemen:

- ad 3a. Bij een te lage p-waarde ( $\sim 50$ ) moet men zich afvragen of
- de probleemstelling voor deze populatie redelijk is dan wel te hoge eisen stelt,
  - in het onderwijs aan de leerstof misschien onvoldoende aandacht wordt besteed,
  - het item wel redelijk leesbaar is met betrekking tot taalgebruik, lay-out. Was er misschien een tekening bij nodig geweest of is de gegeven tekening onduidelijk of fout?
  - het item misschien een constructiefout bevat waardoor een bepaald alternatief dat als onjuist is bedoeld (dus als afleider) toch voor de leerlingen het goede antwoord is.

- ad 3b. Als één a-waarde (het percentage leerlingen dat de afleider als het goede antwoord beschouwt) groter of gelijk is aan de p-waarde, zou het dan kunnen zijn
- dat het item een "val" bevat waar velen inlopen? of wijkt de afleider misschien nauwelijks van het juiste antwoord af?
  - dat het goede antwoord bij nadere inspectie misschien toch niet zo onaanvechtbaar is als de constructeurs hebben gemeend?
  - dat er in het onderwijs onvoldoende tegen deze misvatting of fout is gewaarschuwd?
- ad 3c. Bij niet te hoge p-waarden hoopt men dat de a-waarden niet gelijk aan nul zullen zijn. Een a-waarde  $< 5$  heeft in deze groep leerlingen niet overeenkomstig de bedoeling gewerkt namelijk als een gangbare fout of een voor de handliggende misvatting.

Het kan zijn dat deze afleider een "clue" bevat, dat is een aanwijzing, waardoor het voor bijna iedere leerling duidelijk is dat dit alternatief zeker niet het goede antwoord kan zijn.

- ad 3d. Het item onderscheidt niet of onvoldoende tussen goede en minder goede leerlingen.
- bevat het item misschien constructiefouten zodat ook de goede leerlingen op een dwaalspoor komen?
  - is het item te gemakkelijk? In dat geval kan het geen onderscheid aanbrengen want praktisch alle leerlingen geven het goede antwoord.
  - is het item veel te moeilijk, zodat de leerlingen maar eens een gokje wagen?

4. Graag een suggestie voor verbetering van dit item.

## 8.2. Materiaal

### 8.2.1. Begrippenlijst itemanalyse (zie blz. 90/91).

Hieronder volgen enkele begrippen die bij meerkeuze-toetsen gebruikt worden.

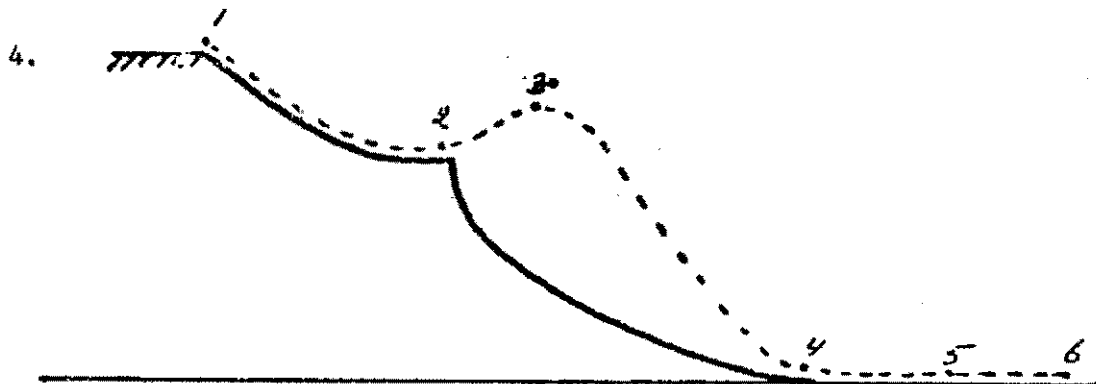
Item	Een meerkeuze-vraag met bijbehorende alternatieven.
Score	Het aantal punten dat een leerling op een toets behaald heeft d.w.z. het aantal goed beantwoorde items. (item goed beantwoord = 1 punt; item fout beantwoord = 0 punt).
Alternatief	Een bij een meerkeuze-vraag behorend goed of fout antwoord.
Afleider	Een bij een meerkeuze-vraag behorend <u>fout</u> antwoord.
p-waarde	Het percentage leerlingen dat het goede antwoord heeft gekozen. (Hoe hoger de p-waarde, hoe gemakkelijker het item. Gewenste waarde: $49 < p\text{-waarde} < 75$ ).
a-waarde	Afleider-waarde: het percentage leerlingen dat deze afleider als het goede antwoord beschouwde.
Voorbeeld	Gegeven een 4-keuze-item met de alternatieven A, B, C en D. Laat A het goede antwoord zijn; B, C en D zijn dan de afleiders. Heeft nu b.v. 70% van de leerlingen alternatief A gekozen en 12, 8 en 10% de respectievelijke afleiders B, C en D, dan heeft dit item: een p-waarde van 70 en drie a-waarden van respectievelijk 12, 8 en 10.
$R_{it}$ -waarde	De $r_{it}$ (item-totaal-correlatie) is de correlatie tussen itemscore (fout = 0; goed = 1) en de totaalscore op de toets. De $r_{it}$ van een item geeft aan in hoeverre een item goed is beantwoord door de goede (d.w.z. hoog scorende) leerlingen en fout is beantwoord door de minder goede (d.w.z. laag scorende) leerlingen. Hoe hoger de $r_{it}$ , hoe beter het item onderscheid maakt tussen goede en minder goede leerlingen.

Norm voor  $r_{it}$  0,40 en hoger zeer goed  
0,30-0,39 redelijk goed  
0,20-0,29 rand geval, item moet verbeterd worden  
0,19 en lager slecht

De psychometrische eisen waaraan een "meettechnisch" goed item moet voldoen:

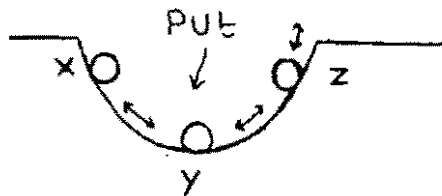
- 1) 49  $\langle$ p-waarde  $\langle$  75
  - 2) geen a-waarde  $\langle$  5
  - 3)  $r_{it}$ -waarde  $\geq$  0,30
- } getallen duiden procenten aan

8.2.2. De te beoordelen items



4. Bovenstaande figuur stelt een skischans voor. De skiër beweegt zich volgens de door de stippellijn aangegeven baan. Bij punt 4 begint hij te remmen, zodat hij in punt 6 stilstaat. De bewegings (kinetische) energie van de skiër is het grootst in
- A punt 1.
  - B punt 2.
  - C punt 3.
  - D punt 4.
5. Dit vraagstuk heeft eveneens betrekking op opgave 4. De zwaartekrachtenergie (potentiële) van de skiër neemt toe
- A van 1 naar 2
  - B van 2 naar 3
  - C van 3 naar 4
  - D van 5 naar 6
6. Wanneer een hoeveelheid water van  $0^{\circ}\text{C}$  overgaat in ijs zal
- A de temperatuur van het water stijgen
  - B de temperatuur van het water dalen
  - C er warmte vrijkomen
  - D er warmte nodig zijn

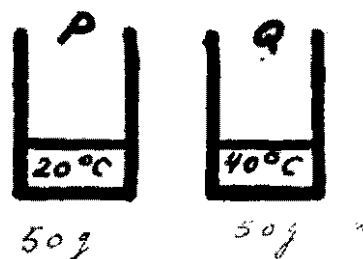
10. Een kogel rolt heen en weer in een put, zoals is aangegeven in nevenstaande tekening. De kogel verlaat de put niet. De wrijving van de kogel met de lucht en met de wand van de put wordt buiten beschouwing gelaten.



De totale energie van de kogel (de som van de zwaartekracht-energie en de bewegingsenergie) is

- A het grootst in punt x
  - B het grootst in punt y
  - C het grootst in punt z
  - D in alle punten even groot.
20. De grootte warmte heeft dezelfde eenheid als de grootte
- A temperatuur
  - B vermogen
  - C arbeid
  - D kracht

23. Men beschikt over 2 gelijke bekers P en Q, zie nevenstaande figuur. De bekers P en Q bevatten elk 50 g water.



De temperatuur van P en het water in P bedraagt 20 °C.

De temperatuur van Q en het water in Q bedraagt 50 °C

Beide bekers hebben dezelfde warmtecapaciteit.

We voeren nu de proeven I en II uit:

Proef I de inhoud van P gieten we bij Q, dit geeft direct na roeren eindtemperatuur  $t_I$ .

Proef II de inhoud van Q gieten we bij P, dit geeft direct na roeren eindtemperatuur  $t_{II}$ .

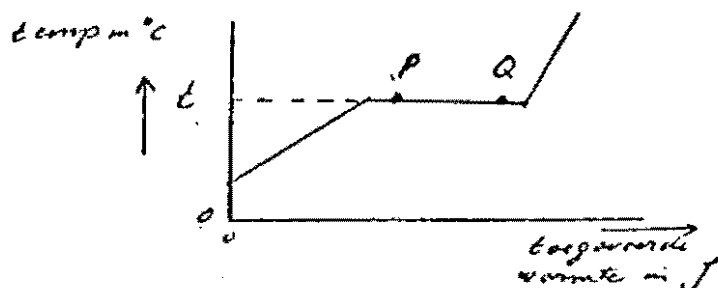
Welke van de onderstaande bewering over de eindtemperaturen  $t_I$  en  $t_{II}$  is juist?



- A  $t_I < t_{II}$
- B  $t_I = t_{II}$
- C  $t_I > t_{II}$
- D  $t_I$  wordt groter of kleiner dan  $t_{II}$ , er zijn te weinig gegevens om dit uit te maken.

28. Aan een bepaalde hoeveelheid vloeistof voert men regelmatig warmte toe.

De temperatuur als functie voor de toegevoerde warmte wordt gegeven door de grafiek in onderstaand diagram.



De temperatuur  $t$  die hoort bij de toestand aangegeven door punt P is gelijk aan de temperatuur die hoort bij punt Q, omdat

- A de vloeistof (bij die temperatuur  $t$ ) stolt
- B de vloeistof (bij die temperatuur  $t$ ) kookt
- C de toegevoerde warmte bij beide punten P en Q gelijk is
- D de toegevoerde warmte bij beide punten P en Q uitsluitend gebruikt wordt om de moleculen sneller te doen bewegen.

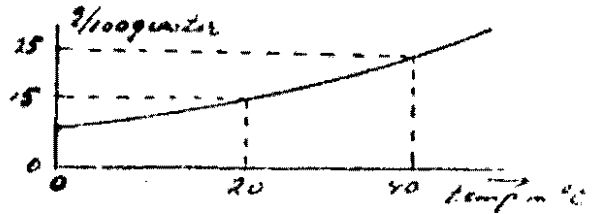
29. Dit vraagstuk heeft eveneens betrekking op de gegevens van opgave 28.

Gegeven zijn onderstaande beweringen I en II.

- I In de toestand behorende bij punt P is meer vloeistof aanwezig dan in de toestand behorende bij punt Q.
- II In de toestand behorende bij punt P kookt de vloeistof.

- A I en II zijn beide juist.
- B I is juist en II is onjuist
- C II is juist en I is onjuist
- D I en II zijn beide onjuist.

30. In nevenstaand diagram is de oplosbaarheid van stof P (in grammen /100 g water) uitgezet tegen de temperatuur.



In 50 g water van 0 °C kan men van stof P oplossen

- A 0 g
- B 4 g
- C 8 g
- D 15 g

31.



Door het draadstuk P Q R, zie bovenstaande figuur, loopt een elektrische stroom. P Q is dikker dan Q R. De lengten van P Q en Q R zijn gelijk.

Beschouw de volgende beweringen:

- I De stroomsterkte in P Q is gelijk aan de stroomsterkte in Q R.
- II De warmteontwikkeling in P Q is gelijk aan de warmteontwikkeling in Q R.

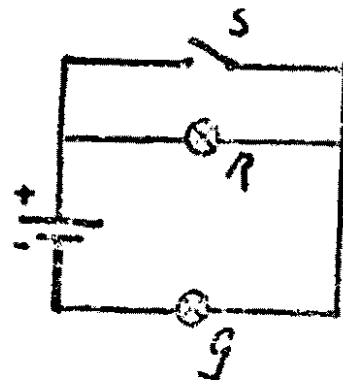
- A I en II zijn beide juist
- B alleen I is juist
- C alleen II is juist
- D I en II zijn beide onjuist

32a. In de getekende schakeling is S een geopende drukschakelaar, G een groen- en R een rood lampje (zie nevenstaande figuur).

Men houdt nu S neergedrukt.

Hierna zal

- A alleen lampje G branden
- B alleen lampje R branden
- C geen der lampjes branden
- D zowel lampje G als lampje R branden.

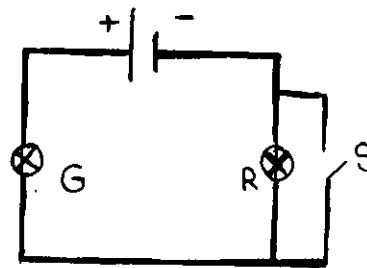


32b. In de getekende schakeling is S een geopende drukschakelaar, G een groen- en R een rood lampje (zie nevenstaande figuur).

Men houdt nu S neergedrukt.

Hierna zal

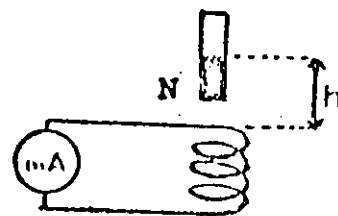
- A alleen lampje G branden
- B alleen lampje R branden
- C geen der lampjes branden
- D zowel lampje G als lampje R branden.



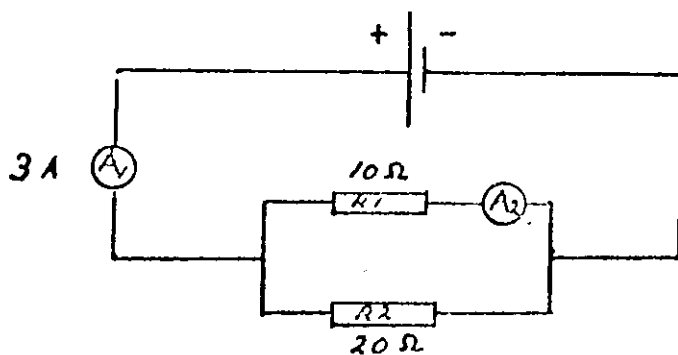
33. Door een spoel welke verbonden is met een milli-ampèremeter laat men vanaf een bepaalde hoogte h een magneet vallen. Men meet de maximale uitslag van de milli-ampèremeter.

Vergroting van de uitslag van de meter zal NIET optreden als men

- A een spoel met meer windingen gebruikt
- B de magneet laat vallen met de zuidpool naar beneden
- C een sterkere magneet laat vallen
- D de magneet van een grotere hoogte laat vallen



34.



In bovenstaande schakeling zijn twee weerstanden R1 en R2 parallel geschakeld.  $R_1 = 10 \Omega$ , en  $R_2 = 20 \Omega$

De ampèremeter A1 geeft een stroomsterkte aan van 3A.

We verwaarlozen de weerstanden van de ampèremeters.

Welke stroomsterkte zal de ampèremeter A2 aangeven?

- A 1,0 A
- B 1,5 A
- C 2,0 A
- D 3,0 A

Beoordelingsresultaten van de items op blz. 92 t/m 96, verkregen met een proefgroep van circa 500 leerlingen per schooltype.

De letters A, B, C en D geven de vier mogelijke keuzen per item aan. Het \* geeft de p-waarde aan. Alle getallen zijn opgegeven in %. De cijfers onder 0 geven het percentage van de leerlingen aan, dat het item niet heeft beantwoord.

Itemanalyse

no. 4.	A	B	C	D	O	Σ
VWO-3	2	12	11	57*	15	0,51
HAVO-3	3	37	20	34	4	0,44
no. 5.						
VWO-3	3	52*	10	1	24	0,60
HAVO-3	2	35*	37	2		0,49
no. 6.						
HAVO-3	3	53	37*	4		0,27
HAVO-3	3	30	60*	5		0,36
LTO-T-4	3	47	44*	5		0,41
no. 10.						
VWO-3	10	13	3	49*	25	0,55
HAVO-3	30	30	10	26*	4	0,38
no. 20						
VWO-3	38	22	36*	4	1	0,34
HAVO-3	60	20	16*	3	1	0,22
no. 23.						
VWO-3	5	72	18*	3	2	0,22
no. 28.						
VWO-3	4	60*	10	21	5	0,14

no.	A	B	C	D	O	$\Sigma$
no. 29 VWO-3	45*	17	18	14	5	0.19
no 30. VWO-3	28	37*	19	9	7	0.15
no 31. HAVO-3	7	48*	6	39		0.36
MAVO-3	5	43*	-4	47		0.32
LTO-T-4	2	57*	3	57		0.25
Versie A						
HAVO-3	15*	7	9	69		0.33
MAVO-3	24*	6	7	61		0.29
no. 32						
Versie B						
HAVO-3	33*	5	5	54	3	0.34
MAVO-3	32*	3	4	61		0.34
no. 33						
HAVO-3	11	43*	7	22	17	0.00
MAVO-4	12	45*	7	29	7	0.32
LTO-T-4	11	50*	6	30	3	0.31
no. 34						
HAVO-3	24	23	35*	14	4	0.45

### 8.3. Verslagen

#### 8.3.1. *Groep Schröder*

De schatting van de p-waarden leverde als resultaat dat men als docent veel te hoge verwachtingen had.

De oorzaak van het grote verschil tussen geschatte p-waarde en de gescoorde p-waarde moet bij het onderwijs worden gezocht.

Waarschijnlijk is de didaktiek nog niet goed ingespeeld.

Verbeteringen van de vragen zijn nauwelijks te geven.

#### 8.3.2. *Groep Vervoort*

In de discussie is volgens het beoordelingsprotocol een aantal van de voorgelegde meerkeuzevragen onder de loep genomen. De detailkritiek overheerste en het is moeilijk een grote lijn in het debat aan te wijzen. Twee opmerkingen lijken de moeite van het vermelden waard.

1. Item 5 is te moeilijk voor HAVO-leerlingen. We vermoeden dat het begrip potentiële energie voor HAVO-leerlingen te moeilijk is en zoeken de oorzaak niet zozeer in de vraagstelling. Twee afleiders zijn slecht. Wij vinden, dat in gevallen waarin een afleider zo goed funktioneert als hier, de andere niet zo goed hoeven te zijn.  
Het mag niet te vaak voorkomen, maar het is geen reden om zo'n item uit het hele werk te verwijderen.
2. Een meerkeuzevraag is op twee manieren te redigeren:
  - a de situatie beschrijven, de vraag stellen en 4 antwoorden eronder zetten (zie b.v. item 1)
  - b de situatie beschrijven en in 4 regels de vragen stellen (item 6)

Wij vermoeden dat de redactie als b) het voor de leerlingen moeilijker maakt te begrijpen wat er gevraagd wordt dan redactie a). Met andere woorden: wie zijn leerlingen liever natuurkundige dan taalkundige vragen stelt, kiese redactie a).

#### 8.3.3. *Groep van Genderen*

Nadat de groep zich aan elkaar had voorgesteld, werden een aantal afspraken gemaakt:

- a) Er werd een beperkt aantal items (5 stuks) uitgekozen voor de analyse.

- b) Per item was maar weinig tijd beschikbaar, zodat bij veel discussie over een bepaald onderdeel de rest maar overgeslagen werd.
- c) Suggesties ter verbetering van een item zullen alleen bij voldoende tijd gegeven worden.

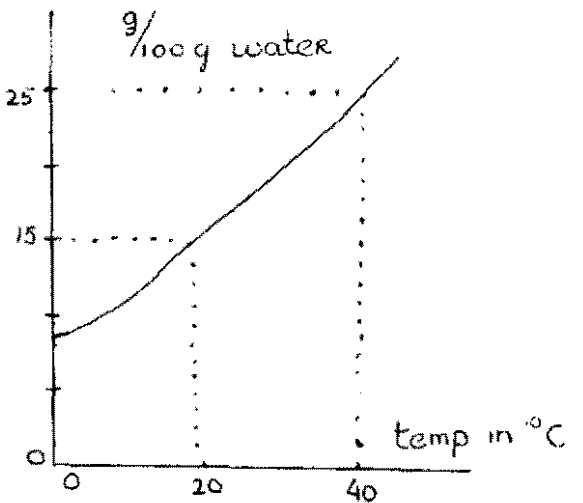
Werkwijze: De groep dacht dat de volgende werkwijze het meest effectief zou zijn:

- a) Eerst goed doorlezen van het item
- b) Schatten van het percentage antwoorden, welke de leerlingen gegeven zouden hebben.
- c) Doorgeven van deze gegevens aan de rapporteur
- d) Berekenen van het gemiddelde
- e) Vergelijken van dit gemiddelde en de resultaten van de proefopname.

Voorbeeld: Item nr. 30

In nevenstaand diagram is de oplosbaarheid van de stof P (in grammen/100 gr water) uitgezet tegen de temperatuur. In 50 g water van 0 °C kan men van stof P oplossen

- A 0 gram
- B 4 gram
- C 3 gram
- D 15 gram



	1	2	3	4	5	6	7	8	9	10	gem.	resultaten itemanalyse
A	3	10	3	5	5	5	15	0	5	5	5,3	28
B*	60	50	60	70	40	45	30	50	50	60	51,5	37
C	30	30	30	20	50	45	50	20	40	30	44,5	19
D	7	10	7	5	5	5	5	30	5	5	8,4	9
0												7
												$r_{it} = 0,15$

Verklaring: B\* = p-waarde

A, C en D = a-waarden afleiders

0 = percentage leerlingen dat de vraag niet heeft ingevuld

Zoals aan de getallen duidelijk te zien is, is het eigenlijk bijzonder moeilijk om te voorspellen hoe leerlingen de items maken.

De itemanalyse van de leerlingen geeft aan een p-waarde  $\approx 50$ , dat wil zeggen dit item is te moeilijk, terwijl de  $r_{it}$ -waarde  $< 0.20$ , dus dit item onderscheidt niet of onvoldoende dat wil zeggen ook slimmere leerlingen maken hier een fout.

Een mogelijke verbetering zou kunnen zijn:

- a) Formulering uitgebreider
- b) Het diagram uitzetten op grafiek papier

Van de andere items volgt hieronder alleen het gemiddelde van de groep, de resultaten van itemanalyse en de  $r_{it}$ -waarden met mogelijk de voorgestelde verbeteringen en verdere opmerkingen.

Item 6

	Gem.	I.A.		
		HAVO-3	MAVO-3	LTO-T4
a	5	3	3	3
b	23	53	30	47
c*	65*	37*	60*	44*
d	7	4	5	5
		$r_{it} = 0,27$	0,36	0,41

Opmerkingen:

1. De resultaten op MAVO lagen in de lijn der verwachtingen
2. Merkwaardige verschillen HAVO-MAVO
3. Kiezen van b waarschijnlijk door slordigheid in lezen
4. "Misschien is het onderwerp temperatuur en warmte toch te abstract".

Item 23

	Gem.	I.A.	
		V.W.O-3	
a	17	5	
b	47	72	
c	22*	18*	
d	14	3	

$r_{it} = 0,22$

Opmerkingen:

1. Te veel en verwarrende tekst
2. Voor deze opgave is meer vraagstukkenstraining vereist
3. Aanbieding (lay-out) zeer slecht. Misschien zou een foto deze item kunnen verbeteren.



Item 32 Versie A

	Gem.	I.A.	HAVO-3	MAVO-3
a	48		15*	24*
b	7		7	6
c	7		9	7
d	38		69	61
		$r_{it} = 0,33$		0,29

Opmerking:

1. Tijdens de lessen zouden de leerlingen zo'n probleem eens, na beantwoording, experimenteel moeten oplossen.

Item 32 Versie B

	Gem.	I.A.	HAVO-3	MAVO-3
a	40*		33*	32*
b	10		5	3
c	13		5	4
d	37		54	61
		$r_{it} = 0,33$		0,34

Item 34

	Gem.	I.A.	HAVO-3
a	27		24
b	15		23
c	51*		35*
d	7		14
o			4
		$r_{it} = 0,45$	

Opmerking:

1. De gegevens in de tekst moeten dezelfde nauwkeurigheid hebben als in de figuur.

Daar de beschikbare tijd verstreken was kon de groep zich niet beraden over andere items, verbeteringen, etc.

#### 8.3.4. *Groep Jetses*

Besproken werden de items 4,5 en 6.

Bediscussieerd werden de afwijkingen in de analyse.

Van item 4 kwam vooral de fysische realiteit ter discussie in verband met de voorstelling die iedereen via de T.V. ervan bezit.

Het verbaasde ons daardoor niet dat B een hoge score behaalde.

Item 5: De groep kon geen verklaring vinden voor de hoge score van C bij HAVO-3

Item 6: Er worden twee duidelijk verschillende feiten getest, namelijk het constant blijven van de temperatuur en het vrijkomen van warmte.

Omdat Broekman binnen kwam kreeg de discussie een ander karakter:

- De betekenis van de formule  $r_{it} = \frac{\bar{y}_i - \bar{y}}{s_y} \frac{\sqrt{p}}{q}$  werd geppord te doorgronden.
- De mogelijkheid om voor iedereen aanvaardbare doelstellingen op te stellen, lijkt bijna een utopie. De mogelijkheid om op grond hiervan dan items te construeren is dus ook voorlopig niet duidelijk aanwezig.

Voor p-geklopperde toetsen zijn niet zo absoluut als men zou hopen, omdat de categorieën vermoedelijk afhankelijk zullen blijken te zijn van het voorgaande onderwijsproces.

Misschien is het zinvol met een beperkt aantal categorieën een proef te nemen.

#### 8.3.5. *Groep de Jong*

De volgende procedure is gevolgd:

Eerst zelf voorspellen welk percentage gescoord heeft op de diverse alternatieven, daarna bekijken wat in werkelijkheid gescoord is, om tenslotte na te gaan waar de voorspelling (of het item) de mist in ging (want dat was soms het geval).

Vraag 4:

Voorspelling voor VWO-3

A	0	5	15	5	5	5	5
B	10	15	20	25	15	20	10
C	5	10	5	10	10	5	25
D	70	70	60	35	70	70	60

Vraag 5:

Voorspelling voor VWO-3

A	0	0	20	10	10	10	8
B	75	100	50	65	75	60	60
C	0	0	20	20	10	10	15
D	10	0	10	5	2	20	0

De grootste moeilijkheid bij deze vragen zit waarschijnlijk in het verschil tussen de werkelijke situatie en de hier bedoelde (ideale) situatie.

In deze zin zou de vraag dan nog verbeterd kunnen worden

Vraag 20:

A	5	20	20	20	10	20	2	30	20
B	15	15	25	5	5	25	10	5	5
C	70	60	45	40	80	50	86	60	60
D	10	5	10	35	5	5	2	5	10

Het probleem bij deze vraag zit waarschijnlijk eerder bij de uitleg dan bij de leerling. Waarschijnlijk kunnen leerlingen op dit niveau nog niet het juiste begrip warmte aan.

Conclusie: onderwijs over dit onderwerp moet aangepakt worden. Ten slotte is naar aanleiding van vraag 28 gediscussieerd over de vraag of een item met een lage  $r_{it}$ -waarde niet in de tekst zou thuishoren.

CITO zal deze er wel uit willen hebben omdat de vraag niet discrimineert en daardoor de test onbetrouwbarder maakt, maar een leraar zou hem wel willen gebruiken om na te gaan of een (b.v.) eenvoudig stukje kennis, bij iedereen aanwezig is.

Slotconclusie:

Uit voorspelling gecombineerd met resultaat blijkt: leraren zijn niet goed in staat te voorspellen hoe moeilijk een vraag voor leerlingen is.

## 9. Evalueren in een breed kader

### 9.1. Vraagstelling

Als U zou willen dat de leerlingen van een hoogste of een op-een-na hoogste klas Uw onderwijs *serieus* zouden evalueren, hoe zoudt U dit dan met hen willen aanpakken? (wat voor procedure, wat voor vragen, wat voor type discussie, enz).

Gegeven: er zijn 2 lesuren beschikbaar.

9.2. Materiaal: geen

9.3. Verslag:

*Groep Verhogen*

Er werd een onderscheid gemaakt tussen continue evaluatie, en evaluatie die eenmaal per jaar of trimester plaatsvindt. Het meest ideale is om zo'n sfeer in de klas te hebben dat de leerlingen onmiddellijk reageren, als ze het met de manier van lesgeven, onderwerpen en dergelijke niet eens zijn. Deze gegevens kunnen dan direkt gebruikt worden om veranderingen aan te brengen. "Eenmalige" evaluatie zou bijvoorbeeld als volgt kunnen plaatsvinden: leerlingen brengen zoveel mogelijk punten in die ze willen bespreken. Hiervan wordt een lijst op het bord gemaakt. Leerlingen krijgen allemaal gelegenheid om, direkt (zodat er geen overleg en beïnvloeding door andere leerlingen plaatsvindt) of thuis (zodat hij er rustig over na kan denken) hun visie op deze punten te geven.

De resultaten kunnen gebruikt worden om doelstellingen te bepalen of daarvan bewust te worden, en moeten uiteraard van invloed zijn op het onderwijs daarna. Daarna kan de leraar eventueel zelf punten inbrengen die niet aan de orde zijn gekomen. Resultaten van proefwerken moeten ook wel gebruikt worden om je onderwijs aan te passen, maar de informatie die je hieruit kunt krijgen is veel te summier.

10. Evalueren van complexe natuurwetenschappelijke doelen

10.1. Vraagstelling

Kies een doelstelling van de volgende soort (zie ook lijst C.M.L.N.-doelstellingen, blz. 55/56).

- a. de leerling moet bewust kunnen opteren voor de experimentele methode om een probleem aan te pakken,
- b. bij de leerling moet een zekere bereidheid en bekwaamheid ontstaan om ook buiten de klas zelfstandig kennis te verwerven en toe te passen,
- c. de leerling moet als basis-attitude hebben: eerlijkheid in waarneming,

en ga na hoe het bereiken van deze doelstelling eventueel vastgesteld zou kunnen worden.

10.2. Materiaal: zie 1.2. blz. 55/56

10.3. Verslag:

*Groep De Mink*

We hebben ons bezig gehouden met de volgende complexe natuurwetenschappelijke doelstelling, welke ons inziens het nastreven waard is:

Bij de leerling moet een zekere bereidheid en bekwaamheid ontstaan om ook buiten de klas en buiten het vak natuurkunde de natuurwetenschappelijke denkwijze aan te wenden.

Dit houdt in dat er in de klas ook aandacht geschonken moet worden aan (natuur)wetenschappelijke methodieken en strategieën welke overdraagbaar zijn (een transferfunctie hebben) naar andere vakgebieden en naar maatschappelijke situaties.

Het *toetsen* in hoeverre deze vaardigheden bij een individuele leerling bereikt zijn is een complex probleem waar wij vooralsnog geen duidelijke oplossing voor zien. Wel is het mogelijk een aantal *voorwaarden* voor het onderwijsleerproces te noemen die het bereiken van deze leereffekten bevorderen. We zijn in de discussie bij dit laatste terecht gekomen en hebben daarmee tevens zicht willen geven op een explicitering van deze doelstelling.

Problemen bij het nastreven van deze doelstelling

- de structuur (meso- en makro van het onderwijs, met inbegrip van het toetsingssysteem en het overladen programma, heeft tot gevolg dat het denken en handelen in de klas zich meer konsentrereren op kennis vergaren over een veelheid van onderwerpen.
- een andere faktor die het bereiken van deze doelstelling tegenwerkt vormen de minder creatieve leerlingen. Hoe houd je daar' feeling mee; moet je en kan je in de classesituatie hun creativiteit sturen?

Middelen en voorwaarden om de doelstelling te bereiken

- a. Meer natuurkunde uit het dagelijks leven. Integratie in de les van situaties uit de maatschappij (energie) en uit de leefwereld van het kind (wat ze thuis doen, in de krant lezen etc.)
- b. Probeer vakkenscheiding te minimaliseren door relaties met andere vakken te leggen.
- c. Breng niet alleen gesloten stukjes natuurkunde, maar houdt ook problemen open.

- d. Stel de leerling open vragen over concrete natuurwetenschappelijke verschijnselen.
- e. Oefen de leerling in het stellen van vragen en het formuleren van hypothesen.
- f. Wanneer de leerling eigen activiteiten moet verrichten zoals bij een praktikumopdracht moet het niet alleen gaan om het eindverslag, maar ook om de handelingen en processen die bij de leerling hebben plaatsgevonden.

Bij d, e en f liggen misschien ook toetsingsmomenten in individuele gesprekken en begeleiding

11. Problemen bij beoordelingslijst van leerlingen

11.1. Vraagstelling: zie 11.3. onderstaand

11.2. Materiaal: zie 7.2, blz. 84 e.v.

11.3. Verslag:

*Groep Veth*

Uitgegaan van de vraagstellingen:

- 1. wie stelt de "lijst" op en waarom?
- 2. naar wie gaan de gegevens?
- 3. waarvoor dienen de gegevens op de lijst?

In principe stellen de docenten de lijst op.

a. Deze zijn moeilijk tot invullen te krijgen. Verplichten leidt tot snel en ondoordacht invullen.

Reduceren van de lijst levert echter te weinig gegevens, te weinig waarde

b. Centrale verwerking van alle gegevens van alle leerlingen is onmogelijk

Ondervangingsproblemen:

ad a) In de lerarenopleiding (ev. bijscholing) aandacht besteden aan observatietechnieken

ad b) Beperking tot "probleem leerlingen" na afspraken gemaakt te hebben in een lerarenvergadering.

Centrale map zou gegevens kunnen bevatten, die ter inzage liggen voor docenten en schoolleiding (ouders? leerlingen?)

Voorzichtig met buitenstaanders.

Marktidee:

Presentatie van deze gegevens zou kunnen plaatsvinden op een 'ouderavond-markt', waar de ouders van alle leerlingen de gegevens over hun kind kunnen vernemen.

Waarvoor dienen de lijstjes?

- Om de minder opvallende leerling beter te leren kennen (dus ten bate van de leraar zelf).
- Om meer inhoud te geven aan het dooie cijfer (onder andere ten bate van de ouders).
- Betere determinatie mogelijk (brugklas - tweede klas en vakkenpakket).
- Betrokkenen worden gedwongen meer facetten bij de evaluatie te betrekken.
- Leraar kan eventueel ook meer te weten komen over het eigen lesgeven.

Aanvulling:

Het verdient misschien aanbeveling om als onderdeel van deze lijst het 'vakcijfer' op te nemen, echter gesplitst in bijvoorbeeld de drie laagste categorieën van Bloom (eventueel Klopfer en dergelijke).

Opmerking:

Bij beoordeling van leerlingen dienen de observatie en interpretatie en dus ook het oordeel onderscheiden te zijn.

## DEEL 4: BIJLAGEN

### INHOUD

	blz.
1. Voordracht Smit*	109
1.1. Leerplannen	109
1.2. Examenprogramma's	109
1.3. Functies van het examen	109
1.3.1. Een intern stimulerende functie	109
1.3.2. Een uniformerende en kontrolerende functie	110
1.3.3. Een afsluitende functie	110
1.3.4. Een toelatende functie	111
1.4. Nivoproblemen	111
1.5. Het opstellen van de eindexamens in de natuurkunde en scheikunde	112
2. Voordracht Berendts:*	
De C.M.L.N. en het eindexamen natuurkunde	113
3. Deelnemerslijst	119

\*De Velines-wintervergadering van 11 januari 1975 ging over de eindexamenproblematiek. Voor natuurkunde waren de inleidingen van Drs. J. Smit en Dr. B.Th. Berendts van belang.

Omdat eindexamens een stuk evaluatie vertegenwoordigen treft U in deze bijlage samenvattingen van hun inleidingen aan.

Het eerste stuk van Drs. J. Smit behandelt de kwestie van examenprogramma's in het algemeen, het tweede van Dr. B.Th. Berendts gaat over de recente leerplanvoorstellen van de C.M.L.N. en de bijbehorende examenprogramma's.



•

## DEEL 4: BIJLAGEN

### 1. Voordracht van Smit, gehouden op de Velinesdag d.d. 11 januari 1975

#### 1.1. Leerplannen

Iedere school moet voor elk van de te geven vakken een leerplan opstellen waarin men aangeeft hoe men zich het onderwijs in de verschillende leerjaren denkt. Men dient dus doelstellingen te formuleren en duidelijk aan te geven hoe men deze denkt te realiseren. Bij een dergelijk leerplan behoort een algemene lessentabel, die de verdeling aangeeft van de beschikbare leraarlessen over de verschillende vakken.

In de Wet op het Voortgezet Onderwijs worden in de artikelen 22, 23 en 24 enkele richtlijnen gegeven voor het opstellen van een leerplan met bijbehorende algemene lessentabel. Met inachtneming van deze richtlijnen stelt het bevoegd gezag, de rektor of directeur van de school en de leraren gehoord, het leerplan en de algemene lessentabel vast.

De op verzoek van de minister door een kommissie opgestelde (konsept-) leerplannen voor rijksscholen werden ter kennisneming toegezonden aan alle scholen. De bedoeling was, dat de niet-rijksscholen aan de hand van deze modellen een eigen leerplan zouden samenstellen. De meeste scholen (met name in de sektor vwo-havo) namen de voor rijksscholen opgestelde leerplannen ongewijzigd over. Daartoe waren zij echter niet verplicht.

#### 1.2. Examenprogramma's

De eindexamenprogramma's, die gelden voor alle scholen, zijn gebaseerd op de 'gemiddelde leerplannen'. Zij geven voor elk vak duidelijk aan welke vaardigheden er aan de hand van aangegeven leerstof getoetst gaat worden. Zij vormen een passende afsluiting op het leerproces, zoals dit in de doorsnee leerplannen is vastgelegd. Zijn de leerplannen, de eigenlijke uitgangspunten van het onderwijsgebeuren op de scholen dikwijls 'vergeten stukken', dit kan bepaald niet gezegd worden van de van de leerplannen afgeleide examenprogramma's. Te konstateren is dat de examenprogramma's in toenemende mate de plaats innemen van de leerplannen in het onderwijsproces op de scholen.

#### 1.3. Funkties van het examen

##### 1.3.1. Een intern stimulerende funktie

De resultaten, behaald bij een goed opgesteld examen, geven aan in hoeverre men er in geslaagd is de gestelde doeleinden te verwezenlijken.

Het is hier als toetselement meteen nivo bepalend. Het leerproces wordt zo opgezet dat de verwachte vaardigheden op het examen aanwezig kunnen zijn. Het examen heeft zo een stimulerende functie bij het onderwijsgebeuren. Hoewel examens eigenlijk niet de instrumenten zijn waarmee bepaalde gewenste veranderingen in het onderwijs mee moeten worden aangebracht, is hun invloed op vrijwel alle facetten van het onderwijs bijzonder groot. Examens hebben naast een toetsend karakter ook een sturend karakter gekregen.

#### 1.3.2. Een uniformerende en kontrollerende functie

Wanneer voor het vervullen van bepaalde functies in de maatschappij specifieke vaardigheden of kennis noodzakelijk zijn, is het prettig wanneer het bezit van bijvoorbeeld een mavo-diploma met een bepaald vakkenpakket de aanwezigheid van de gewenste kwaliteiten tot op zekere hoogte garandeert. Diploma's openen soms de weg tot vervolgstudies. Het vervolgonderwijs moet dan een basis hebben waarop verder kan worden gewerkt.

Er is dus vanuit de maatschappij en vanuit het vervolgonderwijs een drang naar uniformering van het onderwijs op scholen van hetzelfde type. Het examen speelt bij deze gewenste uniformering een grote rol. Het examen heeft hier ook een kontrollerende taak. Het garandeert maatschappij en vervolgonderwijs de aanwezigheid van bepaalde kwaliteiten en het licht docenten en leerlingen in over de wijze waarop het onderwijs is overgekomen.

#### 1.3.3. Een afsluitende functie (en wel ten opzichte van de maatschappij)

Leerplannen, examenprogramma's en examens moeten zo opgezet worden dat de leerlingen, die niet verder gaan studeren, een afgerond geheel van kennis en vaardigheden meekrijgen, waarmee zij in de maatschappij kunnen opereren en als volwaardige leden van de samenleving kunnen functioneren (bijvoorbeeld op een verstandige wijze mee kunnen denken en beslissen in milieuproblemen).

De examens moeten dus meewerken aan het leggen van een algemene werkbasis waar vanuit een leerling zich eventueel in zijn maatschappelijke functie verder kan specialiseren en waarmee hij tevens een volwaardige deelnemer aan onze maatschappij wordt.

1.3.4. Een toelatende functie (ten opzichte van het vervolgonderwijs)

Eén van de grondpijlers van het mammoetwetsysteem is het bieden van doorstroming tussen de verschillende schooltypen. Doodlopende wegen moeten zoveel mogelijk vermeden worden. Elk schooltype vormt een schakel in een groot geheel.

In het algemeen moet er nu voor gewaakt worden dat de eisen van het vervolgonderwijs het leerproces niet gaan vervormen. Toch moet er ook opgelet worden dat de aansluiting met het vervolgonderwijs niet wordt geblokkeerd. De examens moeten nu meewerken om verbindingswegen met het vervolgonderwijs niet wordt geblokkeerd. De examens moeten nu meewerken om verbindingswegen tussen de verschillende schooltypen open te houden.

1.4. Nivoproblemen (rond de examens mavo-havo-vwo)

Bij vergelijking van de huidige examens mavo-havo-vwo met de oude examens mulo en hbs valt het volgende op te merken.

Het onderwijs op een mavo-school is moeilijk te vergelijken met dat op een mulo-school. Het mavo wordt afgesloten door een examen in 5 of 6 vakken, het mulo werd afgerond met een examen in 12 vakken. De verandering van mulo en mavo bracht voor menig vak meteen inhoudelijk grote veranderingen mee (bij de natuurwetenschappen kregen het werken met diagrammen en grafieken en het praktikum meer aandacht, terwijl de vraagstukkenkultus werd teruggedrongen).

Gesteld kan worden dat een leerling op het mavo-examen in zijn geheel gezien een gelijkwaardige prestatie levert vergeleken met een leerling die aan het mulo-examen deelnam. Per vak een grotere prestatie (in de regel zijn er nu ook meer uren beschikbaar).

De havo-school kan voor de niet-exakte vakken als opvolger gezien worden van de mms. Voor de exakte vakken is in de oude situatie geen vergelijkbaar schooltype te vinden. Per school vallen er nogal verschillen in aanleg en capaciteiten van de leerlingen te konstateren. Het is buitengewoon moeilijk bij dit schooltype, waarvan het onderwijs nog bezig is uit te kristalliseren, nauwkeurig vast te stellen welke vaardigheden, welk begrip en inzicht men van de examenkandidaat mag verlangen.

Lettend op de positie tussen het vwo en het mavo en rekening houdend met de doorstroming naar het hbo zou op het examen een nivo verwacht mogen worden dat de leerlingen vroeger aan het begin van de eindexamenklassen hbs bereikten (momenteel bereikt slechts 40% van de havo-kandidaten die doorstromen naar het hbo en hto de tweede klas!).

Het vwo heeft vergeleken met de hbs een andere doelstelling (voorbereidend wetenschappelijk onderwijs) en minder examenvakken. Per vak mag van de leerlingen een grotere prestatie verwacht worden.

1.5. Het opstellen van de eindexamens in de natuurkunde en scheikunde

Voor elk van de examens natuurkunde mavo-3, mavo-4, havo en vwo wordt een advieskommissie gevormd; voor scheikunde zijn er advieskommissies voor de examens mavo (3 + 4), havo en vwo. Deze advieskommissies bestaan uit een aantal leraren, die daadwerkelijk betrokken zijn bij het onderwijs op het betreffende schooltype. De advieskommissies havo en vwo natuurkunde en vwo scheikunde worden begeleid door een deskundige, verbonden aan of in relatie staande met het universitaire onderwijs. Ieder jaar worden één of twee leden van elke kommissie vervangen.

Wanneer een advieskommissie het voorstel voor een examen gereed heeft (wat het meerkeuzedeel van het examen scheikunde betreft (havo) met de door de CITO-schrijfgroep opgestelde toets akkoord gaat) vindt een bespreking van het geheel plaats met de vakinspektors en de vakdeskundige leden van de Commissie Vaststelling Opgaven. De Commissie Vaststelling Opgaven stelt uiteindelijk het werk vast.

De advieskommissies stellen tevens kon ept-normen op. Na de examens worden de definitieve normen vastgesteld door de normenkommissies, die bestaan uit vert genwoordigers van de Vakinspektors, de vakinspektors en de vakdeskundige leden van de Commissie Vaststelling Opgaven.

2. Voordracht Berendts, gehouden op de Velinesdag d.d. 11 januari 1975:  
De C.M.L.N. en het eindexamen natuurkunde

Na het interimrapport uit 1969 is nu een tweede rapport van de C.M.L.N. gereed gemaakt, dat inmiddels door de minister is vrij gegeven en zeer binnenkort op grote schaal zal worden verspreid. Ten opzichte van het interimrapport zijn er vele verschillen aan te wijzen.

In 1969 was er een apart deeltje over de *onderbouw*. Dit bevatte een uitvoerig leerstofoverzicht met veel nieuwe gezichtspunten. De bedoeling was, zoals uitdrukkelijk er vóór werd vermeld (maar helaas niet werd begrepen), dat dit overzicht zou dienen als *ideeënbus* voor de leraar, om hem behulpzaam te zijn bij het advies aan het bevoegd gezag van zijn school over het leerplan. Doordat men er ten onrechte een 'must' in zag is dit deel ongunstig ontvangen. In het nieuwe rapport is de tekst van het oude onderbouwrapport ongewijzigd opgenomen als bijlage.

Er aan toegevoegd is een prioriteitenlijst, samengesteld na verwerking van het kommentaar van de leraren, met in een aantal gevallen argumenten. Gehoopt wordt dat duidelijker dan voorheen hierdoor de functie van ideeënbus naar voren komt.

Voor de *andere delen van het rapport* werd een nieuwe tekst opgesteld, rekening houdende met de ontvangen commentaren. Daarbij is in de leerstoflijst steeds met teken-tjes aangegeven van welke onderwerpen men verwacht, dat ze in de onderbouw worden afgehandeld of althans in een eerste ronde aandacht krijgen. Deze delen staan hiermee volledig los van de onderbouw-bijlage. Rekening moest verder worden gehouden met veranderde randkondities, zoals vervanging van het mondeling examen door het schoolonderzoek en het feit, dat niet langer mag worden verwacht dat alleen een zekere elite in de examenklassen komt op grond van hun positieve belangstelling voor het vak, maar een veel grotere groep, die natuurkunde nodig heeft voor toelating tot verdere studie. Voor U van direkt belang zijn de delen IV, V en VI van het rapport, die respektievelijk behandelen de mavo, havo en het vwo en het deel VIII, dat een voorstel inhoudt over de examenprogramma's.

Duidelijk zij gesteld, dat het overal gaat om voorstellen voor leerplannen, waarvan de inhoud van school tot school mag verschillen. Daar alle scholen wel moeten deelnemen aan een centraal schriftelijk examen, dat voor de helft het cijfer voor het betreffende vak bepaalt, moeten natuurlijk alle leerplannen daar op voorbereiden. Daarom heeft de C.M.L.N. ook gekeken, welke konsekventies de voorstellen hadden voor het *examenprogramma*.

Hieruit is een voorstel gekomen, dat aan de minister is voorgelegd. Bij het voorbereiden van het examenprogramma bestond tevens gelegenheid tot coördinatie, voor wat betreft naamgeving en details tussen de rapporten voor de verschillende schooltypen onderling.

Tegen de *bestaande examenprogramma's* zijn een aantal bezwaren aan te voeren:

- er is overlading, zodat een deel van de stof, dat niet tot de voorgeschreven groepen hoort, vaak helemaal niet aan bod komt
- er is onduidelijkheid, met name bij havo (althans tot voor kort)
- vrijheid voor de leraar is er in de praktijk heel weinig bij
- door de groepenkeuze bij vwo weet het vervolgonderwijs nooit precies waar men aan toe is

Op basis van het interimrapport is geschat hoe het zat met de benodigde tijd. Met name voor havo is hier door leraren binnen en buiten C.M.L.N. zo goed mogelijk nagegaan hoeveel lessen nodig zouden zijn voor behandeling van de stof, repetities, e.d. Op grond hiervan is de stof in omvang beperkt, zoveel mogelijk met behoud van het gewenste nivo. Bij deze schattingen is uitgegaan van de volgende uren tabellen:

voor mavo-4: 2, 2, 5 of 2, 3, 4, uren in de klassen 2, 3 en 4  
voor havo : 2, 3, 4 en 4 uren in de klassen 2 t/m 5  
voor vwo : 2, 3, 3, 4 en 4 uren in de klassen 2 t/m 6

Bij de opzet van de leerplannen is gestreefd naar een zekere *ruimte* voor de leraar. Binnen deze ruimte moet de leraar vrijheid hebben bij de keuze van een stuk stof. De uitersten van volledige vrijheid bij het bepalen van de onderwerpen en volledig in details alle te behandelen onderwerpen voorschrijven zijn beide onmogelijk. Gemikt is op een tussenweg, waarbij de praktijk van examens en wensen van het opvolgende onderwijs die ruimte niet zeer groot hebben gelaten, maar naar de mening van de commissie is deze wel aanwezig.

Hier moet plaats zijn voor een hobby van de leraar voor een bepaald onderdeel van ons vak. Door over een dergelijk onderwerp te kunnen spreken zal hij ook enthousiaster les kunnen geven dan bij andere onderwerpen. Dit zal op zijn beurt bijdragen tot verlevendiging van de lessen. De leerlingen zullen er een betere motivatie vóór (of tegen) een verdere studie, waarbij dit vak een rol speelt, door krijgen. Het kan de mogelijkheid geven van een soort afstudeeronderwerp; een stukje natuurkunde, waar delen van een aantal hoofdstukken gebruikt worden en de samenhang duidelijker te geven is. Een belangrijk voordeel is ook dat via deze

ruimte er voortdurend mogelijkheid is om het onderwijs bij-de-tijd te houden. Een nieuw onderwerp kan eerst de keuze zijn van enkelen; daarna van meerderen die het idee overnemen, om tenslotte bij gebleken geschiktheid en volgroeide inhoud en didaktiek opgenomen te worden in de leerstofkern in ruil voor een verouderd onderwerp.

Bij *vwo* is naast een *kernleerstof* ingevoerd een aantal *keuzegroepen*. In het voorstel staat nu dat in de laatste klas twee keuzeonderwerpen van elk 15 uren (all in) behandeld moeten worden. De opzet uit het interimrapport van drie onderwerpen van 10 uren elk geeft per onderwerp te weinig armslag. Met 4 uren per week betekent dit circa 2 maanden van de zesde klas.

Wat de examenprocedure voor deze keuzeonderwerpen betreft is gekozen voor het volgende systeem. De inspectie wijst elk jaar tijdig enkele (waarschijnlijk twee) onderwerpen aan, waarvan de school er één doet. Hiervoor worden vragen opgesteld bij het schriftelijk deel van het examen. Voor de bepaling van het tweede onderwerp is de school geheel vrij. Het wordt geëxamineerd bij het schoolonderzoek, zodat de keuze in het begin van het examenjaar, tegelijk met de overige gegevens over de wijze waarop men het schoolonderzoek denkt te doen, aan de inspecteur moet worden doorgegeven.

Wat het tweede onderwerp betreft bestaat bijvoorbeeld ook de vrijheid dat leerlingen in dezelfde klas verschillend kiezen. In feite kiest de leraar samen met zijn klas. Het is bijvoorbeeld denkbaar dat de leerlingen, die verder willen gaan in een richting waarbij natuurkunde een hoofdrol speelt, best een moeilijke kluit aan kunnen, terwijl de overige leerlingen beter gediend zijn met een meer traditioneel onderwerp of juist een aan hun situatie aangepast thema.

Van een aantal onderwerpen is de inhoud al vastgelegd, terwijl er teksten voor klaar zijn en experimenten mee zijn uitgevoerd, waarmee de haalbaarheid is getest. Daarbij is het beslist niet zo dat het alleen gaat om onderwerpen die op de universiteit thuis horen. In feite kan men drie categorieën onderscheiden:

- stukjes oude leerstof die uit de examenlijst zijn geschrapt, maar die door veel leraren nog graag behandeld worden
- oudere onderwerpen op een modernere wijze aangevuld en in een nieuw jasje gestoken
- nieuwe onderwerpen die tot nu toe niet tot de leerstof van het vwo hoorden



Voorbeelden van de eerste groep zijn: mechanica (een samenstel van diverse onderwerpen), geometrische optica, golfoptica (aanvulling op een aantal wel behandelde punten hiervan). Tot de tweede categorie zijn te rekenen: stroming van vloeistoffen en gassen (laminaire en turbulente stroming, wet van Bernoulli, lifteffekt bij vliegtuigen), trillingen en golven (uitgebreider dan in de kernleerstof, met vergelijkingen op diverse terreinen van de natuurkunde). Tot de nieuwe onderwerpen horen: de vaste stof; astrofysica; fysische elektronica (elektronenlens, elektronenstraalbuis, massaspektrometer, versnellers, cyclotron, e.d.); relativiteitstheorie; elementaire deeltjes; quantummechanica en meteorologie (in twee versies, namelijk fysische meteorologie: vorming van wolken, mist, stormen, e.d. en dynamische meteorologie, waar het gaat over fronten, weerkaartjes, e.d.).

Als het nieuwe examenprogramma zou worden ingevoerd in 1978 zal van de meeste onderwerpen gebruik gemaakt kunnen worden. De beschikbare teksten maken zelfwerkzaamheid van de leerlingen mogelijk, waardoor het behandelen van twee onderwerpen parallel in een klas realiseerbaar wordt. In de toekomst zullen min of meer kontinu nieuwe onderwerpen aan de lijst worden toegevoegd; zo wordt op het moment een onderwerp biofysica overwogen.

Voor *havo* is het doel, ruimte scheppen voor eigen inbreng, gelijk. De uitvoering is hier anders gedacht. De leraar dient een keuze te maken uit de volgende mogelijkheden:

- Meer dan normaal aandacht besteden aan leerlingenpraktika, hetgeen hij in het schoolonderzoek tot uiting kan brengen, door bijvoorbeeld één tentamen als praktikum in te richten.
- Hij kan de leerlingen oefenen in zelfstandige formulering van verbanden en wetten, door van bepaalde problemen buiten de eigenlijke leerstof uit te gaan en de leerlingen door analogie met wel bekende wetten te laten zoeken naar de samenhang. Hij kan bijvoorbeeld op elk schoolonderzoektentamen een op deze wijze gekonstrueerde vraag stellen.
- Hij kan een onderwerp buiten de leerstof kiezen voor uitgebreidere behandeling, zoals de keuzegroepen op het vwo, met alle voordelen als daar genoemd. Daarover kan een heel tentamen van het schoolonderzoek gaan.
- Tenslotte zijn er in de leerstoflijst een aantal punten 'naar rechts' geplaatst, waaruit hij een keuze kan maken. Dit zijn zaken die direkt aansluiten bij wel tot de examenlijst horende onderwerpen.

Examinering kan verspreid door het schoolonderzoek heen geschieden. Op deze wijze heeft de havo-leraar meer vrijheidsgraden dan op het vwo. Ook wordt hier geen omvang genoemd van de extra onderwerpen; wel wordt in het rapport voorgesteld dat de examinering voor circa een kwart van het schoolonderzoek zal bestaan uit deze extra's. De beperkte tijd op de mavo heeft er toe geleid dat geen poging is gedaan een soortgelijke scheiding ook daar door te voeren. Het mavo-programma zal dus uit een vaste reeks onderwerpen bestaan.

Voor het opstellen van de *onderwerpenlijst voor het examen* zijn de onderwerpen voor de schooltypen in vier kolommen naast elkaar geplaatst (mavo-3, naast mavo-4, havo en vwo). Op deze wijze ontstaat horizontaal een beeld er van hoe ver een onderwerp op de verschillende schooltypen wordt uitgediept. Hierbij is het havo-programma de spil geweest.

Duidelijk blijkt uiteraard van links (mavo-3) naar rechts (vwo) een opklimming in moeilijkheidsgraad en abstraktienivo. Om dit nog duidelijker te illustreren dan in nuchtere onderwerpenlijsten mogelijk is, zal voor een paar onderwerpen de stof voor de diverse schooltypen uitvoeriger worden opgeschreven. Een begin hiermee is reeds gemaakt voor het onderwerp magnetisme. Dit moet voor de praktijk een steun zijn om te bepalen hoever men moet trachten te komen bij de behandeling. Vaak blijkt het verschil uit toevoegingen bij de omschrijving die een beperking inhouden (zoals 'kwalitatief'; 'alleen voor een homogeen veld' e.d.).

Uit de naast elkaar geplaatste lijsten volgt ook dat het aantal onderwerpen voor mavo beperkt is vergeleken met de andere schooltypen. Toch zijn er onderwerpen die wel in de mavo-lijst staan en niet in de andere. Het betreft hier dan zaken die meestal op havo en vwo toch wel aan de orde zullen komen, maar daar niet specifiek voor het examen bestudeerd behoeven te worden (voorbeelden: wet van Archimedes, smeltwarmte, foto-toestel). Bij vwo ziet men veel meer nadruk op de wiskundige formulering dan bij havo. Moeilijke onderwerpen als isothermen, massadefekt, e.d. komen op de havo-lijst niet voor. Daar staat tegenover dat voor havo een aantal praktische zaken genoemd zijn, zoals werkingen van licht, kortsluiting en overbelasting, elektronica en vaste stoffen. Het is niet zo dat alle informatie voor de praktijk in de examenlijst zit; men moet wel steeds teruggrijpen op de uitgewerkte leerplanvoorstellen voor het betreffende schooltype.

In het examenvoorstel is ook getracht een formulering te geven van de *vaardigheden* die van een leerling verwacht worden, dus wat hij kan beginnen met de verworven kennis. Deze opsomming is concentrisch opgebouwd, dus eerst wat verwacht wordt van een mavo-leerling, daarna hetgeen meer verwacht wordt van iemand, die examen havo doet en tenslotte de extra bedrevenheid die men onderstelt bij een vwo-er. Een 25-tal vaardigheden worden hier onderscheiden.

De mavo-leerling moet, als voorbeeld, met een bekende relatie de waarde van een grootte kunnen berekenen als de nodige gegevens verstrekt worden. Een havo-leerling moet ook zelf de nodige gegevens selekteren, interpoleren en grootte-orde van de fout kunnen schatten. De vwo-leerling moet bovendien een eenvoudige dimensieberekening kunnen uitvoeren. Zo in een lijst geplaatst is dit allemaal erg droog en weinig operationeel. Om voor de praktijk steun te verlenen wordt momenteel aan de afdeling didaktiek van de V.U. een poging gedaan een aantal vragen te selekteren, waarmee de diverse vaardigheden getoetst kunnen worden. De verwachting is dat hiervan op korte termijn resultaten beschikbaar zullen zijn, zodat er bij de beslissingen van de Rijksexamenprogramma-kommissie rekening mee gehouden kan worden.

De oproepen namens de minister aan de diverse organisaties, die leden moeten noemen voor die examenprogramma-kommissie zijn al verstuurd, ook aan Velines. Gehoopt wordt dat er op korte termijn zaken gedaan kunnen worden, zodat het nieuwe programma in 1978 (optimisten hoopten al 1977) kan worden gevolgd. We hopen dat het geheel in de praktijk zal blijken te voldoen, zodat een behoorlijk peil van onderwijs gehandhaafd kan worden en de leerlingen later iets kunnen bereiken met de verworven kennis en inzichten in de eigenschappen van de ons omringende natuur.