

## 5 Toetsing

### 5.1 Inleiding

Toetsing is te zien als het slotakkoord van een lessenserie, al zal tijdens een lessenserie op een aantal momenten ook sprake zijn van *formatieve* toetsing. In een onderwijsleerproces wordt voortdurend gemeten of de leerling al dan niet op het juiste niveau zit. Van deze metingen hangt voor leerlingen veel af: er wordt bepaald aan welk niveau van onderwijs de leerling kan/mag deelnemen, of de leerling een stap mag maken naar een hoger leerjaar, of de leerling mag slagen voor het eindexamen aan het einde van de schoolcarrière enzovoort. Als meetinstrument maken leraren gebruik van toetsen. Voor deze toetsen worden namen gebruikt zoals proefwerk, schriftelijke overhoring, repetitie, eindexamen, schoolexamen, tentamen en praktijktoets. In alle gevallen komt het echter bij deze *summatieve toetsing* op hetzelfde neer: de leerling krijgt een aantal vragen en/of opdrachten voorgelegd die binnen een afgebakende tijd afgerond moeten worden. Het resultaat geeft een aanduiding van het niveau waarop de leerling ten tijde van de toets functioneerde op een bepaald vlak.

Elke natuurkundige weet dat een meting niet betrouwbaar is als hiervoor geen goede meetinstrumenten worden gebruikt. Een toets is ook een meetinstrument, en kan alleen betrouwbare resultaten geven als de vragen in de toets ook daadwerkelijk toetsen wat ze beogen. Ook de samenstelling van de toets uit de beschikbare goede toetsvragen speelt daarbij een rol. Daarbij beperken we ons in dit hoofdstuk tot *theorietoetsen* (en niet bijvoorbeeld practicumtoetsen).

De *centrale vraag* voor dit hoofdstuk is: welke eigenschappen heeft een goede toetsvraag, en hoe geef je met die toetsvragen een theorietoets op een verantwoorde manier vorm?

### 5.2 Toetsvraagconstructie

In deze paragraaf geven we voorbeelden van verschillende vormen van *open* en *gesloten* toetsvragen, samen met de aandachtspunten ofwel *constructieregels* voor het construeren van deze twee vraagsoorten. Het met behulp van dit soort vragen samenstellen van een toets komt in de volgende paragrafen aan de orde.

#### 5.2.1 Open vragen

Open vragen zijn vragen waarop een leerling zelf het antwoord moet formuleren. Een open vraag doet daarmee een beroep op de volgende drie vaardigheden van de leerling:

- Leesvaardigheid: de leerling moet de benodigde informatie en de vraag kunnen lezen en begrijpen.
- Vakbegrip: de leerling moet de achterliggende natuurkunde in de vraag kunnen herkennen, deze op een bepaald niveau beheersen en weten te koppelen aan de, in de vraag geschetste, context of situatie.
- Schrijfvaardigheid: de leerling moet het juiste antwoord kunnen formuleren.

Bij een open vraag is het van groot belang dat de vraag ook daadwerkelijk toetst wat de vraagsteller beoogt. Voor een toets natuurkunde is dit over het algemeen een stuk vakbegrip, wat dan ook de nadruk moet krijgen in de vraagstelling. Een te groot beroep op de lees- en schrijfvaardigheid van de leerling moet zoveel mogelijk worden vermeden.

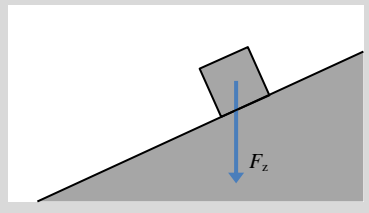
### In- en aanvulvraag

- 1 Twee eenheden die worden gebruikt voor temperatuur zijn ... en ...
- 2 In een gloeilamp wordt ... energie omgezet in ... en ...

Figuur 174 – Voorbeelden van in- of aanvulvragen.

### Kort-antwoordvraag

- 1 Wat is het kookpunt van water, uitgedrukt in Kelvin?
- 2 In de figuur hieronder is een voorwerp op een helling getekend. De zwaartekracht op het voorwerp is getekend. Construeer de normaalkracht die op het voorwerp werkt.



Figuur 175 – Voorbeelden van kort-antwoordvragen.

### In- of aanvulvraag

Bij een *in- of aanvulvraag* krijgt de leerling de opdracht om een deel van een gegeven tekst in of aan te vullen. Twee voorbeelden van in- of aanvulvragen zijn weergegeven in figuur 174.

In- of aanvulvragen worden vaak toegepast bij vragen waarin het beoogde doel is om een stuk (parate) vakkennis bij leerlingen te toetsen. Voor dit doel zijn deze vragen ook zeer geschikt. Leerlingen kunnen zo gedwongen worden om gebruik te maken van de juiste vaktermen. Daarnaast is er snel inzicht te krijgen in het al dan niet kunnen toepassen van deze vaktermen binnen een bepaalde situatie of context.

### Kort-antwoordvraag

Bij een *kort-antwoordvraag* wordt van de leerling een beknopt antwoord verwacht. Dat antwoord kan bijvoorbeeld bestaan uit een enkelvoudige zin, een aantal woorden, een getal, een formule of een eenvoudige tekening. Twee voorbeelden van kort-antwoordvragen zijn weergegeven in figuur 175.

Ook kort-antwoordvragen zijn goed bruikbaar voor de toetsing van (parate) vakkennis bij leerlingen. Door de vragen binnen een context te gebruiken, ontstaan er ook mogelijkheden om dit vraagtype toe te passen in situaties waarin het doel meer ligt bij het toetsen van inzicht in het vak bij leerlingen. Bij kort-antwoordvragen wordt de leerling vaak gevraagd om een zeer concreet antwoord te geven, dat niet al teveel ruimte voor interpretatie biedt aan de leerling. Dit vraagtype doet daarmee geen groot beroep op de schrijfvaardigheid van de leerling. Dat maakt dit vraagtype ook geschikt voor situaties waarin de leerling gevraagd wordt om een erg complexe situatie te analyseren.

Het nakijken van kort-antwoord vragen is veelal eenduidig en verloopt voor de leraar over het algemeen soepel en snel.

### Lang-antwoordvraag

Bij een *lang-antwoordvraag* wordt van de leerling een lang(er) antwoord verwacht. Dat antwoord kan bijvoorbeeld bestaan uit een uitleg, een beredenering, een berekening of bepaling, een afleiding van een formule of een constructie. Twee voorbeelden van lang-antwoordvragen zijn weergegeven in figuur 176.

#### Lang-antwoordvraag

- 1 Een bal wordt in verticale richting omhoog getrapt met een beginsnelheid van 15 m/s. Wrijvingskrachten op de bal mogen verwaarloosd worden. Bereken de maximale hoogte die de bal bereikt.
- 2 Op een optische bank wordt een voorwerp voor een lens geplaatst. Het beeld is zichtbaar op een scherm. De afstand van het voorwerp tot de lens is gelijk aan 2 keer de brandpuntsafstand van de lens. Door het voorwerp verder van de lens af te schuiven, verandert de afmeting van het beeld. Beredeneer of het beeld nu groter of kleiner wordt.



Figuur 176 – Voorbeelden van lang-antwoordvragen.

Lang-antwoordvragen worden veel gebruikt in het vak natuurkunde. Berekeningen, bepalingen en beredeneringen zijn namelijk aan de orde van de dag. Met dit vraagtype is goed te meten of een leerling een bepaalde situatie kan inschatten, er de juiste natuurkundige concepten aan kan koppelen en deze op de juiste wijze weet te combineren.

Het beantwoorden van een lang-antwoordvraag vraagt van de leerlingen wel een grote mate van schrijfvaardigheid, en daarvan moet een leraar zich bewust zijn. Zeker als het gaat om vragen waarbij leerlingen iets moeten uitleggen of een beredenering op papier moeten zetten, kan er een enorme diversiteit aan antwoorden ontstaan. Het beoordelen van de antwoorden op lang-antwoordvragen is daarmee complex en tijdsintensief. Een nauwkeurig geformuleerd correctievoorschrift is daarbij een vereiste.

### Betoogvraag

Kernsplijting is een proces dat gebaseerd is op bepaalde natuurkundige principes. Rond kernsplijting (kernenergie) is nogal wat maatschappelijke discussie. Want deze vorm van energieopwekking heeft een aantal voordelen, maar ook een aantal nadelen.

Schrijf een betoog waarin je persoonlijke visie rondom het gebruik van kernsplijting als energiebron tot uitdrukking komt. Benoem in het betoog minimaal de volgende zaken:

- fysische achtergronden en technische werking van kernsplijting;
- voordelen van het gebruik van kernsplijting als energiebron;
- nadelen van het gebruik van kernsplijting als energiebron;
- afweging van de voordelen en de nadelen ten opzichte van elkaar.

Het betoog moet een omvang hebben van minimaal 1000 en maximaal 1500 woorden.

Figuur 177 – Voorbeeld van een betoog- of opstelvraag.

### Betoog- of opstelvraag

Bij een *betoog- of opstelvraag* wordt van de leerling verlangd dat er een uitgebreid verhaal op papier komt. Dat verhaal bestaat uit een samenhangende tekst met een structuur van inleiding, middendeel en slot. In bepaalde gevallen kan het ook gaan om een zeer gedetailleerde tekening of een berekening. Een voorbeeld van zo'n vraag is weergegeven in figuur 177.

Betoog- of opstelvragen worden slechts zeer zelden binnen het schoolvak natuurkunde toegepast, in tegenstelling tot een aantal andere schoolvakken. De beoordeling van een betoog- of opstelvraag is een lastige klus voor de leraar. Een correctievoorschrift is alleen in hoofdlijnen op te stellen. Het vraagtype vraagt namelijk om enige creativiteit van de leerling, waardoor niet kan worden vastgelegd wat te verwachten antwoorden zijn. Het correctievoorschrift beperkt zich dan tot het noemen van de aspecten die (minimaal) in het betoog of opstel moeten voorkomen.

### 5.2.2 Constructie van open vragen

Voor de constructie van open vragen is het aan te bevelen om de volgende *constructieregels* in acht te nemen.

#### Eenduidige vragen

Om als eerste constructieregel te stellen “dat je moet vragen wat je wilt weten”, lijkt een open deur. Toch zijn veel vragen in toetsen en lesmethoden allerm minst eenduidig in wat ze van de leerling vragen. Zowel voor de leerling als voor de leraar is dit geen wenselijke situatie. De leerling weet namelijk niet wat er van hem of haar verwacht wordt. En de leraar krijgt antwoorden die hij of zij niet voor ogen had bij de constructie van de vraag, maar die wel (deels) goed kunnen zijn.

Een advies dat wat dit betreft nog weleens wordt gegeven, is om eerst het antwoord te formuleren en pas daarna de vraag. Deze werkwijze is echter sterk af te raden. Er bestaat namelijk een kans dat er slechts één beoogd antwoord in ogeschouw genomen wordt, en dat de vraag naar aanleiding van alleen dit antwoord wordt geformuleerd. Het risico is dat de vraag dan zodanig geformuleerd wordt, dat deze niet meer goed leesbaar is. De meest effectieve werkwijze is de vraag formuleren, daarna het antwoord formuleren en dan nogmaals kritisch naar de vraag kijken om te bepalen of een herformulering noodzakelijk is.

Bij een niet-eenduidige vraagformulering is vaak sprake van een zogenaamd *klemtoon-probleem*. In figuur 178 staat daarvan een voorbeeld.

#### Klemtoon leggen

Beschouw de volgende vraag:

Waarom bepaalt het KNMI elk uur de temperatuur op een groot aantal meetpunten?

Deze vraag heeft geen eenduidige formulering. Door de klemtoon op een andere plaats in de vraag te leggen, ontstaat steeds een andere vraag. Voorbeelden hiervan zijn (het deel van de vraag waar de klemtoon ligt, is steeds cursief weergegeven):

- *Waarom* bepaalt het KNMI elk uur de temperatuur op een groot aantal meetpunten?
- Waarom bepaalt *het KNMI* elk uur de temperatuur op een groot aantal meetpunten?
- Waarom bepaalt het KNMI *elk uur* de temperatuur op een groot aantal meetpunten?
- Waarom bepaalt het KNMI elk uur *de temperatuur* op een groot aantal meetpunten?
- Waarom bepaalt het KNMI elk uur de temperatuur *op een groot aantal meetpunten*?

Op deze vraag is voor leerlingen geen eenduidig antwoord te geven, omdat voor hen niet duidelijk is welke van deze vragen wordt bedoeld.

Dit probleem is op te lossen door een deel van de informatie aan de leerlingen weg te geven. De onduidelijkheid over wat de informatie en wat de vraag is, wordt daarmee weggenomen. Een alternatieve formulering van de vraag hierboven zou kunnen zijn:

Het KNMI bepaalt elk uur de temperatuur op een groot aantal meetpunten.

- Geef een reden waarom dat elk uur gebeurt.
- Geef een reden waarom dat op een groot aantal meetpunten gebeurt.

Door de informatie en de vraag van elkaar te scheiden, ontstaat duidelijkheid in de vraagstelling en kan van de leerling een eenduidig antwoord worden verwacht.

Figuur 178 – Voorbeeld van een vraag met een klemtoon-probleem, en de oplossing voor dat probleem. Bron: *CITO – Toetsen op school*.

Het multi-interpretabel zijn van een vraag is een van de grootste uitdagingen in de constructie van een goede toetsvraag. Het is aan te bevelen om de ‘klemtooncontrole’ regelmatig toe te passen.

### Antwoordrestricties

Het komt regelmatig voor dat leerlingen in het antwoord op een vraag een aantal zaken moeten opsommen. Hierbij is het van belang om een concrete opdracht te geven. Als het de bedoeling is dat de leerling bijvoorbeeld drie redenen geeft, dan moet dat in de vraagformulering worden opgenomen. Met formuleringen als “enkele” of “een aantal” redenen is een vraag niet concreet genoeg.

### Onderscheid tussen informatie en vraag

In een toets natuurkunde is een veelheid aan vragen opgenomen voor de leerling. Uniek is echter dat elke vraag gesteld wordt vanuit een bepaalde situatie en/of context. Want binnen de natuurkunde proberen we de wereld om ons heen te begrijpen, en dat komt ook in de toetsing van het vak tot uitdrukking. In de praktijk betekent dit dat er in een toets natuurkunde behoorlijk wat informatie is opgenomen. Informatie die bedoeld is om leerlingen een goed beeld te geven van de situatie en/of context.

Bij de constructie van een vraag is het zeer aan te bevelen om de informatie los te koppelen van de vraag. Op die manier is het voor de leerling duidelijk welke tekst informatief bedoeld is (de stam) en welke tekst er gebruikt is om vragen te stellen (de vraag). In figuur 179 is een voorbeeld weergegeven van het loskoppelen van informatie en vraag.

### Informatie en vraag loskoppelen

Beschouw de volgende vraag:

Een hoeveelheid water wordt verwarmd van een temperatuur van 20 °C naar een temperatuur van 80 °C.

Leg uit dat er een kleinere hoeveelheid warmte nodig is voor het verwarmen van een kleinere hoeveelheid water bij dezelfde begin- en eindtemperatuur van het water.

In deze vraag wordt een bepaalde situatie geschetst (een hoeveelheid water wordt verwarmd van een gegeven begin- naar een eindtemperatuur). Vervolgens wordt de situatie enigszins gewijzigd (de hoeveelheid water wordt verkleind) en wordt gevraagd hoe een andere grootte (de hoeveelheid warmte) mee varieert. In de vraag (vanaf “Leg uit...”) staat nogal wat informatie opgenomen, waardoor deze een aanzienlijk beroep doet op de leesvaardigheid van de leerling. In een betere formulering worden de informatie en de vraag van elkaar losgekoppeld:

In een experiment wordt een hoeveelheid water verwarmd van een temperatuur van 20 °C naar een temperatuur van 80 °C. Hiervoor is een bepaalde hoeveelheid warmte nodig. Het experiment wordt herhaald met een kleinere hoeveelheid water. Er is nu minder warmte nodig om de stof te verwarmen van een temperatuur van 20 °C naar een temperatuur van 80 °C.

Leg dit uit.

Figuur 179 – Voorbeeld van het loskoppelen van informatie en vraag.

In het voorbeeld van figuur 178 is een vergelijkbare aanpak gehanteerd. Een bijkomend voordeel van het duidelijk onderscheiden van de informatie en de vraag is dat een leerling bij het vervolg in de toets geen informatie hoeft te zoeken in de vragen, maar dat hij of zij zich kan beperken tot de informatieteksten.

### Concrete vragen

Op vaag gestelde vragen zijn vaag geformuleerde antwoorden van leerlingen te verwachten. Bij de constructie van een vraag is het dan ook van belang om zo

### Concrete vragen stellen

Beschouw de volgende vraag:

Kun je aangeven welk verband er bestaat tussen het volume waarin een gas is opgesloten en de gasdruk?

Iedereen zal onderkennen dat een antwoord als “Ja” (of “Nee”) het juiste antwoord op deze vraag kan zijn. Het spreekt voor zich dat dit niet het beoogde antwoord is. Wél concrete formuleringen van deze vraag kunnen zijn:

- Welk verband bestaat er tussen het volume waarin een gas is opgesloten en de gasdruk?
- Geef aan welk verband er bestaat tussen het volume waarin een gas is opgesloten en de gasdruk.

Op deze vragen bestaat wel een eenduidig antwoord.

Figuur 180 – Voorbeeld van een niet concreet gestelde vraag en de mogelijke herformuleringen.

concreet mogelijk te zijn. Een onvoldoende concreet geformuleerde vraag wordt over het algemeen snel herkend, maar toch komen ze veel voor. In figuur 180 is een voorbeeld weergegeven van een niet concreet gestelde vraag en een alternatief hiervoor.

De noodzaak van het stellen van concrete vragen lijkt overduidelijk. Toch is er, vooral in lesmethodes, vaak sprake van niet concreet geformuleerde vragen zoals “Kun je aangeven/uitleggen/vertellen/...”. Een reden hiervoor zou kunnen zijn dat niet concreet geformuleerde vragen bij de leerling vriendelijker overkomen dan een wél concreet geformuleerde vraag. Toch heeft de concretisering van vragen nadrukkelijk de voorkeur.

Het is voor de leerling dus van belang dat een vraag concreet is. Dit uit zich ook in het gebruik van een aantal *standaardformuleringen* voor vragen. Door steeds dezelfde formulering te gebruiken, ontstaat er voor de leerling veel duidelijkheid en kan hij of zij snel inschatten wat er bij een bepaalde vraag van hem of haar wordt verwacht. In figuur 181 en 182 staat een aantal standaardformuleringen voor kort- en lang-antwoordvragen.

Formulering	Toelichting
Noem	Vraag naar iets concreets (plaats, kenmerk, enzovoort).
Geef (aan)	Vraag naar concrete zaken.
Hoe	Kwantitatieve vragen en vragen naar concrete zaken.
Wie	Identificatievraag.
Wat	Vraag naar een definitie.
Waar	Vraag naar een plaatsaanduiding.
Welke	Vraag naar een object of categorie.
Wanneer	Vraag naar een moment of omstandigheden.
Arceer	Het antwoord moet grafisch worden weergegeven.
Citeer	Het antwoord is een citaat uit een gegeven informatiebron.

Figuur 181 – Standaardformuleringen voor kort-antwoordvragen.

Formulering	Toelichting
Geef (aan)	Vraag naar iets abstracts, zoals een reden of een oorzaak.
Bereken	Niet alleen de uitkomst wordt beoordeeld, maar ook de uitgevoerde berekening.
Bepaal	Niet alleen de uitkomst wordt beoordeeld, maar ook de uitgevoerde verwerking van een informatiebron (bijvoorbeeld een grafiek) en de uitgevoerde berekening.
Laat met een berekening zien	De uitkomst van de berekening staat al in de vraag of in de stam van de vraag aangegeven.
Stel op	Opstellen van vergelijkingen.
Leid af	Op basis van verstrekte gegevens moet het gevraagde worden afgeleid.
Onderzoek	Onderzoeken van een serie eigenschappen of beweringen.
Bewijs	Op deductieve wijze aantonen dat een gestelde bewering correct is.
Toon aan	Aannemelijk maken van een gegeven antwoord.
Leg uit (waarom)	Een redenering of argumentatie bestaande uit enkele afzonderlijke denkstappen.
	Daarnaast bestaan er vergelijkbare formuleringen met een vergelijkbare inhoud als de bovenstaande: beschrijf, beredeneer, beargumenteer, beoordeel, vat samen, teken, construeer, maak een schatting.

Figuur 182 – Standaardformuleringen voor lang-antwoordvragen.

Het grote voordeel van het consequent hanteren van dezelfde formuleringen is duidelijkheid voor de leerling in wat er van hem of haar als antwoord wordt verwacht. In de syllabi voor havo en vwo zijn deze formuleringen ook opgenomen, en worden daar de *examenwerkwoorden* genoemd.

### Heldere informatie

In het voorgaande is vooral het belang van een helder geformuleerde vraag aan de orde geweest. Het spreekt echter voor zich dat ook de informatie rondom de

vraag voor de leerling helder moet zijn. Het is daarbij van belang dat de informatie goed begrepen kan worden door een leerling met het juiste niveau en de juiste voorbereiding. De informatie dient de situatie of de context te beschrijven en mag niet (teveel) afleiden van de essentie.

Een beknopte tekst die de essentie goed aangeeft, is vanzelfsprekend aan te bevelen. Het gebruik van visuele informatie – zoals figuren, tekeningen, foto's of grafieken – is aan te raden als deze de situatie of context duidelijker weten te maken. Ook in een toets zegt een figuur vaak meer dan woorden. In figuur 183 wordt een voorbeeld uitgewerkt waarin het presenteren van de informatie goed tot uitdrukking komt.

### Visualiseren

Beschouw de tekst van het eerste deel van een eindexamenopgave:

Hoewel er van dinosauriërs vrij veel bekend is, weten we van de meeste dinosauriërs weinig over het geluid dat ze maakten. Een uitzondering hierop is de Parasaurolophus. Deze dinosaurus was in het bezit van een grote hoorn boven op de schedel. Deze hoorn diende als klankkast om het geluid te versterken.

Bij een volwassen mannetje is de hoorn 1,8 m lang. Eén uiteinde van deze hoorn is open, het andere uiteinde is gesloten. De luchttemperatuur in de hoorn bedraagt 20 °C. Bereken de frequentie van de grondtoon.

De hoeveelheid tekst in deze vraag is aanzienlijk. Slechts een zeer klein deel van de tekst is echt van belang om de vraag te kunnen beantwoorden. Het restant van de tekst is bedoeld om een context te introduceren. Het herkennen van de natuurkunde binnen een context is een essentiële vaardigheid, waarop de leerling ook getoetst moet worden. Het inperken van de tekst is dus geen optie: de context zou daarmee komen te vervallen. De examenmakers hebben destijds gekozen voor de volgende aanvulling:



Hoewel er van dinosauriërs vrij veel bekend is, weten we van de meeste dinosauriërs weinig over het geluid dat ze maakten. Een uitzondering hierop is de Parasaurolophus, zie nevenstaande figuur. Deze dinosaurus was in het bezit van een grote hoorn boven op de schedel. Deze hoorn diende als klankkast om het geluid te versterken.

Bij een volwassen mannetje is de hoorn 1,8 m lang. Eén uiteinde van deze hoorn is open, het andere uiteinde is gesloten. De luchttemperatuur in de hoorn bedraagt 20 °C. Bereken de frequentie van de grondtoon.

Door een figuur toe te voegen, wordt de context ineens veel duidelijker, en kan de leerling zich een voorstelling maken van de situatie die aan de orde is.

Figuur 183 – Voorbeeld van heldere (visuele) informatie. Bron: examen havo 2012, tweede tijdvak.

Het toepassen van extra informatiebronnen is niet altijd even vanzelfsprekend, maar wel altijd het overwegen waard. In de eerste versie van de vraag uit figuur 183 (zonder figuur) wordt een groot beroep gedaan op de leesvaardigheid van de leerling, terwijl dit in de tweede versie (met figuur) minder het geval is door een aanvullende bron van informatie.

### Hoeveelheid tekst

Een toets natuurkunde is bedoeld om binnen het vakgebied conceptueel begrip te toetsen, en niet de leesvaardigheid van de leerling. Dan is het van belang om de hoeveelheid tekst te beperken. Het is de uitdaging om enerzijds de situaties, contexten en vragen helder beschreven te krijgen, maar anderzijds de leerling ook niet te overvragen in de te lezen hoeveelheid tekst. In figuur 184 wordt een derde versie van de opgave uit figuur 183 gepresenteerd, met als doel de hoeveelheid tekst in te perken.

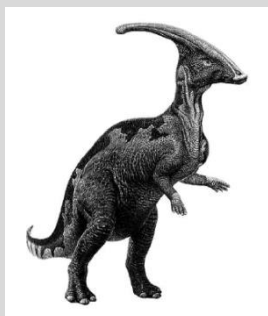
De daadwerkelijke hoeveelheid tekst is in figuur 184 uiteraard niet veranderd. De hoeveelheid tekst zal door een leerling echter wel ervaren worden als meer beperkt, vanwege de tussenvraag en de focus die daarmee bij de leerling gelegd wordt op het beantwoorden van een vraag en de natuurkunde die daarbij aan de orde is.

Een vanzelfsprekende tweede methode om de hoeveelheid tekst te beperken, is het daadwerkelijk schrappen van zinnen, zinsdelen of woorden. Daarbij is het

wel belangrijk om na te gaan of de essentie van de vraag en een helder beeld van de situatie of context niet komen te vervallen.

### Tekst inperken

Beschouw de tweede versie van de opgave uit figuur 183. Het eerste deel van de opgave is bedoeld om inzicht te krijgen in de context, maar het doet wel een beroep op de leesvaardigheid van de leerlingen. Een oplossing hiervoor is het volgende:



Hoewel er van dinosauriërs vrij veel bekend is, weten we van de meeste dinosauriërs weinig over het geluid dat ze maakten. Een uitzondering hierop is de Parasaurolophus, zie nevenstaande figuur. Deze dinosaurus was in het bezit van een grote hoorn boven op de schedel. Deze hoorn diende als klankkast om het geluid te versterken.

**1** Van welk natuurkundig verschijnsel is hier sprake?

Bij een volwassen mannetje is de hoorn 1,8 m lang. Eén uiteinde van deze hoorn is open, het andere uiteinde is gesloten. De luchttemperatuur in de hoorn bedraagt 20 °C.

**2** Bereken de frequentie van de grondtoon.

Door aan de eerste paragraaf (daar waar de context wordt geïntroduceerd) een (eenvoudige) vraag te koppelen, wordt de leerling gedwongen na te denken over de context en de koppeling te leggen met de relevante natuurkunde. De leerling zal hierdoor minder tekst ervaren en wordt sneller naar de essentie geleid.

Figuur 184 – Voorbeeld van een manier om de ervaren hoeveelheid tekst in een opgave te beperken.

### Koppeling tussen vragen

In een toets is het van belang dat elke vraag op zichzelf staat. Het is voor een leerling erg vervelend als hij of zij een bepaalde vraag niet weet te beantwoorden en daardoor ook in de problemen raakt bij vervolgvragen. De koppeling tussen vragen moet zoveel mogelijk worden vermeden. In figuur 185 staat daarvan een voorbeeld.

### Gekoppelde vragen vermijden

In dit voorbeeld wordt verder gegaan op de tweede vraag van de examenopgave uit figuur 183 en 184. Het vervolg van de opgave luidt:

Bij een volwassen mannetje is de hoorn 1,8 m lang. Eén uiteinde van deze hoorn is open, het andere uiteinde is gesloten. De luchttemperatuur in de hoorn bedraagt 20 °C.

**2** Bereken de frequentie van de grondtoon.

Onderzoek heeft uitgewezen dat een mannelijke Parasaurolophus een toon kon produceren met een frequentie van 240 Hz. Dit is een boventoon van de grondtoon uit de vorige vraag.

**3** Beredeneer of dit de eerste, de tweede, de derde, de vierde of de vijfde boventoon is.

Tussen deze twee vragen bestaat een duidelijke koppeling. Het antwoord op vraag 2 is nodig om het antwoord op vraag 3 te kunnen berekenen. Deze situatie is niet wenselijk. Als een leerling niet in staat is geweest het antwoord op vraag 2 te vinden, loopt het geheel hier voor hem of haar spaak. Er bestaat echter de mogelijkheid dat de leerling op vraag 3 het antwoord wel weet, maar het aantonen hiervan wordt hem of haar bemoeilijkt. Dit is op de volgende manier op te lossen:

Bij een volwassen mannetje is de hoorn 1,8 m lang. Eén uiteinde van deze hoorn is open, het andere uiteinde is gesloten. De luchttemperatuur in de hoorn bedraagt 20 °C. De frequentie van de grondtoon bedraagt 48 Hz.

**2** Toon dat aan.

Onderzoek heeft uitgewezen dat een mannelijke Parasaurolophus een toon kon produceren met een frequentie van 240 Hz. Dit is een boventoon van de grondtoon van 48 Hz.

**3** Beredeneer of dit de eerste, de tweede, de derde, de vierde of de vijfde boventoon is.

Door de uitkomst van vraag 2 weg te geven, kan de leerling bij vraag 3 probleemloos verder, en daar opnieuw aantonen over de juiste kennis en vaardigheden te beschikken.

Figuur 185 – Voorbeeld van een manier om de koppeling tussen vragen te vermijden.

In het voorbeeld van figuur 185 is ervoor gekozen om het antwoord op de vraag weg te geven en de leerling te laten aantonen dat dit antwoord juist is. Er zou de

indruk kunnen bestaan dat de vraag hiermee eenvoudiger is geworden. Dit is uitdrukkelijk niet het geval. Mocht de leerling aan de hand van het gegeven antwoord de vraag gaan oplossen, dan zal hij of zij toch alle benodigde denkstappen moeten doorlopen – weliswaar in een andere volgorde dan wanneer het antwoord niet gegeven is. Het doel van de vraag blijft de toetsing van natuurkundig inzicht, en daarin vindt geen wijziging plaats.

Het geven van het antwoord en dit de leerling laten aantonen, kan het best op een zo expliciet mogelijke manier gebeuren. De leerling zelf een antwoord laten kiezen, is voor de meeste leerlingen teveel gevraagd en dus niet aan te raden. De opmerking “Als je bij de vorige vraag geen antwoord hebt gevonden, reken dan verder met...” scheidt bij de leerling vaak verwarring. Het blijft dan namelijk de vraag of het gegeven antwoord juist is, in de buurt van het juiste antwoord ligt of ver van het juiste antwoord verwijderd ligt. Door de uitkomst te laten aantonen, wordt al deze twijfel vermeden.

### 5.2.3 Gesloten vragen

Gesloten vragen zijn vragen waarop het antwoord gegeven is in de vorm van een aantal alternatieven, waarbij de leerling het juiste (of beste) alternatief moet aangeven. Dat doet een beroep op de volgende twee vaardigheden van de leerling:

- Leesvaardigheid: de leerling moet de benodigde informatie, de vraag en de antwoordalternatieven kunnen lezen en begrijpen.
- Vakbegrip: de leerling moet de achterliggende natuurkunde in de vraag kunnen herkennen, deze op een bepaald niveau beheersen en weten te koppelen aan de, in de vraag geschetste, context of situatie en de gegeven antwoordalternatieven.

In vergelijking met een open vraag is schrijfvaardigheid voor de leerling in een gesloten vraag niet vereist. Net als bij een open vraag is het bij de constructie van een gesloten vraag van belang dat de vraag ook daadwerkelijk toetst wat de vraagsteller beoogt. Voor een toets natuurkunde is dit over het algemeen een stuk vakbegrip, wat dan ook de nadruk moet krijgen in de vraagstelling. Een te groot beroep op de leesvaardigheid van de leerlingen moet zoveel mogelijk worden vermeden.

#### Meerkeuzevraag

De binnen het vak natuurkunde meest gebruikte gesloten vraag is de *meerkeuzevraag* (of *multiple-choice-vraag*). In dit type vraag wordt voor de leerlingen een vraag gekoppeld aan een geschetste situatie. Vervolgens wordt er een aantal mogelijke antwoorden gepresenteerd, waaruit de leerling het juiste antwoord moet kiezen. In figuur 186 staat een voorbeeld van zo'n meerkeuzevraag.

Dit vraagtype is goed bruikbaar om (snel) te meten of een leerling al dan niet het juiste conceptuele begrip heeft. Een voordeel is dat de focus van de leerling kan liggen op het inschatten van de situatie en de natuurkundige inhoud die in de vraag wordt getoetst. Het formuleren van een antwoord is niet aan de orde. Er wordt dan ook geen beroep gedaan op de schrijfvaardigheid van de leerling. Daardoor is er voor de leraar geen sprake van eventuele onduidelijkheden in het antwoord van de leerling. Het niet hoeven opschrijven van een berekening, bepaling of beredenering heeft ook nadelen. Een antwoord op een meerkeuzevraag is namelijk goed of fout, terwijl in de beredenering van de leerling wel degelijk juiste antwoordelementen kunnen voorkomen.

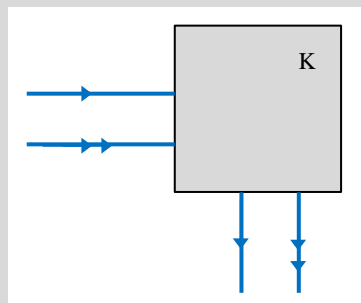
De beoordeling van een toets met meerkeuzevragen vraagt van de leraar vanzelfsprekend veel minder tijd dan de beoordeling van een toets met open vragen. In bepaalde situaties kan dit wenselijk zijn. Hierbij moet wel rekening worden gehouden met het feit dat de constructie van een gesloten vraag veel tijd vraagt. Naast de beschrijving van de situatie of de context en de formulering van de vraag, moeten nu namelijk ook de mogelijke antwoorden worden geformuleerd. De formulering van het juiste antwoord is daarbij meestal niet het probleem, maar het formuleren van de andere antwoordalternatieven kan arbeidsintensief zijn. Het is aan de leraar om de afweging te maken of een toets met gesloten vragen wenselijk is.

#### Meerkeuzevraag

Onder een ondoorzichtig stuk karton K kan zich een van de volgende voorwerpen bevinden:

- een vlakke spiegel;
- een rechthoekig prisma;
- een lens.

De loop van de lichtstralen wordt door het voorwerp veranderd, zoals weergegeven in onderstaande figuur.



Welk voorwerp kan er onder het stuk karton zitten?

- A Alleen een vlakke spiegel.
- B Alleen een rechthoekig prisma.
- C Een vlakke spiegel of een rechthoekig prisma.
- D Een vlakke spiegel of een lens.

Figuur 186 – Voorbeeld van een meerkeuzevraag.



### Stellingvraag

Bestudeer de volgende twee stellingen:

**1** Als twee vingers van dezelfde hand in contact komen met twee draden, waar een spanning tussen staat van 230 V, krijg je een schok. Deze situatie is veel gevaarlijker dan wanneer er contact is met een vinger van elke hand.

**2** Een spanning van 230 V wil zeggen dat aan een elektron in een gesloten stroomkring een energie wordt gegeven van 230 J.

Welke van deze twee stellingen is juist?

- A Alleen 1.
- B Alleen 2.
- C Zowel 1 als 2.
- D Geen van beiden.

Figuur 187 – Voorbeeld van een stellingvraag.

### Combinatievraag

In onderstaand overzicht staat een aantal eenheden.

- A ampère
- B kelvin
- C kilogram
- D meter
- E mol
- F seconde

Vul de onderstaande zinnen aan met de gegeven eenheden (A t/m F).

- 1** De standaardeenheid voor temperatuur is ...
- 2** De standaardeenheid voor tijd is ...
- 3** De standaardeenheid voor stroomsterkte is ...
- 4** De standaardeenheid voor hoeveelheid stof is ...

Figuur 189 – Voorbeeld van een combinatievraag.

### Stellingvraag

In een *stellingvraag* worden twee stellingen voorgelegd, en soms zelfs meer dan twee. De leerling moet aangeven welke van de stellingen juist is/zijn. De antwoordalternatieven zijn bij elke vraag hetzelfde. In figuur 187 staat een voorbeeld van zo'n stellingvraag.

Het gebruik van dit vraagtype is af te raden. Er wordt een groot beroep gedaan op de leesvaardigheid van de leerling, en daarmee is het uiterst twijfelachtig of dit type gesloten vraag wel in voldoende mate het vakinhoudelijke begrip van de leerling toetst. De formulering van de stellingen moet uiterst zorgvuldig gebeuren, en een kleine onvolkomenheid of onzorgvuldigheid kan al snel leiden tot onduidelijkheid in de stelling (en dus in de vraag).

### Herordeningsvraag

In een *herordeningsvraag* wordt een aantal onderdelen gepresenteerd die een bepaalde relatie met elkaar hebben. Het is aan de leerling om die onderdelen in de gevraagde volgorde te herordenen. Een voorbeeld van een herordeningsvraag staat in figuur 188.

### Herordeningsvraag

In het verleden zijn er verschillende modellen voor een atoom in omloop geweest. Deze modellen zijn:

- 1** Atoommodel van Thomson
- 2** Atoommodel van Rutherford
- 3** Atoommodel van Schrödinger
- 4** Atoommodel van Dalton
- 5** Atoommodel van Bohr

Hoe heeft de ontwikkeling van het atoommodel plaatsgevonden in chronologische volgorde?

- A 1 – 4 – 5 – 2 – 3
- B 4 – 1 – 2 – 5 – 3
- C 4 – 1 – 2 – 3 – 5
- D 1 – 4 – 3 – 2 – 5

Figuur 188 – Voorbeeld van een herordeningsvraag.

Een herordeningsvraag wordt door veel leerlingen ervaren als een puzzel, en daardoor niet door alle leerlingen gewaardeerd. Het formuleren van antwoordalternatieven bij dit vraagtype is een behoorlijke uitdaging. Het doel is namelijk dat alle alternatieven de leerling voldoende moeten 'uitnodigen' om ervoor te kiezen. Opvallend onjuiste antwoordalternatieven moeten dan zoveel mogelijk worden vermeden. De herordeningsvraag kent in de natuurkunde niet heel veel toepassingen, en wordt dan ook niet vaak gebruikt.

### Combinatievraag

In een *combinatievraag* is het voor de leerling de uitdaging om overeenkomstige uitspraken en/of gegevens met elkaar te combineren. Dit vraagtype wordt ook wel een *matching-vraag* genoemd. In figuur 189 staat een voorbeeld van een combinatievraag over grootheden en eenheden.

Een combinatievraag wordt, over het algemeen, door een leerling als niet al te moeilijk ervaren. Het vraagtype toetst voornamelijk vakkennis, en daarop kan een leerling zich goed voorbereiden. De uitdaging voor de leraar is om bij een combinatievraag te komen tot een valide normering. Door de keuze voor een bepaald alternatief beperkt de leerling de keuzemogelijkheden, waardoor de gokkans hoger is naarmate de leerling vordert in de vraag. Er moet dus goed worden nagedacht over de manier van puntentoekenning bij zo'n vraag. De combinatievraag kent in de natuurkunde niet heel veel toepassingen, en wordt dan ook niet vaak gebruikt.

### Juist/onjuist-vraag

Eerder in deze paragraaf is de stellingvraag aan de orde geweest, en is aangegeven dat deze nogal een beroep doet op de leesvaardigheid van de leerling. Door

### Juist/onjuist-vragen

Geef van de onderstaande stellingen aan of deze juist of onjuist zijn.

- 1 De standaardeenheid voor snelheid is de kilometer per uur (km/uur of km/h).
- 2 Als het volume waarin een gas is opgesloten verkleind wordt en de gasdruk blijft gelijk, zal de temperatuur van het gas stijgen.
- 3 Een voorwerp wordt in verticale richting omhoog geschoten. In het hoogste punt dat het voorwerp bereikt, werken er op het voorwerp geen krachten.

Figuur 190 – Voorbeeld van juist/onjuist-vragen.

### Negatieve vraagstelling

In de volgende vraag is een negatieve vraagstelling moeilijk te vermijden:

Welke van onderstaande eenheden is geen standaardeenheid?

- A ampère
- B kelvin
- C seconde
- D volt

Als een leerling over het woord “geen” heen leest (of hier wellicht “een” leest), ontstaat een compleet andere vraag. De oplossing is echter eenvoudig:

Welke van onderstaande eenheden is **geen** standaardeenheid?

- A ampère
- B kelvin
- C seconde
- D volt

Door het essentiële woord in de vraag vet te drukken (of te onderstrepen of in hoofdletters te schrijven), ontstaat er voor de leerling duidelijkheid in wat er van hem of haar wordt verwacht.

Figuur 191 – Een alternatief voor een negatieve vraagstelling in een gesloten vraag.

een combinatie van stellingen aan de leerling voor te leggen, ontstaan er combinaties die door de leerling niet altijd als even logisch worden ervaren. De stellingen kunnen echter ook afzonderlijk aan de leerling worden voorgelegd, waarbij de leerling moet aangeven of de stelling al dan niet juist is. Op dat moment is een *juist/onjuist-vraag* (of *waar/onwaar-vraag*) ontstaan. In figuur 190 staat een aantal voorbeelden van juist/onjuist-vragen.

Het gebruik van juist/onjuist-vragen kan een leraar (en een leerling zelf) snel inzicht verschaffen in het al dan niet beheersen van de leerstof. De uitdaging daarbij is de formulering van de stelling. Het spreekt voor zich dat deze door de leerling slechts op één manier geïnterpreteerd mag worden. Dit vraagtype doet dan ook een beroep op de leesvaardigheid van de leerling. Toch worden juist/onjuist-vragen in de natuurkunde regelmatig toegepast om kennishiaten bij leerlingen snel te kunnen opsporen.

## 5.2.4 Constructie van gesloten vragen

Voor de constructie van gesloten vragen is het aan te bevelen om de volgende *constructieregels* in acht te nemen.

### Informatie en vraag

In paragraaf 5.2.2 is al veel gezegd over de constructie van een goede open vraag. Net als bij een open vraag is er bij een gesloten vraag sprake van een hoeveelheid informatie: de beschrijving van een situatie of context. De aanpak hierbij is voor beide soorten vragen sterk vergelijkbaar. En net als een open vraag moet een gesloten vraag eenduidig en concreet zijn. Een belangrijk verschil is wel dat er in een gesloten vraag meestal alleen sprake is van een vraag, terwijl er in een open vraag ook vaak een opdracht voorkomt zoals bereken, bepaal of leid af. In deze paragraaf geven we enkele, speciaal voor de constructie van gesloten vragen relevante aanvullingen op de constructieregels voor open vragen.

Bij een gesloten vraag is het van groot belang, nog sterker dan bij een open vraag, dat er sprake is van een directe en korte vraag. Daarna volgen de antwoordalternatieven. Via een directe en korte vraag is de relatie tussen de vraag en de antwoordalternatieven veel eenvoudiger te handhaven, en daarmee ontstaat duidelijkheid voor de leerling.

Bij een gesloten vraag wordt een groot beroep gedaan op de leesvaardigheid van de leerling. Het is dan ook van belang om zowel de informatie als de vraag beknopt te formuleren. Daarbij is het zeer aan te bevelen om de vraag positief te formuleren. Daar waar een negatieve formulering gewenst is, is het nodig om dit voor de leerling te expliciteren. In figuur 191 staat daarvan een voorbeeld.

### Antwoordalternatieven

Het bijzondere van gesloten vragen is dat het juiste antwoord beschikbaar is voor de leerling. Het is aan de leerling om dit juiste antwoord te selecteren uit een aantal aangeboden alternatieven. In de formulering van de alternatieven is het aan te bevelen om eerst het juiste antwoord op te stellen. Naar aanleiding van dit antwoord kunnen dan de alternatieven worden geformuleerd. Het juiste antwoord wordt ook wel de *antwoordsleutel* genoemd, en de andere alternatieven vormen de zogenaamde *afleiders*. Het optimale aantal afleiders is vastgesteld op drie (of eventueel twee). In combinatie met de antwoordsleutel ontstaat er dan een vierkeuzevraag (of driekeuzevraag). De gokkans is met dit aantal alternatieven acceptabel, en het beroep op de leesvaardigheid van de leerling is nog redelijk beperkt. De uitdaging is vooral om alle antwoordalternatieven ‘aannemelijk’ te maken voor de leerling, en dat is niet eenvoudig. Als er sprake is van een vierkeuzevraag waarbij een van de antwoordalternatieven nooit door een leerling gekozen zal worden, is er eigenlijk sprake van een driekeuzevraag – en dat heeft invloed op de gokkans.

Voor het formuleren van de antwoordalternatieven in een gesloten vraag moeten de volgende constructieregels in acht worden genomen.

- Er bestaat een vaste *volgorde* waarin de alternatieven worden gepresenteerd. De standaardvolgorde daarbij is *alfabetisch*, maar in bepaalde gevallen kan een

chronologische volgorde of een volgorde ‘van klein naar groot’ beter zijn.

- Het *uiterlijk* van de alternatieven moet ruwweg hetzelfde zijn. Dit geldt bijvoorbeeld voor de tekstlengte en de grammaticale structuur per alternatief. Een sterk afwijkend uiterlijk van een van de alternatieven kan de leerling in verwarring brengen, en hem of haar doen besluiten om dit alternatief als zeker juist (of onjuist) antwoord aan te merken.
- Als er *subtiële (maar essentiële) verschillen* bestaan tussen de afleiders is het aan te bevelen om deze voor de leerling duidelijk te maken, bijvoorbeeld door bepaalde woorden in de alternatieven vet te drukken of te onderstrepen. Op die manier wordt het beroep op de leesvaardigheid van de leerling enigszins beperkt.
- De *verschillen* tussen de afleiders mogen niet te groot zijn. Als een van de alternatieven een wezenlijk ander antwoord is, bestaat de kans dat de leerling dit zal aanmerken als zeker onjuist (of juist) antwoord, waarmee het aantal alternatieven impliciet verkleind wordt (of er juist geen alternatieven aan de orde zijn voor de leerling).
- Bij een gesloten vraag is het de bedoeling dat slechts één antwoord het juiste is. De alternatieven moeten elkaar dan ook *uitsluiten*.
- Het gebruik van *absolute woorden* in een gesloten vraag brengt een risico met zich mee. Absolute woorden zijn woorden zoals “altijd”, “nooit” en “zeker”. Bij het gebruik van deze woorden moet er bij de constructie absolute zekerheid bestaan over het juiste antwoord en het onjuist zijn van alle alternatieven. Vooral in gesloten vragen met stellingen (bijvoorbeeld in een juist/onjuist-vraag) is het gebruik van absolute woorden een risico. Er kan een situatie ontstaan waarbij het beoogde juiste antwoord toch niet als geheel juist bestempeld kan worden en/of een afleider toch (gedeeltelijk) juist kan zijn.
- Als het formuleren van voldoende afleiders een grote uitdaging blijkt te zijn, wil men als laatste alternatief nog weleens kiezen voor een alternatief in de vorm van “Alle bovengenoemde antwoorden zijn juist” of “Geen van bovengenoemde antwoorden is juist”. Over het algemeen is deze aanvulling af te raden, zeker als het juiste antwoord al benoemd is en dus door de leerling gekozen kan worden.
- Ook bij de alternatieven is het streven om de *tekst per alternatief* niet te lang te maken. Gesloten vragen doen al een flink beroep op de leesvaardigheid van de leerling, en te lang geformuleerde alternatieven bevorderen de kwaliteit van de vraag dan niet.

In figuur 192 staan enkele meerkeuzevragen waarbij de bovenstaande constructieregels in acht zijn genomen.

### Constructieregels toepassen

Een metalen draad (draad 1) is aangesloten op een spanningsbron. Hierdoor loopt er een bepaalde stroomsterkte door de draad. De draad wordt vervangen door een draad met een **2 maal zo grote diameter** (draad 2).

Wat geldt er voor de stroomsterkte door draad 2?

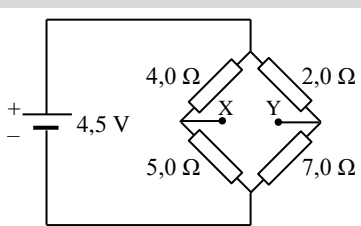
- A De stroomsterkte is **4** maal zo **groot** als die bij draad 1.
- B De stroomsterkte is **2** maal zo **groot** als die bij draad 1.
- C De stroomsterkte is **2** maal zo **klein** als die bij draad 1.
- D De stroomsterkte is **4** maal zo **klein** als die bij draad 1.

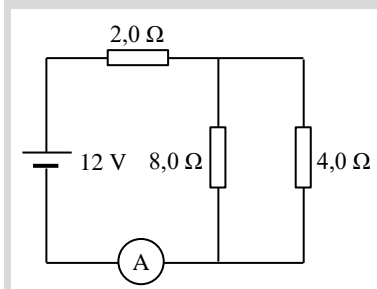
In deze vraag is de essentiële informatie, hier een verandering van de diameter van de draad, vet gedrukt. De antwoordalternatieven lijken sterk op elkaar en zijn allen ‘aanneemelijk’ voor de leerling. Ook hierin zijn de essentiële verschillen vet gedrukt.

Op een spanningsbron zijn vier weerstanden aangesloten op de manier zoals weergegeven in de figuur hiernaast. De spanningsbron geeft een spanning af van 4,5 V. De waarden van de weerstanden staan in de figuur.

Hoe groot is het potentiaalverschil tussen de punten X en Y in de figuur?

- A  $-1,5$  V
- B  $-1,0$  V
- C  $+1,0$  V
- D  $+1,5$  V





De tekstuele informatie in deze vraag wordt beperkt door de introductie van een figuur. Daar waar relevant worden waarden in de tekst herhaald, maar de leerling zal de figuur goed moeten bestuderen. De alternatieven zijn oplopend geordend.

Gegeven is de schakeling zoals weergegeven in de figuur hiernaast. In de schakeling zijn drie weerstanden en een ampèremeter aangesloten op een spanningsbron. Deze bron geeft een spanning af van 12 V. De waarden voor de weerstanden staan in de figuur.

Welke waarde geeft de ampèremeter aan?

- A 0,40 A
- B 0,90 A
- C 2,6 A
- D 4,5 A

Ook in deze vraag is de hoeveelheid tekstuele informatie beperkt door de introductie van een figuur. In de situatie waarin een elektrische schakeling aan de orde is, is dit ook een logische keuze. De alternatieven zijn weer oplopend geordend, en zijn gebaseerd op veelvoorkomende denk- en rekenfouten bij leerlingen.

Het is echter zeer de vraag of deze vraag wel zo geschikt is als gesloten vraag. Er bestaat een gerede kans dat de leerling wel degelijk weet hoe hij of zij aan het antwoord kan komen, maar hierin een (veelvoorkomende) fout maakt. Het antwoord is in deze vraagvorm dan volledig fout, terwijl de beredenering van de leerling (deels) juist kan zijn. Deze vraag zou dan ook beter als open vraag gesteld kunnen worden, zodat ook de gedachtegang van de leerling door de leraar te volgen en te beoordelen is.

Figuur 192– Voorbeelden van meerkeuzevragen waarop de constructieregels voor gesloten vragen zijn toegepast.

## 5.3 Toetsconstructie

Het construeren van toetsvragen is een complex proces, zoals in de vorige paragraaf gebleken is. Dat geldt ook voor het construeren van een complete toets. In deze paragraaf geven we eerst een overzicht van de *kwaliteitscriteria* voor een goede toetsvraag en/of toets, gevolgd door een globale beschrijving van het proces van toetsontwikkeling in de *toetscyclus*.

### 5.3.1 Kwaliteitscriteria

Om te controleren of een geconstrueerde toetsvraag en/of toets van voldoende kwaliteit is, kunnen de volgende – hieronder verder uitgewerkte – criteria worden gebruikt: *validiteit*, *betrouwbaarheid* en *transparantie*. Dit zijn de drie belangrijkste kwaliteitscriteria, die elk op hun beurt opgesplitst kunnen worden in enkele deelcriteria. Deze kwaliteitscriteria gelden op zowel microscopisch niveau (een enkele toetsvraag) als macroscopisch niveau (een volledige toets).

#### Validiteit

Voor een toetsvraag of toets is het belangrijk dat deze *valide* is. Daarbij zijn twee deelcriteria te onderscheiden: *vormvaliditeit* en *inhoudsvaliditeit*.

**Vormvaliditeit** – Een toetsvraag of toets is *vormvalide* als deze meet wat vooraf, door de constructeur, beoogd is. Met een toets voor het vak natuurkunde heeft een leraar als doel het conceptuele begrip van leerlingen te meten. Als de toetsvragen echter bijvoorbeeld een te groot beroep doen op de leesvaardigheid van de leerling, is het zeer de vraag of dit conceptuele begrip wordt gemeten. Een leerling kan namelijk wel conceptueel het juiste niveau hebben, maar minder leesvaardig zijn dan noodzakelijk is voor het maken van de toets. Op dat moment zal de leerling niet het resultaat behalen dat zijn of haar conceptuele begrip goed weergeeft. De toets is daarmee niet voldoende vormvalide.

**Inhoudsvaliditeit** – Een toets is *inhoudsvalide* als de opgaven en opdrachten uit de toets het te toetsen onderwerp inhoudelijk dekken. Als een leerling een toets krijgt over een bepaald hoofdstuk en de vragen worden alleen maar gesteld over een specifiek onderdeel (bijvoorbeeld een paragraaf uit het hoofdstuk), is er geen



Figuur 193 – Het belang van de inhoudsvaliditeit van een toets.

sprake van inhoudsvaliditeit. Een dergelijke toets doet geen recht aan de inspanning van de leerlingen bij hun voorbereiding op de toets. Een hulpmiddel om de inhoudsvaliditeit van een toets te bewaken is de *toetsmatrijs* (zie paragraaf 5.4.2).

### Betrouwbaarheid

Bij het kwaliteitscriterium *betrouwbaarheid* spelen de volgende deelcriteria een rol: *objectiviteit*, *specificiteit*, *moeilijkheidsgraad* en *toetsomvang*.

**Objectiviteit** – In de beoordeling van een toetsvraag of toets zou er geen verschil mogen bestaan tussen verschillende beoordelaars. Elke beoordelaar zou tot dezelfde afweging moeten komen aangaande de antwoorden die een leerling geeft. Hiertoe is een scherp geformuleerde vraag binnen een goed omschreven situatie of context noodzakelijk. Ook het correctievoorschrift, de criteria waarop een antwoord beoordeeld wordt, moet eenduidig zijn. Door deze eenduidigheid na te streven, ontstaat er een bepaalde mate van *objectiviteit* – en dat wordt door de leerlingen als rechtvaardig beschouwd.

Daarnaast is het van belang om de toetsing vorm te geven vanuit vooraf opgestelde leerdoelen. Voor een leerling schept dat veel duidelijkheid. Als er alleen in het onderwijsaanbod (de lessen) aandacht is voor wat er van de leerling wordt verwacht, kan dit bij hem of haar onduidelijkheid creëren – en dat wordt vaak als onrechtvaardig beschouwd.

**Specificiteit** – Een toetsvraag is specifiek als een leerling een goede kans heeft om de vraag correct te beantwoorden. Voorwaarden hiervoor zijn dat de leerling over een voldoende intellectueel niveau beschikt, de leerstof begrepen heeft en zich goed heeft voorbereid. Als een leerling de leerstof niet goed begrepen heeft, maar een bepaalde vraag wel volledig goed weet te beantwoorden, is de vraag mogelijk niet voldoende specifiek. Het spreekt voor zich dat bij deze *specificiteit* een bepaalde mate van subjectiviteit een rol speelt, maar het is wel aan te bevelen om hiermee rekening te houden.

**Moeilijkheidsgraad** – Afzonderlijke toetsvragen kunnen door leerlingen worden ervaren als moeilijk of makkelijk. Het vraagt veel ervaring met toetsconstructie om de moeilijkheidsgraad van een vraag voor de afname van de vraag goed in te schatten. Toch is het nodig om vooraf per vraag zo'n inschatting te maken om tot een evenwichtig samengestelde toets met een goede *moeilijkheidsgraad* te kunnen komen. Voor de samenstelling van een toets is dan de in figuur 194 weergegeven handreiking niet ongebruikelijk. Achteraf – na afname en correctie van de toets – geeft de cijferspreiding (zie figuur 195 voor een voorbeeld) een indicatie van de moeilijkheidsgraad van de toets.

Een toets *differentieert* goed als deze een duidelijk onderscheid maakt tussen leerlingen die de stof goed en minder goed beheersen. Achteraf geeft de cijferspreiding daarvan een eerste indicatie, en kan een *toetsanalyse* meer informatie geven (zie paragraaf 5.4.3).

De moeilijkheidsgraad van afzonderlijke toetsvragen is in te schatten met behulp van *taxonomieën* als OBIT, RTTI en RTI (zie paragraaf 2.7.1 en 5.4.1), waarmee (toets)vragen zijn te classificeren naar het daarin van de leerling gevraagde niveau van beheersing van de leerstof. Door het opnemen van deze classificatie in de *toetsmatrijs* (zie paragraaf 5.4.2) kan vooraf een eerste indruk van de moeilijkheidsgraad van de toets ontstaan.

**Toetsomvang** – Het is ten slotte voor een betrouwbare meting van belang dat er sprake is van *voldoende opgaven*: de toets is niet te kort én niet te lang. Een hulpmiddel daarbij is het vooraf inschatten van de tijd die een leerling (nodig) heeft per vraag.

### Transparantie

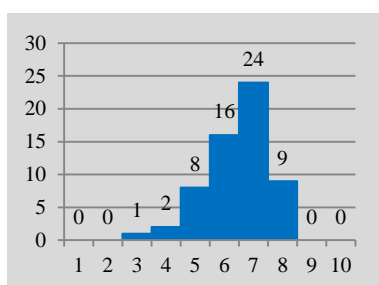
Leerlingen (en ouders) willen in deze tijd graag weten waar ze aan toe zijn in het onderwijs, en dat is een terechte wens. Men verwacht een bepaalde mate van *transparantie* van de scholen. Met name de toetsing ligt dan onder het vergrootglas, omdat dit het meetinstrument is dat de vorderingen van een leerling vastlegt. Dit komt tot uitdrukking in de deelcriteria aanvaardbaarheid en duidelijkheid.

**Aanvaardbaarheid** – Een toets is *aanvaardbaar* als leerlingen en leraren er blijk van geven de toets te accepteren als valide en betrouwbaar. De toets voldoet

### Toetssamenstelling

Voor de samenstelling van een valide toets is de volgende handreiking niet ongebruikelijk: 20% (zeer) moeilijke vragen, 20% (zeer) eenvoudige vragen en 60% vragen van een gemiddelde moeilijkheidsgraad. Met deze opbouw kunnen leerlingen met een voldoende niveau een voldoende resultaat op de toets scoren. Zwakkere leerlingen hebben de mogelijkheid een aantal vragen (toch) goed te beantwoorden en kunnen zo laten zien dat ze zich hebben voorbereid. De sterke(re) leerlingen hebben via de moeilijke vragen de mogelijkheid om zich te onderscheiden. Afwijken van de gegeven handreiking is natuurlijk altijd toegestaan, afhankelijk van de doelen die met een toets worden nastreefd.

Figuur 194 – Handreiking voor de samenstelling van een toets op het criterium moeilijkheidsgraad.



Figuur 195 – Cijferspreiding na correctie van een toets.

### Puntentoekening

Bij de eindexamens wordt het maximaal te behalen aantal punten bij een vraag aangegeven, zodat de leerling een indruk heeft van de omvang van de vraag en de score die ermee te behalen is. Een voordeel daarvan bij een toets is dat het duidelijkheid geeft voor de leerling tijdens het maken van de toets en de bespreking achteraf – en daarmee de transparantie bevordert. Een nadeel is dat het bij de leerling ook spanning kan opleveren (want wat als een leerling een grote vraag niet goed weet te maken?) en dat een scoretoekenning achteraf niet meer kan worden aangepast (bijvoorbeeld als een bepaalde vraag in de toetsanalyse een vreemde score te zien geeft).

Het al dan niet noteren van het te behalen aantal punten bij de vragen in een toets is een beslissing van de individuele leraar en/of de vaksectie.

Figuur 197 – Het in de toets opnemen van de puntentoekening.

aan de verwachtingen (zie *Fokke & Sukke* in figuur 193) en sluit aan bij de leerling. Daarbij kan het ook helpen als de toets ‘ergens over gaat’, zoals weergegeven in figuur 196.

**Duidelijkheid** – Het moet via opgestelde leerdoelen helder zijn waarop de leerling wordt beoordeeld. Deze leerdoelen moeten in de toetsinstrumenten te herkennen zijn. En bij de correctie van een toets moet duidelijk zijn welke criteria er zijn gehanteerd en welke norm er is toegepast. Dat vraagt om het al eerder genoemde eenduidige correctievoorschrift, met daarin ook de toe te kennen punten voor delen van antwoorden die juist zijn. Deze duidelijkheid is essentieel om tot een goede samenwerking te komen met zowel leerlingen als ouders.

Daarnaast kan ervoor worden gekozen om de normering voor een deel al in de toets zelf op te nemen, zoals op de in figuur 197 aangegeven manier.

### Relevantie

Om leerlingen voldoende te motiveren voor onderwijs is het essentieel om dit onderwijs voor hen relevant te laten zijn. Het onderwijs moet ergens over gaan dat de leerlingen interesseert en waarin ze zich willen ontwikkelen. Toetsing is een onderdeel van het onderwijsleerproces, en ook toetsing moet dus een voldoende mate van *relevantie* in zich hebben.

Ook voor het maken van een toets heeft de leerling een bepaalde mate van motivatie nodig. Door de vragen in de toets voor de leerling relevant te maken, ontstaat deze motivatie intrinsiek (buiten de extrinsieke motivatie die een toets doorgaans toch al oproept bij een leerling).

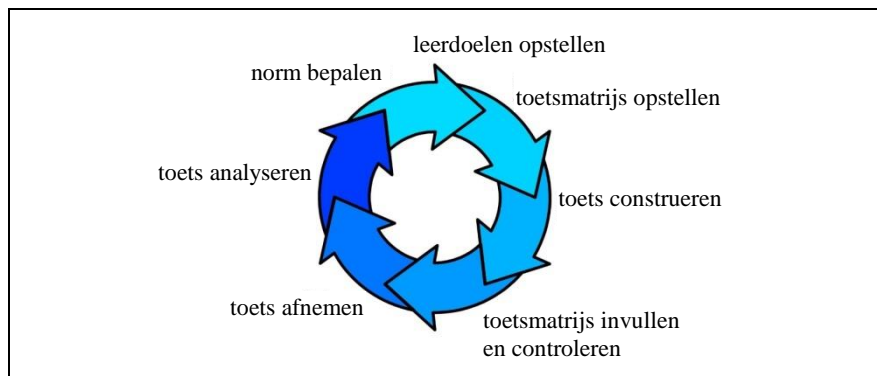
Het ultieme doel van een toetsconstructeur is de leerlingen een interesse en noodzaak te laten ervaren in het beantwoorden van de vraag en/of het oplossen van een gesteld probleem. De Nederlandse eindexamens vormen een goed voorbeeld van relevante toetsvragen. Vrijwel alle vragen worden hierin gesteld vanuit een, over het algemeen voor leerlingen aansprekende, context.

Figuur 196 – Relevantie als aanvullend criterium voor toetsvragen.

## 5.3.2 Toetscyclus

Toetsing wordt in het onderwijs van steeds groter belang, waardoor de meetinstrumenten (de toetsen) van een goede kwaliteit moeten zijn. Alleen dan kunnen er betrouwbare uitspraken worden gedaan en conclusies worden getrokken naar aanleiding van de resultaten van een afgenomen toets.

Om tot een toets te komen die in voldoende mate aan de kwaliteitscriteria validiteit, betrouwbaarheid en transparantie voldoet, is het aan te bevelen om de in figuur 198 weergegeven *toetscyclus* te doorlopen bij de voorbereiding, constructie, afname en analyse van een toets. De verschillende stappen in deze toetscyclus worden in deze paragraaf kort beschreven. De daarin te gebruiken toetsgereedschappen – *taxonomieën*, *toetsmatrijs* en *toetsanalyse* – komen daarna in paragraaf 5.4 aan de orde.



Figuur 198 – De toetscyclus.

**Leerdoelen opstellen** – De toetscyclus start altijd met het formuleren van de *leerdoelen*. Vooraf moet namelijk duidelijk zijn welke doelen er met het geboden

onderwijs worden nagestreefd en op basis van welke te behalen doelen de leerling moet worden getoetst. Vaak worden de leerdoelen al opgesteld bij de start van het onderwijsleerproces. In een later stadium vormen zij ook de basis voor een te construeren toets.

**Toetsmatrijs opstellen** – Om overzicht te krijgen van de opbouw van de toets is een *toetsmatrijs* een geschikt gereedschap. In een toetsmatrijs zijn de leerdoelen te herkennen, en vaak ook de concepten die in de toets aan de orde worden gesteld. Ook kan er een wenselijke puntentelling worden gegeven om een goede balans in de toets te krijgen in relatie tot de leerdoelen en concepten. Vaak wordt er in de toetsmatrijs een *taxonomie* geïntegreerd om de typen vragen te kunnen duiden en hiermee voorafgaand aan de afname van de toets een inschatting te kunnen maken van de moeilijkheidsgraad.

**Toets construeren** – Het *construeren* van de toets is het belangrijkste en meest arbeidsintensieve onderdeel van de toetscyclus. Voor een goede toetsing is namelijk een goed meetinstrument noodzakelijk. In paragraaf 5.2 is uitgebreid ingegaan op de constructie van goede toetsvragen.

**Toetsmatrijs invullen en controleren** – Na het samenstellen van de toets kunnen de eigenschappen van de vragen uit de toets worden ingevuld in de toetsmatrijs. Deze eigenschappen bestaan uit de koppeling naar de leerdoelen en/of concepten, de puntentoekekening per vraag en de beoordeling volgens de taxonomie. De ingevulde toetsmatrijs geeft een beeld van de validiteit en betrouwbaarheid van de toets. Op basis van deze gegevens kan de toets nog worden aangepast.

**Toets afnemen** – Het *afnemen* van een toets bij leerlingen lijkt een eenvoudige activiteit, maar de leraar moet met veel zaken rekening houden. Zo moeten de hulpmiddelen voor leerlingen (bijvoorbeeld een tabellenboek) aanwezig zijn, moet er voor de leerlingen voldoende tijd beschikbaar zijn en moet de toets in een rustige omgeving gemaakt kunnen worden. Ook moet rekening worden gehouden met leerlingen die in aanmerking komen voor extra faciliteiten, zoals tijdverlenging. Het is zeer aan te bevelen om te zorgen voor een zo optimaal mogelijke toetsafname, zodat er geen sprake is van beïnvloeding van de resultaten door een niet optimale toetsorganisatie.

**Toets analyseren** – Na de afname en beoordeling van de toets is het verstandig om een *toetsanalyse* los te laten op de resultaten van de leerlingen per vraag. De data die op deze manier ontstaan, kunnen inzicht geven in de kwaliteit van de afzonderlijke vragen en de toets als geheel. Daarbij wordt gebruik gemaakt van analyse-instrumenten als de *p-waarde* (voor de moeilijkheidsgraad van een vraag), de *RIT-* of *RIR-waarde* (voor het onderscheidend vermogen van een vraag) of de *Cronbach Alpha* (voor de betrouwbaarheid van de toets als geheel).

**Norm bepalen** – Tot slot moet er een *normering* voor de toets worden vastgesteld. Het is aan te bevelen om vooraf een voorgenomen norm vast te stellen, maar het is goed mogelijk dat deze norm op basis van de toetsanalyse moet worden aangepast: niet alle vooraf gemaakte inschattingen zullen correct geweest zijn.

## 5.4 Toetsgereedschap

In deze paragraaf gaan we in op de gereedschappen die in de toetscyclus kunnen worden gebruikt: de (verschillende) *taxonomieën*, de *toetsmatrijs* en de *toetsanalyse* die op de toetsresultaten kan worden losgelaten. Dat laatste vraagt om een zo objectief mogelijke correctie van de toets, wat in paragraaf 5.5 aan de orde zal komen.

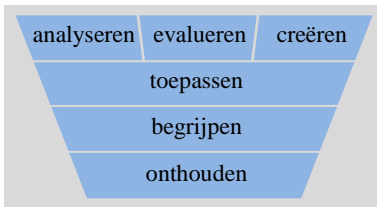
### 5.4.1 Taxonomieën

Een taxonomie in het kader van toetsing geeft een indeling in verschillende soorten kennis en in beheersingsniveau van die kennis. Zo'n taxonomie heeft meestal een hiërarchische opbouw. Bekende taxonomieën zijn de *taxonomie van Bloom*

en de *taxonomie van Krahtwohl*, die te zien is als een uitbreiding van die van Bloom. Voor praktisch gebruik in het (natuurkunde)onderwijs zijn deze taxonomieën aan de complexe kant, en ze worden dan ook meestal gebruikt in een vereenvoudigde vorm als de OBIT-, RTTI- of RTI-taxonomie.

### Taxonomie van Bloom

Eén van de bekendste taxonomieën voor conceptuele kennis is de *taxonomie van Bloom*. In figuur 199 is daarvan de meest recente versie weergegeven. De laagste drie niveaus in de taxonomie van Bloom vormen de lagere-orde denkvaardigheden, met een opbouw van *onthouden*, *begrijpen* en *toepassen*. Het hoogste niveau in deze taxonomie bestaat uit de hogere-orde denkvaardigheden *analyseren*, *evalueren* en *creëren*. In figuur 200 worden de verschillende niveaus voor conceptuele kennis in de taxonomie van Bloom toegelicht, steeds met een voorbeeld van de vaardigheden die van een leerling op een bepaald niveau mogen worden verwacht.



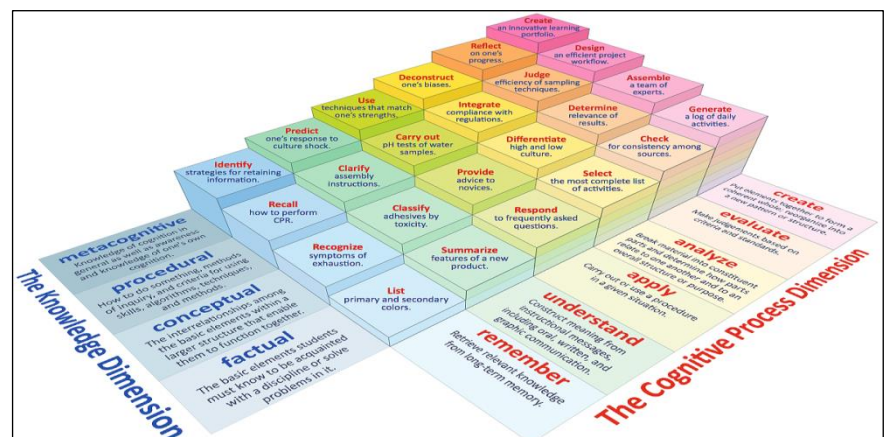
Figuur 199 – Taxonomie van Bloom.

<b>Lagere-orde denkvaardigheden</b>	<b>Onthouden</b>	Een leerling kan informatie reproduceren en/of benoemen.
	<b>Begrijpen</b>	Een leerling kan concepten uitleggen en/of verbanden tussen concepten herkennen.
	<b>Toepassen</b>	Een leerling kan concepten gebruiken in een context.
<b>Hogere-orde denkvaardigheden</b>	<b>Analyseren</b>	Een leerling kan verbanden en relaties tussen concepten onderzoeken.
	<b>Evalueren</b>	Een leerling kan een bepaalde theorie en/of verbanden beoordelen op correctheid en bruikbaarheid.
	<b>Creëren</b>	Een leerling kan ideeën of producten genereren of ontwerpen.

Figuur 200 – Niveaus in de taxonomie van Bloom.

### Taxonomie van Krahtwohl

Een andere bekende taxonomie is de *taxonomie van Krahtwohl*, zoals weergegeven in figuur 201. Deze taxonomie gaat uit van twee dimensies: de *kennisdimensie* en de *denkprocesdimensie*. In de kennisdimensie worden vier vormen van kennis gedefinieerd: feitelijke kennis, conceptuele kennis, procedurele kennis en meta-cognitieve kennis. In de denkprocesdimensie worden zes denkvaardigheden onderscheiden, oplopend in beheersingsniveau: onthouden, begrijpen, toepassen, analyseren, evalueren en creëren – zoals in de taxonomie van Bloom.



Figuur 201 – Taxonomie van Krahtwohl.

Toegepast op het schoolvak natuurkunde kunnen de vormen van kennis uit de taxonomie van Krahtwohl in grote lijnen als volgt worden gekarakteriseerd:

- Feitelijke kennis: verschijnselen zoals faseovergangen, soorten beweging, spiegeling en breking.



- Conceptuele kennis: begrippen en regels zoals kracht, de wetten van Newton en de spiegelingen- en brekingswet.
- Procedurele kennis: procedurele vaardigheden met een stappenplan zoals rekenen met formules, diagrammen gebruiken en meetinstrumenten bedienen.
- Meta-cognitieve kennis: werkwijzen met planmatig handelen zoals onderzoeken, ontwerpen, probleemoplossen, modelleren en waarderen.

In figuur 202 staat een aantal voorbeelden bij de verschillende beheersingsniveaus van de kennisdimensie.

<b>Feitelijke kennis</b> Verschijnselen	De leerling kan: <ul style="list-style-type: none"> <li>• de verschillende faseovergangen opnoemen.</li> <li>• faseovergangen in zijn omgeving herkennen.</li> </ul>
<b>Conceptuele kennis</b> Begrippen en regels	De leerling kan: <ul style="list-style-type: none"> <li>• de drie wetten van Newton opnoemen.</li> <li>• verkeersituaties toelichten met behulp van de drie wetten van Newton.</li> </ul>
<b>Procedurele kennis</b> Vaardigheden	De leerling kan: <ul style="list-style-type: none"> <li>• de eigenschappen van een diagram opnoemen.</li> <li>• een tabel omzetten in een diagram.</li> </ul>
<b>Meta-cognitieve kennis</b> Werkwijzen	De leerling kan: <ul style="list-style-type: none"> <li>• een stappenplan voor onderzoeken opnoemen.</li> <li>• een plan maken voor een onderzoek en dit uitvoeren.</li> </ul>

Figuur 202 – De kennisdimensie uit de taxonomie van Krahtwohl, met voor elk van de soorten kennis een voorbeeld van onthouden en toepassen.

### OBIT, RTTI en RTI

Een op scholen veel gebruikte taxonomie voor conceptuele kennis is de *OBIT-taxonomie*. De letters in OBIT staan achtereenvolgens voor onthouden, begrijpen, integreren en toepassen. De niveaus onthouden en begrijpen zijn van een lagere orde en representeren *reproductief leren*. De niveaus integreren en toepassen zijn van een hogere orde en representeren meer *inzichtelijk leren*. In figuur 203 staat een karakterisering van de niveaus in de OBIT-taxonomie voor het schoolvak natuurkunde.

<b>Onthouden</b>	Bij onthouden gaat het voornamelijk om het leren via reproductie en/of het uit het hoofd leren van vakinhoud.
<b>Begrijpen</b>	Bij begrijpen gaat het bijvoorbeeld om het in eigen woorden kunnen uitleggen van een bepaald concept en/of het beschrijven of ordenen van vakinhoud.
<b>Integreren</b>	Bij integratie gaat het om het leggen van verbanden tussen verschillende concepten of het oplossen van problemen waarin meerdere denkstappen aan de orde zijn.
<b>Toepassen</b>	Bij toepassen van kennis wordt er een bepaalde mate van creativiteit in het denken gevraagd, bijvoorbeeld door het niet specificeren van bepaalde denkstappen en/of door het voorleggen van een nieuwe context waarin kennis moet worden herkend en toegepast.

Figuur 203 – OBIT-taxonomie.

Verreweg de bekendste en meest gebruikte taxonomie voor conceptuele kennis in het voortgezet onderwijs is de *RTTI-taxonomie*. De letters RTTI staan achtereenvolgens voor reproductie, toepassen-1, toepassen-2 en inzicht. In figuur 204 wordt de RTTI-taxonomie nader uitgewerkt voor het schoolvak natuurkunde.

<b>Reproductie</b>	Als de leerling vragen moet beantwoorden op het niveau van reproductie, dan beantwoordt de leerling de vraag puur op basis van herkenning en kennis. Van enige creatieve omgang met het geleerde is geen sprake.
<b>Toepassen-1</b>	In vragen die gesteld worden op toepassingsniveau wordt van de leerling gevraagd om de aanwezige kennis te gebruiken in een

<b>Toepassen-2</b>	bepaalde context. In <i>toepassen-1</i> zal er voor de leerling sprake zijn van een bekende context en/of oplosmethode.
<b>Inzicht</b>	In vragen die gesteld worden op toepassingsniveau wordt van de leerling gevraagd om de aanwezige kennis te gebruiken in een bepaalde context. In <i>toepassen-2</i> zal er voor de leerling sprake zijn van een onbekende context en/of een nieuwe oplosstrategie. Bij een inzichtsvraag krijgt de leerling doorgaans nieuwe en voor hem/haar onbekende informatie voorgeschoteld. De leerling zal deze informatie moeten kunnen verwerken met behulp van aanwezige kennis en creatief moeten zijn in het bedenken van een oplosmethode.

Figuur 204 – RTTI-taxonomie.

Een nadeel van de RTTI-taxonomie is dat het lastig is om een vraag te classificeren als toepassen-1 of toepassen-2. Het feit dat er gebruik wordt gemaakt van bijvoorbeeld een bekende context (toepassen-1) of een onbekende context (toepassen-2) betekent dat de taxonomie niet geheel los staat van het onderwijsleerproces, en dat maakt de objectiviteit twijfelachtig. Het gebruik van de *RTI-taxonomie* kan dan een oplossing zijn. Hierin staat R voor reproductie, T voor toepassen en I voor inzicht. Er wordt dus geen onderscheid gemaakt tussen de twee verschillende niveaus van toepassen, en dat maakt deze aanpassing objectiever en handiger in gebruik.

### 5.4.2 Toetsmatrijs

De opbouw van een toets kan overzichtelijk worden weergegeven in een *toetsmatrijs*. Deze geeft allereerst een indruk van de mate waarin de leerdoelen door de toets worden gedekt. Vaak is er in een toetsmatrijs ook een taxonomie geïntegreerd om een indruk te krijgen van de moeilijkheidsgraad van de afzonderlijke vragen en daarmee ook van de toets als geheel. Er bestaan in de praktijk vele varianten van de toetsmatrijs. In figuur 205 is een voorbeeld weergegeven voor een toets over elektriciteitsleer.

A		B	C										D		
opgave	te behalen punten	paragraaf methode	lading	stroomsterkte	wet van Ohm	weerstand draad	serieschakeling	parallelschakeling	combinatieschakeling	vermogen	elektrische energie	rendement	R	T	I
			4	3	4	3	5	5	5	4	4	2	R	T	I
			± 2	± 1	± 2	± 3	± 2	± 2	± 5	± 2	± 2	± 1			
1															
2															
3															
4															
5															
...															
...															
...															

Figuur 205 – Voorbeeld van een toetsmatrijs (elektriciteitsleer).

In de voorbeeld-toetsmatrijs van figuur 205 is een aantal kolommen aangegeven met de letters A, B, C en D. Hieronder staat een toelichting op de inhoud van deze kolommen.

In de kolommen onder A staan de vraagnummers en het maximaal aantal te behalen punten per vraag.

Kolom B legt een verbinding tussen de toets(matrijs) en de bron die de leerling gebruikt heeft om zich de leerstof eigen te maken. Hierbij kan gedacht worden aan de paragraafnummers waarop een vraag betrekking heeft en/of de leerdoelen die bij een hoofdstuk aan de orde zijn. Als de toetsmatrijs is opgesteld voor een schoolexamen, kunnen in deze kolom de (sub)domeinen uit de syllabus worden geplaatst.

In de kolommen onder C staan in het algemeen de leerdoelen, hier verkort tot de concepten die in de toets aan de orde kunnen of moeten komen. Per concept is steeds aangegeven welk aantal punten in de toets gewenst is, en met welke marge er daarbij rekening mag worden gehouden. Dit kan de leraar vooraf zelf beslissen, maar het kan ook zijn dat er in de vaksectie op school afspraken worden gemaakt over de puntenverdeling die nodig is om tot een valide en betrouwbare toets te komen. Opvallend is dat bij een aantal concepten het wenselijk te behalen aantal punten gelijk is aan de daaronder weergegeven marge. Dit betekent dat deze concepten niet noodzakelijkerwijs in de toets een plaats hoeven te krijgen. Er ontstaat op die manier voor de leraar ook een bepaalde mate van keuzevrijheid en flexibiliteit om een toets naar eigen inzicht inhoudelijk vorm te geven.

De kolommen onder D geven een taxonomie weer. Per vraag kan worden aangegeven in welk deel van de gekozen taxonomie de vraag het best kan worden geplaatst. Bij een volledig ingevulde toetsmatrijs kan de samensteller van de toets dan in een oogopslag zien of de variatie in de toets en de moeilijkheidsgraad van de toets als geheel naar wens is. In figuur 205 is gekozen voor de RTI-taxonomie (reproductie, toepassen en inzicht), maar hier kan elke willekeurige taxonomie worden opgenomen.

In figuur 206 is de toetsmatrijs van figuur 205 ingevuld voor de eerste vijf vragen om een indruk te geven hoe dat er uiteindelijk uit ziet.

opgave	te behalen punten	paragraaf methode	lading	stroomsterkte	wet van Ohm	weerstand draad	serieschakeling	parallelschakeling	combinatieschakeling	vermogen	elektrische energie	rendement			
			4	3	4	3	5	5	5	4	4	2	R	T	I
			± 2	± 1	± 2	± 3	± 2	± 2	± 5	± 2	± 2	± 1			
1	2	2.1	2										x		
2	4	2.4			2		2							x	
3	3	2.6			1					2				x	
4	5	2.5			1				4						x
5	3	2.3	1	2									x		
...	...														
...	...														
...	...														

Figuur 206 – Voorbeeld van een op de toetsmatrijs van figuur 205 gebaseerde invulling.

Voor de constructie van een toets is het aan te bevelen om eerst een toetsmatrijs te ontwerpen. Op die manier wordt er in een vroeg stadium van de toetscyclus al nagedacht over de manier waarop de toets inhoudelijk moet worden opgebouwd. Het opstellen van een toetsmatrijs kan daarnaast een mooi instrument zijn voor een discussie onder vakcollega's, waarbij deze discussie vaak ook zal gaan over de inhoudelijke invulling van het onderwijs voorafgaand aan de toets.

Na de constructie van de toets kan met de ingevulde toetsmatrijs worden gecontroleerd of de toets voldoet aan de gewenste eisen wat betreft validiteit en be-

trouwbaarheid. Ook zorgt het gebruik van een toetsmatrijs binnen de vaksectie op school voor een bepaalde mate van objectiviteit in de toetsconstructie.

### 5.4.3 Toetsanalyse

Na de afname van een toets zijn er data beschikbaar over hoe de leerlingen gescoord hebben op de verschillende onderdelen van de toets en op de toets als geheel. Om een indruk te krijgen van de moeilijkheidsgraad en de betrouwbaarheid van de toets zijn een aantal getalwaarden beschikbaar. In deze paragraaf worden achtereenvolgens de *p-waarde*, de *RIT*- en *RIR-waarde* en de *Cronbach Alpha* besproken.

#### Moeilijkheidsgraad

De moeilijkheidsgraad van een toetsvraag kan worden weergegeven met de *p-waarde*. Deze *p-waarde* heeft een waarde tussen 0 en 1. Een *p-waarde* 0 betekent dat alle leerlingen de vraag fout hadden, terwijl bij een *p-waarde* 1 alle leerlingen de vraag goed hadden. Dus: hoe hoger de *p-waarde* is, des te gemakkelijker is de vraag.

**Open vragen** – De *p-waarde* van een open vraag is te berekenen uit de gemiddelde en de maximale score op een vraag:

$$p = \frac{\text{gemiddelde score}}{\text{maximale score}}$$

Bij de ideale moeilijkheidsgraad van een open vraag wordt ervan uitgegaan dat de vraag een goed onderscheid maakt tussen de ‘kenners’ en ‘niet-kenners’. Bij een *p-waarde* van 0,50 (het midden tussen 0 en 1) splitst de vraag de groep leerlingen in twee gelijke delen. In de praktijk mag rekening worden gehouden met een marge van +/- 0,1 of 0,2.

**Gesloten vragen** – Bij gesloten vragen zoals meerkeuzevragen is de *p-waarde* het percentage leerlingen dat de vraag correct heeft beantwoord:

$$p = \frac{\text{aantal correct}}{\text{aantal totaal}}$$

De ideale moeilijkheidsgraad van een meerkeuzevraag hangt af van de ‘blinde raadkans’ (die er bij een open vraag niet is). Door te gokken kan een leerling een meerkeuzevraag per ongeluk goed beantwoorden. Dit maakt het goed beantwoorden van meerkeuzevragen iets makkelijker, en moet de ideale *p-waarde* dus iets hoger liggen: bij vijf alternatieven is de ideale *p-waarde* 0,60, bij vier alternatieven 0,63, bij drie alternatieven 0,67 en bij twee alternatieven 0,75.

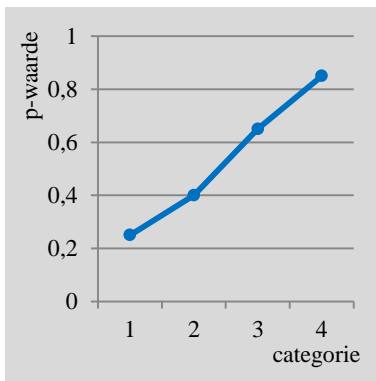
Bij de berekening van de *p-waarde* moet wel worden opgemerkt dat de waarde pas betrouwbaar is als er sprake is van een aanzienlijk aantal leerlingen dat de toets gemaakt heeft. Als richtlijn geldt een minimum van zo’n honderd leerlingen. Bij een kleinere populatie kan de *p-waarde* echter wel degelijk een indruk geven, maar moet deze wel kritisch bekeken worden.

Het aardige van de *p-waarde* is dat deze door leraren vooraf is in te schatten. Dit is een uiterst waardevolle oefening om vooraf de moeilijkheidsgraad van een toets bij benadering al vast te stellen. Ook het gezamenlijk inschatten van de moeilijkheidsgraad van vragen door een groep leraren is een waardevolle activiteit. Er ontstaat daarbij een goed onderbouwde discussie over de toetsing en het onderliggende onderwijs. Het is zelfs het overwegen waard om de ingeschatte *p-waarde* op te nemen in de toetsmatrijs.

Het berekenen van de *p-waarde* van een vraag met de toetsresultaten van de leerlingen is eenvoudig te doen met de hierboven gegeven formules.

#### Onderscheidend vermogen

Het onderscheidend vermogen van een toetsvraag wordt gegeven door de *RIT*- en *RIR-waarde*. Het is wenselijk dat leerlingen die een goede score behalen op de totale toets ook een goede score behalen op een afzonderlijke toetsvraag. De correlatie tussen de score op de toetsvraag en de score op de totale toets moet dus in orde zijn. Aan de *RIT*- en *RIR-waarde* is te zien of dit in voldoende mate het geval is, zoals uitgewerkt in figuur 209.

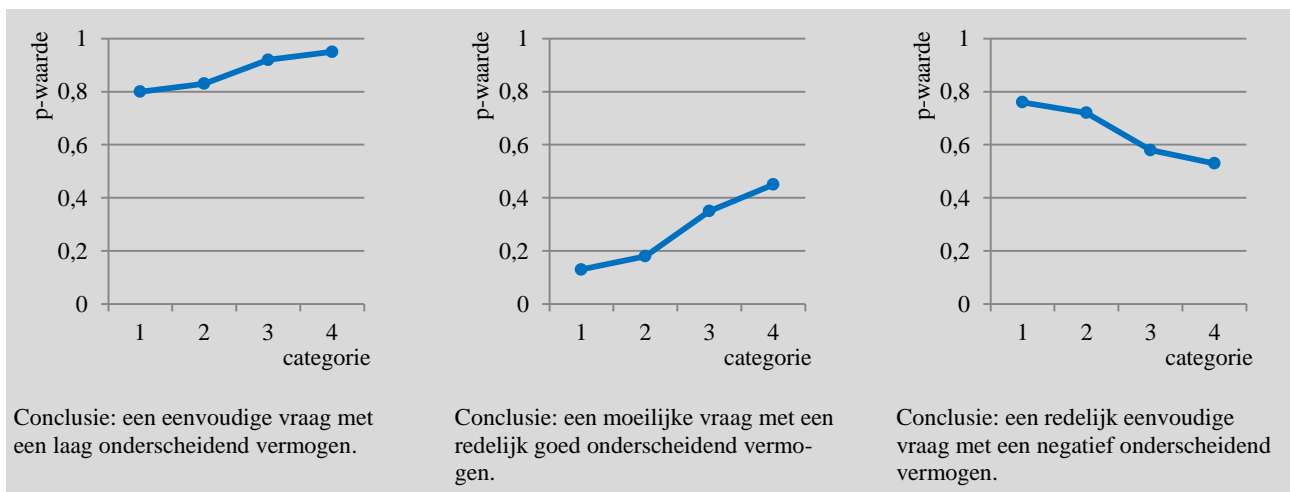


Figuur 207 – Grafische weergave van toetsresultaten: de gemiddelde p-waarde van een toetsvraag in de vier op grond van de totaalscore op de toets onderscheiden categorieën leerlingen.

Het berekenen van de RIT- of RIR-waarde van een toetsvraag met de toetsresultaten van de leerlingen is ingewikkeld. Er zijn wel spreadsheet-programma's voor beschikbaar. Maar met de hieronder weergegeven manier van het grafisch weergeven van de p-waarde van een toetsvraag is al vrij eenvoudig een goede indruk te krijgen van hoe het met het onderscheidend vermogen van die vraag zit.

De toetsresultaten van de leerlingen zijn per vraag grafisch weer te geven zoals in figuur 207. Om tot een dergelijke grafiek te komen wordt de groep leerlingen ingedeeld in vier categorieën. Categorie 1 bestaat uit de 25% leerlingen met de slechtste score op de toets. Categorie 4 bestaat uit de 25% leerlingen met de beste score op de toets. De categorieën 2 en 3 worden op een analoge manier gevuld. Per categorie wordt de p-waarde van een toetsvraag berekend. Deze p-waarde staat uitgezet op de verticale as, en zo ontstaan er vier meetpunten in de grafiek. De ligging van de meetpunten geeft een beeld van het onderscheidend vermogen van de toetsvraag. In het voorbeeld van figuur 207 is dit onderscheidend vermogen goed. In de grafiek is namelijk te zien dat leerlingen die goed scoren op de volledige toets (categorie 4) ook goed scoren op deze vraag, en dat leerlingen die slecht scoren op de volledige toets (categorie 1) ook slecht scoren op deze vraag. De richting van de grafieklijn die door de meetpunten kan worden gedacht, is dus een maat voor het onderscheidend vermogen van de vraag.

Als de lijn in de grafiek een niet al te grote helling heeft, is het onderscheidend vermogen van de vraag gering. Liggen de meetpunten laag op de verticale as, dan zal er sprake zijn van een lage p-waarde over de volledige populatie en is de conclusie dat de vraag een hoge moeilijkheidsgraad heeft. Als de grafieklijn aflopend is, is het onderscheidend vermogen van de vraag negatief. Zwak scorende leerlingen op de volledige toets scoren dan op de vraag beter dan goed scorende leerlingen, wat een onwenselijke situatie is. In figuur 208 staan nog enkele andere voorbeelden van de grafische weergave van de toetsresultaten van leerlingen en de conclusie die daaruit valt te trekken.



Conclusie: een eenvoudige vraag met een laag onderscheidend vermogen.

Conclusie: een moeilijke vraag met een redelijk goed onderscheidend vermogen.

Conclusie: een redelijk eenvoudige vraag met een negatief onderscheidend vermogen.

Figuur 208 – Voorbeelden van de grafische weergave van toetsresultaten en de conclusie die daaruit valt te trekken over de moeilijkheidsgraad en het onderscheidend vermogen van een toetsvraag.

### RIT- en RIR-waarde

De RIT- en RIR-waarde van een toetsvraag geven een indruk van het onderscheidend vermogen van een vraag. Met de RIT-waarde wordt de correlatie (aangeduid met de letter R) bepaald tussen de score op een afzonderlijke toetsvraag (vandaar de letter I van Item) en de totaalscore op de toets (vandaar de letter T). De RIR-waarde wordt op een vergelijkbare manier berekend, met dat verschil dat men nu op zoek is naar de correlatie (R) tussen de score op een toetsvraag (I) en de restscore van de toets (vandaar de letter R). De restscore is de totaalscore op de toets, verminderd met de score op de vraag waarvoor de RIR-waarde berekend wordt. De RIT- en de RIR-waarde zijn in de meeste gevallen nagenoeg gelijk.

De RIT- of RIR-waarde ligt ergens tussen  $-1$  en  $+1$ . Een positieve waarde wil zeg-

### Cronbach Alpha

De Cronbach Alpha geeft een indruk van de betrouwbaarheid van een toets als geheel. De waarde ligt tussen 0 (niet betrouwbaar) en 1 (volledig betrouwbaar). Boven de waarde 0,80 is een toets voldoende betrouwbaar, onder een waarde van 0,70 is de kwaliteit van de toets (zeer) twijfelachtig.

Bij de Cronbach Alpha moet wel de kanttekening worden geplaatst dat deze voornamelijk betekenis heeft bij toetsen die bestaan uit redelijk veel vragen die vergelijkbare competenties toetsen, en als er geen sprake is van een strakke limitering van de tijd die leerlingen voor het maken van de toets hebben.

Figuur 210 – De Cronbach Alpha als maat voor de betrouwbaarheid van een volledige toets.

gen dat er een positief verband bestaat tussen de score op een toetsvraag en de totaal-score op de toets. In de praktijk betekent dit dat leerlingen die goed scoren op de totale toets ook goed scoren op de betreffende vraag. Leerlingen die slecht scoren op de totale toets scoren ook slecht op de betreffende vraag. Deze situatie is ook wenselijk: het onderscheidend vermogen van de vraag is voldoende. Is er echter sprake van een negatieve RIT- of RIR-waarde, dan betekent dit dat leerlingen die goed scoren op de totale toets juist slecht scoren op de betreffende vraag. Leerlingen die op de totale toets slecht scoren, doen het op die vraag juist goed. Het spreekt voor zich dat dit geen wenselijke situatie is, en dat het onderscheidend vermogen van de vraag zwak is.

In de praktijk betekent een RIT-of RIR-waarde boven 0,35 dat de vraag een goed onderscheidend vermogen heeft. Een RIT- of RIR-waarde onder 0,15 betekent een onvoldoende onderscheidend vermogen, met daarbij de opmerking dat een negatieve RIT- of RIR-waarde niet acceptabel is. In dat geval is er echt iets met de inhoud of kwaliteit van de vraag aan de hand. Bij een RIT- of RIR-waarde tussen 0,15 en 0,35 mag er twijfel zijn over het onderscheidend vermogen. Het is dan zonder meer aan te bevelen om de vraag nog eens goed te bekijken. Het inschatten van een RIT- of RIR-waarde voorafgaand aan de afname van een toets is, in tegenstelling tot de p-waarde, onmogelijk. De RIT- en RIR-waarde zijn volledig gebaseerd op de data die volgen uit de toetsresultaten van een groep leerlingen.

Figuur 209 – De RIT- en RIR-waarde als maat voor het onderscheidend vermogen van toetsvragen.

### Betrouwbaarheid

Bij de RIT- en RIR-waarde gaat het om de betrouwbaarheid van afzonderlijke toetsvragen. Er is ook een maat voor de betrouwbaarheid van een toets als geheel: de *Cronbach Alpha*. De vraag is echter hoe relevant die Cronbach Alpha is voor een gemiddelde toets over de leerstof van een hoofdstuk in het leerboek, gezien de in figuur 210 genoemde beperkingen.

## 5.5 Toetspraktijk

In deze paragraaf gaan we in op het *nakijken* van de toets als een van de onderdelen van de toetscyclus. Bij die beoordeling van de toetsresultaten van de leerlingen is het gewenst om zo objectief mogelijk te zijn, maar in de praktijk is er sprake van (menselijke) *beoordelaarseffecten*. En dan ligt er nog een vraag naar wat te doen bij het *teruggeven* en *bespreken* van de gemaakte toets.

### Toets nakijken

Het nakijken van een toets kan niet zonder een vooraf opgesteld *correctiemodel* met het beoogde antwoord op de vragen en een puntenverdeling voor de te onderscheiden aspecten van dat antwoord binnen het beschikbare aantal punten zoals opgenomen in de toetsmatrijs. Als voor een vraag als “De grootheden A en B zijn recht evenredig. Geef op drie verschillende manieren aan wat dat betekent” 2 punten beschikbaar zijn, dan krijgt een leerling bijvoorbeeld 1 punt als twee antwoorden goed zijn, en niets als er maar één antwoord goed is. Bedenk ook hoe je rekenfouten en fouten met afronden, significantie, eenheden en dergelijke rekent. Daarvoor is de examennorm voor het omgaan met dit soort fouten bruikbaar, maar vergeet niet dat een punt aftrek bij een toets met een maximum van 40 punten zwaarder telt dan bij een examen met een maximumscore van 100.

Dan begint de correctie. Schrijf beslissingen over de normering van vaak voorkomende fouten in een andere kleur in het correctiemodel. Dat bevordert de consistentie van de correctie. Het schrijven van toelichtingen bij fouten van leerlingen is voor de leerling prettig en heel nuttig, maar kost zeer veel tijd. Veel gemaakte fouten kunnen beter klassikaal besproken worden.

Bij het nakijken van een toets kunnen *beoordelaarseffecten* optreden: menselijke reacties die afbreuk doen aan de daarbij gewenste objectiviteit. Een overzicht van een deel van deze effecten staat in figuur 211. Het kan heel lastig zijn om deze effecten te vermijden, maar het is in elk geval van belang dat je je als leraar bewust bent van het bestaan ervan.

Hoewel het lastig kan zijn om beoordelaarseffecten te vermijden, zijn er wel een paar mogelijkheden die daarbij kunnen helpen: een toets per vraag nakijken,

per vraag steeds de volgorde van het nakijken van de toetsen van verschillende leerlingen wisselen, en ‘blind’ nakijken (dus zonder naar de naam van de leerling te kijken).

<b>Eerste indruk</b>	Een beoordelaar kan de neiging hebben om te snel te willen oordelen, zeker als hij of zij onder tijdsdruk moet werken.
<b>Logische fout</b>	Als de beoordelaar een gunstige/ongunstige indruk krijgt bij het beoordelen van de antwoorden op de eerste vragen kan dit, ten onrechte, worden doorgezet naar de antwoorden op de volgende toetsvragen. Of: als een leerling een bepaalde vraag goed/fout maakt, de conclusie trekken dat het antwoord op een andere vraag dan ook wel goed/fout zal zijn.
<b>Sympathie en antipathie</b>	De beoordelaar kan een goede/slechte klik hebben met een bepaalde leerling en daardoor deze leerling sneller goed/streng beoordelen.
<b>Projectie</b>	Een beoordelaar kan bepaalde, positieve of negatieve, kenmerken toeschrijven aan een leerling en hiernaar handelen.
<b>Stereotype</b>	Bij stereotypering worden aan een leerling eigenschappen toegekend op basis van de groep waartoe de leerling behoort. Een bekend voorbeeld hiervan is een verschillende benadering van jongens en meisjes in de beoordeling.
<b>Mildheid en strengheid</b>	Een milde/strengere beoordelaar heeft de neiging om steeds te beoordelen boven/onder het (te verwachten) gemiddelde van de toets.
<b>Centrale tendentie</b>	Bij centrale tendentie wil de beoordelaar steeds een score toekennen die op of in de buurt van het (te verwachten) gemiddelde van de toets ligt.
<b>Normverschuiving</b>	Als leerlingen op een bepaalde vraag van de toets steeds slecht/goed scoren, dan kan de beoordelaar de neiging hebben steeds milder/strenger te gaan beoordelen.
<b>Sequentie-effect</b>	Als een grote hoeveelheid foute/goede antwoorden gevolgd wordt door een (deels) goed/fout antwoord, dan kan de beoordelaar de neiging hebben een onevenredig hoge/lage beoordeling te geven.

Figuur 211 – Beoordelaarseffecten.

De ervaring leert dat de normering op enkele punten bijgesteld moet worden na het corrigeren van zo’n vijf toetsen. Doe dat niet al te rigoureuus: het kost veel tijd, en het effect ervan op het eindresultaat is meestal marginaal. Tijdens de correctie zul je nogal eens door het onbehaaglijke gevoel beslopen worden dat de toets te gemakkelijk of te moeilijk was. Begin dan alleen in extreme gevallen opnieuw met nakijken. Maak gewoon de correctie af en bekijk de toets nog eens kritisch. Was hij echt te simpel of te moeilijk? Vraag het bij twijfel aan een vakcollega.

Een verandering van de norm kan achteraf gebeuren door verandering van het aantal punten dat de leerling ook zonder prestatie krijgt (meestal, zoals bij de examens, 10% van het totale aantal punten). Een norm die achteraf aangescherpt wordt, is moeilijk aan leerlingen te verkopen. Ze zullen dan snel het gevoel hebben onrechtvaardig behandeld te worden. Ze zijn daar nogal gevoelig voor, en vergeten snel de keren dat de norm versoepeld werd. Als je een slechte relatie met een klas hebt, is het meestal niet verstandig een norm achteraf aan te scherpen – zelfs niet als dat eigenlijk zou moeten. Dat zou de relatie met de klas blijvend kunnen beschadigen.

Tijdens het corrigeren kom je altijd zwakke punten van de toets op het spoor. Verbeter die na de correctie meteen, dan ligt er voor volgend jaar een betere toets klaar. Uitstellen van het verbeteren tot volgend jaar werkt niet: de meeste verbeteringen zijn dan allang weer vergeten. Het noteren van later aan te brengen verbeteringen kost evenveel tijd als het verbeteren zelf. Houd ook aantekening van de scores van de leerlingen, zodat je later weet of het een (te) moeilijke of juist een gemakkelijke toets was.

### Toets teruggeven

Bedenk vooraf of de leerlingen de toets mogen houden. Als je de toets inneemt, kan hij verbeterd en in een volgend jaar nog eens gebruikt worden.

Deel de toetsen zelf uit en zeg daarbij iets tegen elke leerling, maar zeker tegen de leerling die een onvoldoende terugkrijgt. Laat merken dat je begrijpt wat dat voor die leerling betekent. Zo’n opmerking kan perspectief bieden, bij-

voorbeeld: “Je werkt er wel aan, maar toch heb je de meeste kennisvragen fout. Kijk de volgende keer goed naar wat je moet weten, dan gaat het misschien beter.” Of: “Het is nog wel niet voldoende, maar het gaat naar mijn idee steeds beter. Gewoon volhouden, dan komt het wel goed.” Bij sommige leerlingen zie je zelf ook geen perspectief. Bij die leerlingen kun je de gezamenlijkheid van het probleem benadrukken, bijvoorbeeld met een opmerking als: “Tja, het is helaas weer mis. Ik heb het heel goed nagekeken maar het is nou eenmaal niet meer dan dit. Ik vind het heel jammer. Laten we ervoor zorgen dat het in ieder geval niet slechter wordt.” Er zijn natuurlijk ook leerlingen die niet werken. Die categorie krijgt dan bijvoorbeeld te horen: “Het is een 4, en we weten allebei hoe dat komt. Misschien wordt het tijd dat je daar eens iets aan gaat doen.” Dan ligt het probleem bij de leerling. Dit is hét moment om irritante leerlingen met een onvoldoende uiterst vriendelijk te benaderen: “Ik weet niet waarom je een onvoldoende hebt, maar het zou wel eens iets te maken kunnen hebben met het feit dat we steeds ruzie hebben. Ik wil daar best eens over praten. Als jij dat ook wilt moet je maar eens langskomen.” Dat doet hij overigens meestal niet. Dat je een leerling die het goed gedaan heeft moet prijzen hoeft geen betoog. Een simpel woord als “Een 8, prima” is al voldoende.

Maak geen sarcastische of cynische opmerkingen bij het teruggeven van een toets, tenzij je precies weet waarom je dat bij die ene leerling doet. Soms werkt het, maar meestal wordt het je zeer kwalijk genomen, met name door andere leerlingen die het horen. Voor leerlingen valt dat gedrag van een leraar in de categorieën ‘machtsmisbruik’ en ‘natrappen’.

### Toets bespreken

Je kunt de toets bespreken voor of na het uitdelen. Vóór het uitdelen is de betrokkenheid van de leerlingen bij de bespreking vaak te gering, na het uitdelen te groot. In het laatste geval lukt het bijna nooit om de bespreking in alle rust en stilte te laten verlopen. Streef daar dan ook niet naar, maar zeg bijvoorbeeld wel: “Normaal wil ik dat het stil is als ik iets uitleg. Bij het bespreken van een toets ben ik daar wat minder streng in.” Het is verstandig om alleen de veel gemaakte fouten te bespreken. Schrijf daarbij de normering op het bord. Een andere manier is om de leerlingen hun toets zelf met behulp van het (uitgedeelde) correctievoorschrift te laten analyseren op gemaakte fouten.

Na de klassikale bespreking van de toets en de norm komen leerlingen die het niet eens zijn met hun beoordeling vaak allemaal tegelijk naar je toe. Ga daar dan niet op in en verwijs ze naar een later moment, als de klas aan het werk is. Stuur iedereen naar zijn plaats en laat ze één voor één bij je komen of loop zelf rond. Zorg ervoor dat iedereen die wat tegen je wil zeggen dat ook kan doen. De leerling hoeft het dan niet eens te zijn met je uiteindelijke beslissing, maar moet die wel kunnen aanvaarden. Als hij dat niet kan, verstoort dat je relatie met die leerling. Biedt de leerling dan bijvoorbeeld aan om het werk door een vakcollega te laten nakijken, en stel hem voor dat jullie beiden zich bij zijn of haar beslissing zullen neerleggen.

Sommige leraren vragen de leerlingen terug te komen in de pauze. Dat verdient echter geen aanbeveling: leraar en leerlingen hebben een pauze nodig, en het klassenmanagement moet zo zijn dat die pauzes ook als zodanig gebruikt kunnen worden. Als er in de les geen tijd beschikbaar is of als er veel leerlingen met commentaar zijn, kunnen ze ook hun commentaar op het blaadje schrijven, zodat de leraar er dan thuis eerst nog eens rustig naar kan kijken.

## 5.6 Afsluiting

De centrale vraag voor dit hoofdstuk was: welke eigenschappen heeft een goede toetsvraag, en hoe geef je met die toetsvragen een theorietoets op een verantwoorde manier vorm?

Het construeren van toetsvragen en toetsen is een complex proces, waarbij het belangrijk is om de toetscyclus zorgvuldig te doorlopen – met expliciete aandacht voor de constructieregels en kwaliteitscriteria voor toetsvragen en toetsen



en voor de inzet van de gereedschappen voor toetsanalyse. Daarnaast kan het nuttig zijn om de toetsconstructie in samenwerking met collega's uit te voeren. Deze samenwerking kan bestaan uit het volledig samen construeren van een toets. Het is echter ook mogelijk een collega te vragen een geconstrueerde toets te controleren en/of van feedback te voorzien: het zogenaamde *vier-ogen-principe*.

Bij de bepaling van de waardering die de leerlingen voor een leraar hebben, speelt de manier waarop deze omgaat met toetsing en beoordeling een belangrijke rol. Gelegenheid tot vragen stellen in de les ervoor, snel nakijken, duidelijke toetsen en toegankelijk zijn voor bezwaren tegen de correctie zijn daarbij belangrijk. Bij dat laatste moeten leerlingen het gevoel hebben dat er serieus naar hen geluisterd is. De leraar doet er dan ook goed aan zoveel mogelijk te laten zien op welke wijze hij tot zijn beoordeling is gekomen, en ruime aandacht te besteden aan de bezwaren van leerlingen daartegen.

Dit alles is te zien als een verdere uitwerking op het volgende in paragraaf 2.8 genoemde *leerprincipe* uit de leerpsychologie: *terugkoppelen*.

**Terugkoppelen** – Sluit na de voorafgaande regelmatige terugkoppeling (feedback) over de vorderingen van de leerlingen in de richting van de leerdoelen het onderwijsleerproces af met een valide, betrouwbare en transparante toets die ook de leerling inzicht geeft in wat hij of zij weet en kan en de beoogde leerdoelen in voldoende mate heeft bereikt.

