

## De ontwikkeling van een diagnostisch instrument voor de vaststelling van specifieke intuïtieve ideeën over stroom en spanning

P.Licht

Vakgroep Didactiek en Practica Natuurkunde  
Vrije Universiteit Amsterdam

### Summary

*A number of educational studies point out that pupils use their intuitive ideas and beliefs about the natural world not only before but also during and after formal education. It is difficult to trace and diagnose these ideas in classroom teaching. In this article we describe the development of a diagnostic test, as a first step to determine the possible existence of certain coherent intuitive ideas about current and voltage; ideas which are already traced by previous research among other pupils. The intermediate result of this test development is a test of 28 questions including 4 more or less reliable scales concerning 'current consumption', 'the battery as a constant current supply', 'local and sequential reasoning' and 'the lack of or an incorrect discrimination between current and voltage'.*

### 1. Het onderzoekskader

Eerder beschreven wij al in dit tijdschrift een onderwijsstrategie die dient als leidraad voor ons onderzoek (Licht, 1986a). Het is niet onze bedoeling nogmaals uitvoerig op deze strategie in te gaan. Maar voor een goed begrip van het vervolg moeten we er in deze inleiding enkele opmerkingen over maken. De strategie wordt gekenmerkt door twee componenten, te weten een conceptuele en een contextuele component. Binnen de *conceptuele component* werden in eerste instantie twee niveau's onderscheiden, een intuïtief niveau en een (vakwetenschappelijk) theoretisch niveau. In tweede instantie is daar een (tussenliggend) beschrijvend niveau aan toegevoegd, dat leidt tot op zichzelf staande empirische generalisaties, die betrekking hebben op de relatie tussen zichtbaar of meetbaar te maken kenmerken van concepten in een beperkt aantal contextsituaties (Licht, 1987).

Binnen de *contextuele component* wordt onderscheid gemaakt tussen school- en praktijkcontexten. Door deze opsplitsing van de strategie in twee componenten wordt het mogelijk te spreken over concepten op intuïtief, beschrijvend en theoretisch niveau binnen school- en praktijkcontexten.

In dit artikel richten we ons op de ontwikkeling van een diagnostisch instrument voor de vaststelling van intuïtieve ideeën binnen bepaalde schoolcontexten. Het gaat ons om de leerlingideeën over stroom en spanning binnen eenvoudige elektrische schakelingen met lampjes en batterijen. In de door ons voorgestelde onderwijsstrategie is het namelijk van belang in een vroeg stadium van het onderwijsleerproces zicht te krijgen op de mogelijk aanwezige intuïtieve ideeën; dit met de bedoeling met het onderwijs in te haken op deze ideeën. Het te ontwikkelen diagnostisch instrument zou dan deel kunnen uitmaken van het begin van een leerstofeenheid over stroom en spanning. Op grond van de resultaten dienen leerlingen dan op een gedifferentieerde manier verwezen te kunnen worden naar een bepaald onderwijsprogramma.

Maar zover is het nog niet. In dit artikel beschrijven we een aantal stappen in de ontwikkeling van een betrouwbaar en valide diagnostisch instrument. Uitspraken over de mate van betrouwbaarheid baseren we op psychometrische gegevens. Uitspraken over de mate van validiteit zijn in die zin prematuur, dat we nog geen ervaring hebben met bepaalde onderwijsprogramma's waarin nader wordt ingegaan op de gediagnosticeerde intuïtieve ideeën. We beschikken in deze fase van de ontwikkeling slechts over uitspraken van leraren m.b.t. de kwaliteit van vragen, die ons informatie geven over de inhoudsvaliditeit. In §2 kijken we terug op twee schriftelijke testen waarmee we de begrips- en redeneerproblemen rond stroom en spanning hebben verkend. In §3 komen we tot een lijst van eisen vanuit het onderwijs, het onderzoek en de psychometrie waaraan het te ontwikkelen instrument moet voldoen. In §4 en §5 gaan we in op deze ontwikkeling en op de analyse van de gegevens die met dit instrument zijn verzameld. In §6 komen we tot slot tot enkele conclusies over de kwaliteit van het instrument.

## **2. Wat aan de ontwikkeling voorafging: een eerste en een tweede test**

*De eerste test* die is ontwikkeld om begrips- en redeneerproble-

men op het spoor te komen, is elders uitvoerig beschreven (Licht 1986a; Kuiper e.a., 1985). In deze test ging de meeste aandacht uit naar de leerlingideeën omtrent stroom, het onderscheid tussen stroom en spanning en wat er gebeurt als ergens in een schakeling een verandering wordt aangebracht. De test bestond uit 12 vragen over schakelschema's, in totaal opgebouwd uit 46 vraagonderdelen. De meeste vragen waren afkomstig uit buitenlands onderzoek en hadden tot doel vast te stellen in hoeverre de begrips- en redeneerproblemen van Nederlandse leerlingen vergelijkbaar zijn met die van buitenlandse leerlingen. De analyse van de gegevens richtte zich op frequenties per vraagonderdeel en op enkele kruistabellen van antwoorden waar- tussen enige samenhang werd verwacht. Daar het een eerste verkenning betrof van de aard en omvang van begrips- en redeneerproblemen binnen het onderwerp elektriciteit, kreeg het aspect van samenhangende vragen nog nauwelijks aandacht.

*De tweede test* is ontwikkeld in samenwerking met collega's in vier andere landen. De resultaten van de Nederlandse leerlingen zijn reeds beschreven (Licht en Snoek, 1986). De test bestond uit 13 vragen, onderverdeeld in 41 vraagonderdelen. Het merendeel van de vragen (9) werd gesteld aan de hand van schakelschema's, de overige aan de hand van realistisch getekende schakelingen. De analyse richtte zich op frequenties, kruistabellen van vraagonderdelen en vergelijking van de resultaten in de diverse landen (Van Rhöneck e.a., 1987).

Op grond van de resultaten op de eerste en de tweede test identificeerden wij vier probleemvelden, waar het gaat om de begrippen stroom en spanning, en om het redeneren in probleemsituaties met eenvoudige elektrische schakelingen:

1. het idee van (geheel of gedeeltelijk) stroomverbruik in een lamp of weerstand (dit probleemveld geven we in het vervolg aan met de *code SV* van stroomverbruik);
2. het idee dat een batterij of stopcontact altijd dezelfde hoeveelheid stroom levert, onafhankelijk van de karakteristieken van de schakeling (*code CS* van constante stroom);
3. het lokaal en sequentieel redeneren in serie- en parallelschakelingen. In plaats van een redeneerwijze waarbij alle delen van een schakeling met elkaar in verband staan, denken veel leerlingen dat een verandering in een schakeling alleen lokaal (= ter plaatse van de verandering) of 'stroomafwaarts' (beke-

ken vanaf het punt van verandering) gevolgen heeft (*code LS* van lokaal en sequentieel);

4. veel leerlingen maken ófwel geen onderscheid tussen stroom en spanning ófwel ze schrijven de kenmerken van stroom toe aan spanning en omgekeerd (*code VI*, de symbolen van spanning en stroomsterkte).

Het is mogelijk gebleken met deze eerste en tweede test een globale diagnose te stellen van de aard en omvang van de aanwezige intuïtieve ideeën onder grote groepen leerlingen. Het diagnostisch gebruik bij individuele leerlingen lijkt ons echter riskant, omdat het aantal met elkaar samenhangende vragen te gering is. Een valide verwijzing naar een bepaald onderwijsprogramma is dan onmogelijk. Bovendien hopen we de betrouwbaarheid van het instrument te verhogen door per probleemveld meer samenhangende vragen op te nemen. Met het hier te bespreken instrument willen we de mogelijkheid van het diagnosticeren van individuele leerlingideeën dichterbij brengen.

### **3. Eisen waaraan het diagnostisch instrument moet voldoen**

De aan het diagnostisch instrument te stellen eisen komen voort uit het onderwijs, het onderzoek en de psychometrie.

Vanwege de toepassing in het onderwijs stellen we de volgende eisen:

1. af te nemen in één lesuur en bij voorkeur door de leerlingen zelf na te kijken m.b.v. een sleutelblad;
2. het resultaat op clusters van samenhangende vragen moet gerichte verwijzing mogelijk maken naar een remediërende leeractiviteit.

Daarnaast moeten de verkregen leerlingenantwoorden een bijdrage kunnen leveren aan de beantwoording van de volgende twee onderzoeksvragen:

- a. hoe hardnekkig zijn de intuïtieve ideeën in de loop van het voortgezet onderwijs;

Hierbij valt te denken aan het gebruik van het instrument door leerlingen in alle leerjaren van havo en vwo. Na analyse van de resultaten uit de onderscheiden leerjaren hopen we de onderzoeksvraag te kunnen beantwoorden, o.m. door regressieanalyse toe te passen, waarbij we de resultaten op afzonderlijke en op samenhangende vragen uitzetten tegen het leerjaar (zie ook Thijs, 1987). Geen of een geringe verbetering in resultaat betekent dan dat het gehanteerde intuïtieve idee nogal resistent is

tegen onderwijs. Deze ideeën mogen we dan, met enig recht, preconcepten noemen. Een verbetering in resultaat over de leerjaren betekent dat het intuïtieve idee betrekkelijk eenvoudig gecorrigeerd kan worden en geen speciale aandacht in een onderwijsprogramma behoeft.

b. in hoeverre treden veranderingen op in de intuïtieve ideeën na een onderwijsprogramma rond één van de in par. 2 genoemde probleemvelden?

Hierbij valt te denken aan het meermalen hanteren van het instrument in één leerlingpopulatie. Zowel bij a als b is bruikbaarheid onder leerlingen, die nog geen elektriciteitsonderwijs hebben gehad, een vereiste.

Vanwege de toepassing in het onderzoek stellen we de volgende eisen:

3. bij voorkeur gebruik maken van zo realistisch mogelijk getekende schakelingen. Schakelschema's zijn slechts toelaatbaar als de gebruikte symbolen worden toegelicht;
4. begrippen die in de test niet centraal staan - zoals weerstand en ampèremeter - moeten worden toegelicht;
5. het taalgebruik moet zijn toegesneden op de onderbouw-leerling;
6. het instrument moet worden opgebouwd rond de vier eerder genoemde probleemvelden, gecodeerd met SV, CS, LS en VI. Om tot een betrouwbaar en valide instrument te komen, voegen we tot slot respectievelijk een zevende en achtste eis toe:
7. de met elkaar samenhangende vragen binnen de vier probleemvelden moeten kunnen worden ondergebracht in schalen met een betrouwbaarheidscoëfficiënt groter dan 0,70. Wij menen af te kunnen wijken van de meer gebruikelijke eis (> 0,90), omdat het hier gaat om individuele beslissingen van diagnostische en niet van selectieve aard;
8. minstens 70% van de variantie in de totaalscores van de leerlingen moet 'verklaard' kunnen worden door de variantie in de schaalscores m.b.t. de probleemvelden SV, CS, LS en VI. Met andere woorden: minstens 70% van de gemaakte fouten moet toegeschreven kunnen worden aan de intuïtieve ideeën die ten grondslag liggen aan de te vormen schalen.

Over de validiteit van het instrument merken we slechts op dat drie ervaren natuurkunde leraren het vooraf met elkaar eens moeten zijn over de interpretatie van een bepaald gekozen antwoord, bekeken in het licht van de vier eerdergenoemde probleemvelden.

#### 4. De ontwikkeling van het instrument en de codering van de antwoorden

In globale termen is de ontwikkeling van het instrument, met een veronderstelde samenhang tussen groepen van vragen, als volgt verlopen.

*stap 1:* constructie van het instrument, waarbij er vooraf argumenten zijn om aan te nemen dat de leerlingantwoorden op bepaalde vragen zullen samenhangen (zie §3 over inhoudsvaliditeit);

*stap 2:* statistische analyse van de antwoorden met bijvoorbeeld factoranalyse. Met factoranalyse worden met elkaar samenhangende antwoorden opgespoord. Deze antwoorden laden dan hoog op de betreffende factor. Een belangrijke vraag wordt dan of de statistisch gevonden factor door ons te interpreteren valt. Als de gevonden factorstructuur goed overeenstemt met de bij stap 1 veronderstelde structuur, leidt stap 2 tot grote vooruitgang in de ontwikkeling van een betrouwbaar en valide instrument. Voor het vaststellen van de betrouwbaarheid verwijzen we naar stap 3. Het goed overeenkomen van de factorstructuur vooraf - vastgesteld na overleg met leraren - met de factor-structuur achteraf, vatten we op als een voorlopige indicator voor de validiteit;

*stap 3:* de vragen die vallen in een door ons te interpreteren factor kunnen vervolgens worden samengenomen tot een schaal. De betrouwbaarheid van deze schaal kan nu worden vastgesteld en ook kan per leerling een schaalscore worden berekend. Een schaalscore is gebaseerd op *alle* antwoorden die hoog laden op de betreffende factor en is om die reden betrouwbaarder voor de diagnose dan het antwoord op een enkele vraag;

*stap 4:* er kunnen verschillende redenen zijn waarom men niet tevreden is met het uiteindelijke resultaat. Zo is het mogelijk dat sommige vragen niet duidelijk tot één factor behoren; ook kan de betrouwbaarheid van bepaalde schalen nog te wensen over laten. In dit soort situaties valt te overwegen het instrument bij te stellen en de stappen 1 t/m 3 opnieuw te doorlopen. De ontwikkeling van de derde test kan gezien worden als een eerste serieuze poging om tot een diagnostisch instrument te komen. Vanwege onze ervaringen met de testen één en twee, en vanwege de toepassing in klassen, is gekozen voor een schriftelijke test met overwegend gesloten vragen. De vragen in de derde test zijn gegroepeerd rond de vier eerder genoemde pro-

bleemvelden. De test omvat 28 vragen opgedeeld in 59 gesloten en 21 open onderdelen. De open onderdelen betreffen telkens de vraag om een toelichting op het eerder aangekruiste antwoord in de gesloten vraag. De eerste 24 vragen hebben betrekking op zo realistisch mogelijk getekende schakelingen met lampjes en batterijen. De laatste 4 vragen gaan over schakelschema's, die dezelfde structuur hebben als enkele van de eerder gepresenteerde schakelingen. Dit soort contextwisselingen wordt toegepast om te zien in hoeverre leerlingen consistent zijn in hun redeneringen. Overigens worden de in deze schema's gehanteerde symbolen in deze schema's voor lampje, batterij en ampèremeter wel eerst toegelicht. Niet elk van de vraagonderdelen leidt tot een zelfstandige variabele in de uiteindelijke analyse. De antwoorden op sommige vraagonderdelen worden gecombineerd tot een nieuwe variabele. Een variabele krijgt de waarde 1 als het achterliggende antwoord of combinatie van antwoorden door ons kan worden geïnterpreteerd als voortkomend uit één van de in §2 vermelde leerlingideeën. Een variabele krijgt de waarde 0 als het achterliggende antwoord of combinatie van antwoorden correct is, of door ons niet te interpreteren valt in het licht van de vermelde intuïtieve ideeën. We coderen in het laatste geval het antwoord niet als 'missing value', omdat we willen controleren of we voldoen aan (de psychometrische) eis (8) uit §3. We zijn immers vooral geïnteresseerd in de interpretatie van bepaalde fouten en voegen om die reden niet te interpreteren fouten samen met correcte antwoorden. Op deze manier ontstaat een bestand van 49 variabelen per leerling, gecodeerd in enen en nullen. Dit biedt ons de mogelijkheid om via factoranalyse<sup>1</sup> op zoek te gaan naar een structuur in deze variabelen. Alvorens nader in te gaan op de resultaten willen we per probleemveld een voorbeeld geven van de constructie van variabelen.

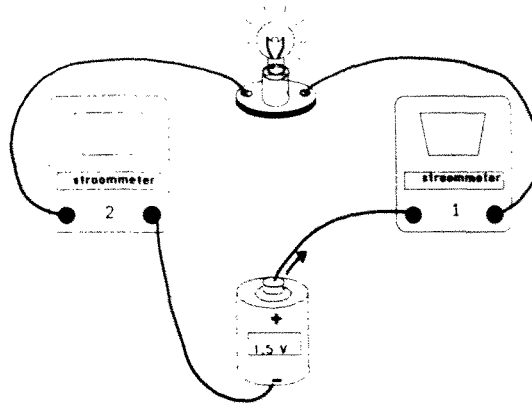
*voorbeeld 1: probleemveld SV (stroomverbruik, fig.1)*

Als vraag 12 wordt beantwoord met c of d krijgt de variabele SV121 de waarde 1 anders de waarde 0. De antwoorden c en d worden door ons geïnterpreteerd als voortkomend uit een stroomverbruikidee.

*voorbeeld 2: probleemveld CS (constante stroom, fig.2)*

Variabele CS71 wordt gebaseerd op een antwoordcombinatie van de twee gesloten vraagonderdelen in vraag 7. Als vraag 7.1 en 7.2 worden beantwoord met a, krijgt variabele CS71 de waarde 1, anders de waarde 0. De antwoorden op de open vraagonder-

*De grootte van een elektrische stroom kun je meten met een stroommeter (in de natuurkunde noemen we zo'n meter een ampèremeter). Je ziet hier een lampje aan beide kanten verbonden met een stroommeter. De twee stroommeters zijn verbonden met een batterij. Het lampje brandt. De pijl geeft de richting van de elektrische stroom aan.*



*Stroommeter 2 meet*

- a. een grotere elektrische stroom dan stroommeter 1;*
- b. een even grote elektrische stroom als stroommeter 1;*
- c. een kleinere elektrische stroom dan stroommeter 1;*
- d. geen elektrische stroom.*

Fig.1: vraag 12

delen van vraag 7 hebben geen invloed op deze codering, maar leiden in bijna alle gevallen tot een ondersteuning van onze interpretatie, vanwege uitspraken als: "alle stroom gaat nu naar lampje 1" of "de stroom hoeft nu niet meer verdeeld te worden". De vraag zou nu kunnen rijzen waarom we niet tevreden zijn met deze ene variabele die duidelijkheid geeft over de mogelijke aanwezigheid van een constant-stroom idee. Het gaat ons er evenwel om vast te stellen of dit intuïtieve idee voor leerlingen ook vruchtbaar is in situaties die enigszins afwijken van de situatie in vraag 7. Pas dan is er naar onze mening sprake van een intuïtief idee waarmee in het onderwijs rekening gehouden moet worden.



In nevenstaande schakeling draaien we lampje 2 los en nemen deze uit de fitting. Lampje 2 is dan uit. Wat gebeurt er met lampje 1? Kruis aan wat juist is.

- a. lampje 1 gaat feller branden;
- b. lampje 1 blijft even fel branden;
- c. lampje 1 gaat zwakker branden;
- d. lampje 1 gaat uit.

Ik heb dit antwoord gekozen, omdat

.....

We draaien nu lampje 2 weer vast. Daarna draaien we lampje 1 los en namen deze uit de fitting. Lampje 1 is dan uit.

Wat gebeurt er met lampje 2? Kruis aan wat juist is.

- a. lampje 2 gaat feller branden;
- b. lampje 2 blijft even fel branden;
- c. lampje 2 gaat zwakker branden;
- d. lampje 2 gaat uit.

Ik heb dit antwoord gekozen, omdat

.....

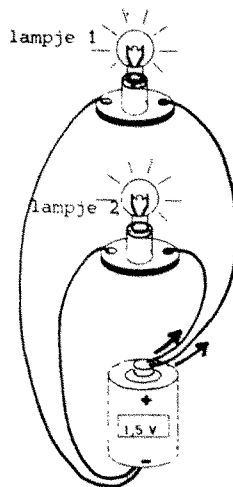


Fig.2: vraag 7

voorbeeld 3: probleemveld LS (lokaal-sequentieel, fig.3)

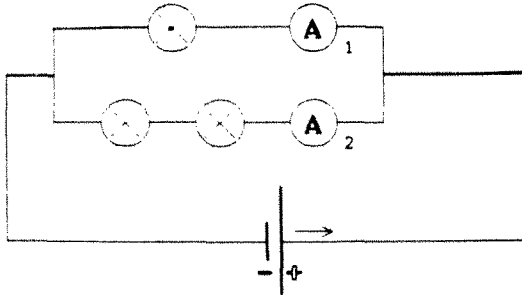
Als vraag 27 wordt beantwoord met "even grote" wordt dit door ons geïnterpreteerd als voortkomend uit een sequentiële redenering in de trant van 'de stroommeters staan vòòr de lampjes', of uit een lokale redenering in de trant van 'de stroom splitst zich bij de vertakking in gelijke delen'. Variabale LS71 krijgt dan de waarde 1.

voorbeeld 4: probleemveld VI (spanning-stroom onderscheid, fig.4)

De variabele VI231 krijgt de waarde 1 als de vraag 23.1 en vraag 23.2 als volgt is beantwoord:

- 23.1 0,6/0,6/0,6 A    en 23.2 1,5/1,5/1,5 volt,
- òf    23.1 0,2/0,2/0,2 A    en 23.2 0,5/0,5/0,5 volt,
- òf    23.1 0,6/0,6/0,6 A    en 23.2 0,5/0,5/0,5 volt.

In het schema hieronder zijn alle drie de lampjes gelijk.



Vergelijk stroommeter 1 met stroommeter 2.

Stroommeter 1 meet een /grotere/even grote/ kleinere/ stroom dan stroommeter 2.

Fig.3: vraag 27

De lampjes in de schakeling hiernaast zijn alle gelijk. Ze zijn aangesloten op een batterij van 1,5 volt. De stroommeter wijst een stroom aan van 0,6 ampère.

Hoe groot is de elektrische stroom door elk lampje?  
door lampje  $L_1$  ..... ampère;  
door lampje  $L_2$  ..... ampère;  
door lampje  $L_3$  ..... ampère.

Hoe groot is de elektrische spanning over elk lampje?  
over lampje  $L_1$  ..... volt;  
over lampje  $L_2$  ..... volt;  
over lampje  $L_3$  ..... volt.

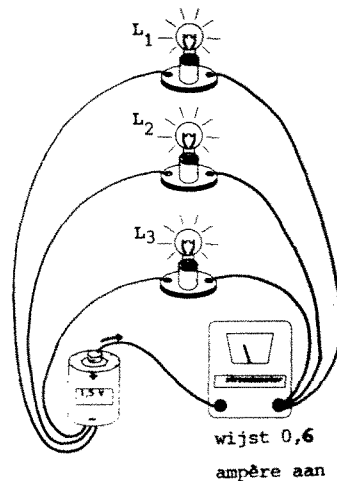


Fig.4: vraag 23

De antwoordcombinaties worden als volgt geïnterpreteerd: in de eerste twee antwoordcombinaties wordt geen onderscheid gemaakt tussen stroom en spanning: elk lampje krijgt alles òf elk

lampje krijgt een-derde deel; in de derde antwoordcombinatie worden stroom en spanning verwisseld. Alle lampjes krijgen alle stroom en een-derde deel van de spanning.

Op deze manier construeren we 12 potentiële variabelen voor een SV-factor, 15 voor een CS-factor, 13 voor een LS-factor en 9 voor een VI-factor. Factoranalyse zal moeten duidelijk maken of deze variabelen ook werkelijk deel uitmaken van de veronderstelde structuur van vier probleemvelden.

### 5. Analyse van de gegevens

In deze paragraaf analyseren we de gegevens van een testafname in klas 2 havo-vwo (N=85) en klas 5 havo (N=60). De test moet immers toepasbaar zijn in alle klassen van havo en vwo. Dit betekent dat de op basis van factoranalyse te construeren schalen, zowel in klas 2 als klas 5, voldoende betrouwbaar dienen te zijn. Een factoranalyse met 4 factoren geeft (na varimax-rotatie) de veronderstelde structuur goed weer (eigenwaarden van de factoren  $> 1$ ). In tabel 1 presenteren we eerst een overzicht van het aantal variabelen met een bepaalde factorlading op de vier factoren, die we kunnen interpreteren als een SV-, CS-, LS- en VI-factor.

Uit de tabel is af te lezen dat ca. twee-derde van de variabelen naar behoren laadt op de vooraf veronderstelde factor. Hierbij komt de LS-factor relatief het duidelijkst naar voren in klas 2 en de SV-factor in klas 5. Tussen de geconstrueerde variabelen komen geen correlaties voor groter dan 0,70. Het heeft dus zin elke variabele afzonderlijk te toetsen op bruikbaarheid binnen één van de te vormen vier schalen. Vanwege de wat verschillende resultaten in beide leerjaren verloopt de overgang van de gevonden vier factoren naar de vorming van schalen, niet standaard. Een variabele wordt namelijk pas in een schaal opgenomen als deze voldoet aan één van de volgende drie criteria:

1. de factorlading op de betreffende factor is in beide leerjaren  $\geq 0,30$  en op de andere factoren  $< 0,15$ ;
  2. de factorlading in beide leerjaren samen is  $\geq 0,70$  en in geen van beide leerjaren  $< 0,10$  en op de andere factoren  $< 0,10$ ;
  3. de factorlading is in één van beide leerjaren  $\geq 0,70$  en in het andere leerjaar niet negatief en op de andere factoren  $< 0,10$ .
- De toepassing van criterium 1 is gebruikelijk. Dit zou echter kunnen betekenen dat variabelen die als representant van een

factor naar voren komen in slechts één van beide leerjaren voor de analyse verloren zouden gaan. In feite hebben we met de toevoeging van de criteria 2 en 3 de toelating van variabelen

Tabel 1: Overzicht van aantallen variabelen met een bepaalde factorlading

factorcode (gepland aantal variabelen)	klas 2			klas 5		
	aantal variabelen met factorlading ≥.30	≥.40	≥.50	aantal variabelen met factorlading ≥.30	≥.40	≥.50
SV (12)	8	6	3	7	7	7
CS (15)	9	9	7	8	7	5
LS (13)	10	9	8	7	5	4
VI (9)	6	4	3	6	4	2

tot een bepaalde schaal wat verruimd. Een en ander leidt tot de opname van variabelen in de vier onderscheiden schalen, zoals weergegeven in tabel 2. Zodra de schalen zijn gevormd, kan de betrouwbaarheidscoëfficiënt en de frequentieverdeling per schaal worden vastgesteld. Omdat het aantal variabelen per schaal verschilt, zien we af van een presentatie van de frequentieverdeling. In tabel 3 presenteren we de betrouwbaarheidscoëfficiënten ( $\alpha$ ) en de gemiddelde schaalscores (+ standaarddeviatie) per leerjaar. Een hoge schaalscore betekent dat de leerlingen op de meeste variabelen uit de betreffende schaal, scoren, en dus in onze interpretatie het achterliggende intuïtieve idee toepassen. Een definitieve vaststelling van de validiteit van de schalen vindt hiermee overigens nog niet plaats.

De betrouwbaarheid van de schalen SV, CS en LS is ruim voldoende ( $\alpha > 0.70$ ), die van de VI-schaal laat te wensen over. Laatstgenoemde schaal is overigens gebaseerd op een gering aantal vragen. Soms is het mogelijk de betrouwbaarheidscoëfficiënt in een van de leerjaren in geringe mate te verhogen door een bepaalde variabele uit de schaal te verwijderen. Dit leidt echter in alle gevallen tot een verlaging van de betrouwbaarheid in het andere leerjaar. We zien dan ook af van het verwijderen van vragen. Zonder daar op dit moment nader op in te gaan, is het interessant te constateren dat de gemiddelde schaalscore per

leerjaar nogal uiteen kan lopen en dat niet gesteld kan worden dat leerlingen in leerjaar 5 het op alle schalen beter doen dan in leerjaar 2 (zie bijvoorbeeld de relatief hoge schaalscore op de CS-schaal in leerjaar 5).

Tabel 2: Schaalvorming op grond van de criteria

schaalnaam	schaalcode	het aantal per schaal toegelaten variabelen op grond van			totaal aantal opgenomen variabelen
		crit.1	crit.2	crit.3	
stroomverbruik	SV	3	5	-	8
constante stroom	CS	6	2	-	8
lok./seq. redeneren	LS	4	3	1	8
geen/foutief onderscheid tussen spanning en stroom	VI	5	-	-	5
<b>totalen</b>		<b>18</b>	<b>10</b>	<b>1</b>	<b>29</b>

Tabel 3: Betrouwbaarheidscoëfficiënten ( $\alpha$ ) en gem. schaalscores (+ standaarddeviaties)

schaalcode	$\alpha$ (klas 2)	$\alpha$ (klas 5)	gem. (klas 2)	gem. (klas 5)
SV	0,75	0,83	4,6 (2,1)	1,8 (2,1)
CS	0,81	0,81	3,3 (2,4)	5,0 (2,6)
LS	0,86	0,84	4,0 (2,8)	3,8 (2,4)
VI	0,62	0,58	2,2 (1,3)	1,7 (1,3)

Tot slot vragen we ons af welke samenhang er bestaat tussen de vier schalen onderling en tussen de vier schaalscores ener-

zijds en de echte testscore anderzijds. Let wel: met de echte testscore bedoelen we de score op alle gesloten vragen in de test, en dus niet de score op door ons achteraf geconstrueerde variabelen.

In tabel 4 presenteren we de correlatie tussen de vier geconstrueerde schalen. Uit deze correlatiegegevens blijkt dat de VI-schaal relatief sterk samenhangt met de CS- en de LS-schaal en in deze vorm nog niet kan leiden tot een onafhankelijke verwijzing naar een bepaald onderwijsprogramma.

Tabel 4: Correlatiecoëfficiënten tussen de vier schaalscores

schaalscore	CS	LS	VI
SV	0,12	0,01	-0,05
CS		0,18	0,30
LS			0,43

Met de gegevens in tabel 5 krijgen we toch al enig inzicht in het percentage leerlingen dat verwezen zou worden naar de diverse onderwijsprogramma's op basis van hun schaalscores. Als verwijzingsnorm hanteren we daarbij het criterium dat een leerling meer dan de helft van de maximale schaalscore heeft (dus SV-, CS-, LS-score > 4; VI-score > 2). Op basis van de gegevens uit tabel 5 lijkt een differentiatie naar verschillende onderwijsprogramma's op grond van de schaalscores mogelijk. Een te verwachten nadeel is dat in leerjaar 2 betrekkelijk veel leerlingen (ca. 35%) verwezen worden naar (bijna) alle vervolgp programma's

In klas 5 blijkt 82% van de variantie in de echte testscore te kunnen worden 'verklaard' met behulp van de variantie in de vier schaalscores. De bijdragen van alle vier schalen is hierbij significant (van SV, CS en LS op 1% niveau, van VI op 5% niveau). In klas 2 blijkt 63% van de variantie in de testscore te kunnen worden 'verklaard' met de vier schaalscores. De bijdragen van drie schalen is hierbij significant (van SV, CS en LS op 1% niveau). Deze resultaten zijn verkregen uit een regressie-analyse, waarbij de behaalde testscore wordt opgevat als afhan-

kelijke variabele en de vier schaalscores tegelijkertijd in de analyse worden opgenomen als de 'verklarende' onafhankelijke variabelen. Het verschil in percentage 'verklarde' variantie tussen beide leerjaren betekent dat in klas 2 in vergelijking met klas 5 meer fouten worden gemaakt die door ons niet te interpreteren zijn in termen van de vier genoemde intuïtieve ideeën. We kunnen deze redenering ook omdraaien: als leerlingen in klas 5 een bepaald intuïtief idee hanteren, lijken ze dit consistent te doen dan leerlingen in klas 2. Een nadere bevestiging van dit resultaat in de nabije toekomst is echter noodzakelijk, aangezien het diagnostisch instrument nu nog te veel gebreken vertoont.

Tabel 5: Percentages leerlingen met een verwijzing naar een bepaald remidiërend onderwijsprogramma en met een bepaald totaal aantal verwijzingen.

	verwijzing naar een programma m.b.t.				totaal aantal verwijzingen				
	SV	CS	LS	VI	0	1	2	3	4
leerjr.2	71	40	45	51	4	32	29	24	11
leerjr.5	20	68	43	32	14	32	35	15	4

## 6. Conclusies over de kwaliteit van het diagnostisch instrument

Onze conclusie is dat we de hier besproken derde test kunnen opvatten als een noodzakelijke en waardevolle stap in de ontwikkeling van een diagnostisch instrument. Om verschillende redenen kan de derde test niet worden beschouwd als eindstap in de ontwikkeling. De hierna te ontwikkelen vierde test kan weliswaar worden opgebouwd rond de vier geconstrueerde schalen, maar zal op de volgende punten afwijken van de derde test:

1. de test kan bestaan uit max. 20 vragen om tot een betrouwbare diagnose te komen. Dit heeft het gunstig neveneffect dat de test in één lesuur is te maken en door de leerlingen zelf is te corrigeren;
2. het aantal open vraagonderdelen kan beperkt blijven, omdat de schaalscores zo'n grote samenhang vertonen met de test-score op de gesloten vragen;

3. de VI-schaal moet worden versterkt door toevoeging van één of twee vragen in deze categorie.

Over de onderzoeksresultaten, verzameld met de te ontwikkelen vierde test, zal later worden gerapporteerd. Dan zal de test ook afgenomen dienen te worden onder grotere groepen leerlingen, verspreid over meerdere scholen.

## Noten

1. We realiseren ons dat factoranalyse het best toepasbaar is bij zogenaamde interval-variabelen. Door de codering in enen en nullen is bij ons geen sprake van dit type variabelen. Het wordt echter steeds meer gebruikelijk factoranalyse slechts op te vatten als een indicator voor samenhang van variabelen en de eisen m.b.t. het type variabelen te versoepelen. Als bovendien de factorstructuur achteraf goed te interpreteren valt, kan dit worden gezien als een bevestiging van het terechte gebruik van deze analyse-methode.

## Literatuur

- Kuiper, J., G.Dulfer, P.Licht & G.Thijs (1985) *Students' conceptual problems in the understanding of simple electricity circuits*, intern rapport VU-Amsterdam.
- Licht, P. (1986a) Begrips- en redeneerproblemen in beginnend elektriciteitsonderwijs, *Tijdschrift voor didactiek der  $\beta$ -wetenschappen*, 4, 88-106.
- Licht, P. (1986b) Waartoe leidt ons elektriciteitsonderwijs? *DBK-nieuws*, 25, 29-47.
- Licht, P. & M.Snoek (1986) Elektriciteit in de Onderbouw, *NVON-maandblad* 11, 32-36.
- Licht, P. (1987) *De ontwikkeling van een onderwijsstrategie om begrips- en redeneerproblemen in het elektriciteitsonderwijs te verminderen*, paper voor ORD '87 te Groningen.
- Rhöneck, C. von, D.M.Shipstone, W.Jung, C.Karrqvist, J.J.Dupin, S.Joshua & P.Licht (1988) European Test of Student Understanding Electricity, *International Journal of Science Education* (in druk).
- Thijs, G.D.(1987) Conceptions of Force and Movement. Intuitive Ideas of pupils in Zimbabwe in comparison with findings from other countries. In: J.Novak e.a., *Proceedings of Cornell Conference on Misconceptions and Educational Strategies, 1987*.