

## Toetsen

## Inhoud

Toetsen.....	1
Waarom toets je? .....	2
Hoe beoordeel je? .....	3
Eisen aan een goede toets .....	4
Toetstechnische kwaliteit.....	6
Taxonomieën.....	7
Categorieën van kennis .....	9
Handelingswerkwoorden .....	10
Kanttekeningen bij het onderscheid tussen reproductie-, toepassings- en inzichtvragen.....	14
Toetsmatrijs.....	17
Cesuur.....	18
Wat is cesuur? .....	18
Verschillende cesuurmethoden .....	18
Toets- en itemanalyse .....	22
Taalgebruik in opdrachten .....	22
Vragen en opdrachten.....	24
Categoriën vragen .....	24
Checklist gesloten vragen.....	27
Voorbeelden van goede meerkeuzevragen .....	28
Voorbeelden van matchingvragen .....	36
Chronologie-vragen .....	38
Open vragen .....	40
Mogelijke problemen bij het nakijken van open vragen.....	41
De rol van attitude bij het beantwoorden van vragen.....	42

Kortantwoord vragen (structured questions) .....	43
Essayvragen .....	44
Instructiemodel essay vragen.....	44
Mondelinge toets/performances .....	45
Opdrachten met bronnen .....	45
Voorbeelden van opdrachten met bronnen .....	47
Examenvragen .....	50
Formulering van opdrachten.....	51
Schrijven van een vraag.....	51
Het schrijven van vragen is een iteratief proces .....	51
Onbedoelde effecten (bias).....	52
Checklist voor het evalueren van toetsvragen.....	52
De psychologie van het beantwoorden van vragen.....	53
Lezen.....	53
Beoordelingsmodellen .....	54
Verbeteren van toetsen .....	54
Standaardisatie.....	54
Standaard procedure.....	54
Moderation.....	54

## Waarom toets je?

Geen onderwijs zonder beoordeling of afsluiting. Het beoordelen van leerlingen op hun geleverde werk/prestaties tussentijds of ter afronding van een les(senreeks kan zowel summatief als formatief plaatsvinden:

### *Summatief*

Het beoordelen geeft het benodigde inzicht om:

- de leerling, zijn ouders, vak- of teamcollega's, of vervolgonderwijs actuele informatie te geven over wat de leerling kent en kan;
- op basis van het geconstateerde beheersingsniveau te kunnen determineren of selecteren.

### *Formatief*

Het beoordelen geeft het benodigde inzicht om:

- het geconstateerde beheersingsniveau te kunnen bijsturen. (zie het hoofdstuk over formatief handelen).

Om inzicht in het beheersingsniveau van de leerling te krijgen, kunnen diverse leerlingprestaties kritisch worden bekeken, zoals:

- een product (bijvoorbeeld een werkstuk, een ontwerp, een kunstwerk of een presentatie) dat de leerling heeft gemaakt, of een prestatie die hij heeft geleverd (een meting verrichten, een turnoefening uitvoeren, een bijdrage leveren aan een discussie, een evenement organiseren of een minionderneming runnen);
- het leer- en werkproces waarmee de leerling tot zijn product of prestatie is gekomen;
- de kennis die de leerling (re)produceert, bijvoorbeeld in een toets;
- de vak- en algemene vaardigheden die de leerling laat zien;
- de houding waar de leerling in bepaalde situaties blijk van geeft.

## Hoe beoordeel je?

Beoordelen kan op verschillende manieren. Mede afhankelijk van wát er wordt beoordeeld, zijn dit de meest gangbare beoordelingsvormen:

- schriftelijke toets met gesloten en/of open vragen (kennis, inzicht, toepassing, analyse, synthese, evaluatie)
- practicum
- mondelinge toets
- presentatie
- leerverslag
- werkstuk
- portfolio
- logboek

Binnen deze vormen kan worden gekozen voor beoordeling d.m.v.:

- cijfers
- woorden, bijvoorbeeld ook met gebruikmaking van niveaubeschrijvingen zoals rubrics (zie het hoofdstuk over rubrics), referentieniveaus, kerndoelen en/of eindtermen.

### **Transparante beoordelingscriteria**

Het is belangrijk dat leerlingen vooraf bekend gemaakt worden met de beoordelingseisen. Leerlingen leren namelijk meer als zij weten wat het leerdoel/de leerdoelen is/zijn, als zij weten aan welke eisen hun prestatie moet voldoen en als zij feedback krijgen op hoe zij het doen in vergelijking met de eisen.

Een goede zet daarbij is leerlingen (en beoordelaars!) aan het begin van een lessenserie voorbeelden te laten zien van goede en slechte prestaties. Dit geeft beiden houvast om de gestelde eisen te kunnen hanteren (zie voor details het hoofdstuk over formatief handelen).

### **Consistentie tussen beoordeling, basisvisie en leerdoelen**

De keuze voor *wat* wordt beoordeeld, *hoe* wordt beoordeeld en *wie* er beoordeelt, moet uiteindelijk altijd zijn ingegeven door de basisvisie van waaruit de les(senreeks) is opgezet en wordt gegeven. De beoordeling moet aansluiten bij de gestelde leerdoelen. Als het leerdoel bijvoorbeeld is dat leerlingen zelf bronnen kunnen zoeken en deze kunnen checken op betrouwbaarheid, dan zal dat zoeken en beoordelen op betrouwbaarheid duidelijk herkenbare items in de beoordeling moeten zijn. Aan deze consistentie tussen beoordeling, leerdoelen en basisvisie van een les(senreeks), schort het nog weleens in de ontwikkel- en onderwijspraktijk. Vooral bij vaardigheden gebeurt het vaak dat deze wel als leerdoel worden gesteld, maar dat de leerlingen daarop niet worden beoordeeld, óf dat er beoordeling plaatsvindt zonder dat de desbetreffende vaardigheid is aangeleerd of inge oefend.

Uiteraard kunnen de prestaties van leerlingen (kennis en vaardigheden) ook op andere wijze getoetst worden, bijvoorbeeld met een portfolio, onderzoeksverslag, werkstuk of een reeks stellingen.

### Eisen aan een goede toets

#### **De belangrijkste criteria voor een kwalitatief goede toets zijn:**

- Validiteit;
- Betrouwbaarheid;
- Specificiteit;
- Moeilijkheid;
- Discriminerend vermogen;
- Transparantie;
- Beschikbare tijd;
- (Taal)technische kwaliteit.

#### **Validiteit**

Bij validiteit wordt onderscheid gemaakt tussen inhoudsvaliditeit en begripsvaliditeit.

Inhoudsvaliditeit wil zeggen dat de toets de behandelde leerstof op verantwoorde wijze afdekt. Hier is de koppeling tussen het leerdoel van het onderwijs en de toets essentieel. Welke kennis moet een leerling in het voorafgaande onderwijs verworven hebben? Wat moet hij met die kennis kunnen doen? Voorafgaand aan het opstellen van de toets moet dat goed in kaart gebracht worden. De kunst is vervolgens om de toetsvragen evenwichtig over de leerdoelen te verdelen. Dekt de toets de behandelde leerstof, zijn de vragen evenwichtig over de gehele stof verdeeld? Komen in een opdracht wel alle aspecten van het geleerde aan de orde. Bij het bepalen van de inhoudsvaliditeit van een toets of opdracht kan het gebruik van een toetsmatrix behulpzaam zijn. Daarin wordt vastgelegd welke inhoudsaspecten bevraagd worden, in welke verhouding die inhoudsaspecten aan de orde komen en op welk niveau die inhoudsaspecten aan de orde komen. Bij de vraag op welk (denk)niveau de inhoud beheerst moet worden kan de gereviseerde taxonomie van Bloom ((Anderson, Krathwohl, Mayer (2001) van dienst zijn.

]

Begripsvaliditeit wil zeggen dat de toets toetst wat hij moet toetsen. Dat lijkt een vanzelfsprekendheid, maar lang niet alle toetsen zijn begripsvalide. Zo kwamen we bij één van de talen een schriftelijke toets tegen waarin een onderdeel 'spreken' was opgenomen. In dat onderdeel moesten leerlingen een aantal Nederlandse zinnen vertalen tot vragen en opmerkingen die in een dialoog met een leeftijdsgenoot gebruikt konden worden. De opmerkingen en antwoorden van de leeftijdsgenoot werden in de doeltaal weergegeven. Een dergelijke manier van toetsen heeft weinig met spreekvaardigheid te maken. Eerder wordt schrijf- of vertaalvaardigheid getoetst. Als spreekvaardigheidstoets was dit onderdeel dus niet (begrips)valide.

### ***Betrouwbaarheid***

Bij het aspect betrouwbaarheid spelen twee belangrijke vragen een rol:

- Komen verschillende beoordelaars tot dezelfde beoordeling?
- Worden vergelijkbare prestaties hetzelfde beoordeeld?

Veel hangt bij dit geval af van de kwaliteit van de vraagstelling (waarover later) en de kwaliteit van het antwoord- of beoordelingsmodel.

Bij een gesloten vraagstelling speelt de betrouwbaarheidsvraag minder dan bij een open vraagstelling, omdat bij een gesloten vraagstelling het gewenste antwoord vaak van tevoren vastgesteld kan worden. Bij een open vraag of opdracht komt het aan op een heldere vraag- of probleemstelling die precies aansluit op de gewenste leerdoelen. Ook is het belangrijk bij open vraagstelling veel aandacht aan het antwoord- of beoordelingsmodel te schenken.

Als het antwoord- of beoordelingsmodel is opgesteld, kan vervolgens gecontroleerd worden of de vraag of opdracht uitnodigt om het goede antwoord te geven of het gewenste gedrag te vertonen.

### ***Specificiteit***

Een toets of opdracht is specifiek als een leerling pas een goede prestatie kan leveren als hij de vastgestelde leerdoelen heeft bereikt, m.a.w. als hij de leerstof beheerst.

Bij veel toetsen worden tegenwoordig bronnen of teksten gegeven waarbij vervolgens vragen worden gesteld. Als die vragen beantwoord kunnen worden door de tekst of de bronnen goed te lezen, zijn die vragen niet specifiek. Ook als leerlingen de stof niet beheersen, kunnen zij op basis van hun leesvaardigheid de vragen beantwoorden.

### ***Moeilijkheid***

Dit criterium betekent, dat het niveau van de toets past bij het gewenste niveau van de leerling en het gevolgde onderwijs. Dat houdt in dat een toets in zijn totaliteit niet moeilijker mag zijn dan de oefeningen die een leerling ter voorbereiding op de toets heeft gemaakt. Natuurlijk mag er best een vraag in de toets zitten die de betere leerling de kans geeft om zich te bewijzen, maar het niveau van de toets in zijn totaliteit moet aansluiten bij het niveau van het voorafgaande onderwijs.

### ***Discriminerend vermogen***

De toets maakt het verschil zichtbaar tussen leerlingen die de stof beheersen en leerlingen die de stof niet beheersen en tussen 'goede' en 'minder goede' leerlingen;

### **Transparantie**

In het onderwijs dat aan de toets of opdracht voorafgaat is het de leerling duidelijk geworden wat van hem wordt verwacht en hoe er getoetst zal worden. Als de leerling de toets gaat maken, is het hem duidelijk hoe zwaar de verschillende vragen of opdrachten 'wegen' en wat er precies van hem verlangd wordt.

### **Beschikbare tijd**

Een toets moet binnen de beschikbare tijd te maken zijn.

### **(Taal)technische kwaliteit**

Taalgebruik in toetsen. Het is van belang om erop te letten dat de leerlingen bij het maken van een toets of een examen geen fouten zullen maken als gevolg van een voor hen te moeilijk of te vaag taalgebruik. Daarbij is het balanceren op een dun draadje, omdat het taalgebruik wel moet passen bij het niveau dat leerlingen op het moment van de toets zouden moeten hebben. Het is niet de bedoeling dat het taalgebruik bewust simpel gehouden wordt. Ook op dit punt is een goede voorbereiding van de toets in het voorafgaande onderwijs belangrijk.

### Toetstechnische kwaliteit

Bij het formuleren van gesloten vragen kunnen deze aanwijzingen gebruikt worden. Bij het formuleren van open vragen komen deze aanwijzingen van pas.

De presentatie van de toets. Voor de presentatie van de toets kunnen de volgende aanwijzingen gelden:

- Maak gebruik van een uniforme presentatie: hetzelfde lettertype, een vaste kantlijn en regelafstand en een vaste opbouw van de opgaven;
- Nummer de vragen en de bladzijden;
- Druk de opgave op één bladzijde af;
- Zorg ervoor dat tekeningen en schema's goed leesbaar zijn;
- Vermeld per vraag het aantal te behalen punten (bij schriftelijke toetsen).

Zorg er bij praktische toetsing voor, dat het de leerling duidelijk is op welke aspecten hij beoordeeld wordt en hoeveel punten er per aspect verdiend kunnen worden.

### **Tips**

- Maak de toets samen met minstens één collega
- Maak de toets voorafgaand aan het behandelen van de stof, zodat alle betrokken collega's weten wat er van de leerlingen wordt verwacht
- Analyseer na afloop de toets, waarbij gekeken kan worden naar de scores per vraag en naar de spreiding in de cijfers
- Stel zo nodig de toets bij, voor hergebruik
- Zorg ervoor dat het beoordelingsmodel geen verschillende interpretaties mogelijk maakt

Een goede toets is valide, betrouwbaar en dus bruikbaar als deze in ieder geval aan de volgende eisen voldoet:

- De opdrachten moeten een goede afspiegeling zijn van de behandelde leerstof.
- Je toetst wat je wilt toetsen. M.a.w. je bent je ervan bewust wat je met de opdracht wilt toetsen en waarom je dat op deze manier doet.
- Je maakt gebruik van handelingswerkwoorden uit het vak. Het verdient de voorkeur door de leerjaren heen dezelfde handelingswerkwoorden te gebruiken.
- De vorm van de toetsen moet voor leerling herkenbaar zijn, wat betekent dat de opgaven te herleiden moeten zijn tot voor leerlingen duidelijke leerdoelen.
- Het moet de leerlingen duidelijk zijn wat er precies van hen wordt verwacht. Is bijvoorbeeld voldoende duidelijk wat van hen verwacht wordt bij vragen als 'leg uit', 'toon aan', 'gebruik in je antwoord de bron'.
- De toets gaat vergezeld van een duidelijk en zo mogelijk eenduidig antwoordmodel. Dit laatste is vooral belangrijk als de toets gemaakt wordt door leerlingen van verschillende docenten.
- Alleen leerlingen die de leerstof beheersen moeten de vragen correct kunnen beantwoorden.
- De opgaven moeten taalkundig correct zijn. Er mag bij leerlingen geen misverstand ontstaan over de bedoeling van de vraag.
- De opdrachten mogen geen betrekking hebben op tekstverklaren
- De toets moet een juiste moeilijkheidsgraad hebben en onderscheid maken tussen verschillende niveaus.
- Er moet voldoende spreiding zijn tussen goede en slechte resultaten (hoge standaarddeviatie)
- De vragen moeten (naar puntenaantal) gelijkmatig over de stof verdeeld zijn, zodat niet één aspect onevenredig veel gewicht krijgt.
- Het aantal opgaven moet voldoende zijn voor een betrouwbare meting.
- De opgaven bevatten informatie over de lengte van het antwoord en de gewenste vorm.
- Aan een goed antwoordelement wordt één scorepunt toegekend (en geen twee of meer).

## Taxonomieën

Bij het samenstellen van een toets kan het handig zijn gebruik te maken van een taxonomie. Deze kan helpen om zicht te krijgen op de denkprocessen en soorten kennis die je wilt toetsen.

Het is goed hierbij te bedenken dat het maken van onderscheid in denkprocessen geen absolute aanwijzing geeft over de moeilijkheidsgraad van een opdracht. Je kunt een opdracht, waarin een beroep gedaan wordt op een of meer denkprocessen maken voor leerlingen van alle niveaus en alle leerjaren (Brookhart, 2010). Voor hogere leerjaren of meer begaafde leerlingen zal de complexiteit van de opdracht groter zijn en/of de bronnen talrijker of minder gemakkelijk toegankelijk. Ook kan de opdracht meer of minder productief gemaakt worden.

Een opdracht waarin leerlingen tonen over veel inhoudelijke kennis te beschikken kan bijvoorbeeld wel degelijk moeilijker zijn dan een eenvoudige analyseopdracht (Zie ook het hoofdstuk Wat maakt een toets moeilijk).

Het is dan ook een misvatting dat het Bloom ging om een rangschikking in lagere en hogere denkprocessen. Deze misvatting werd in de hand gewerkt door de taxonomie de vorm te geven van een piramide. Beter is het daarom om gebruik te maken van onderstaande afbeelding.



Het is dan ook onjuist te denken dat pas kunt beginnen met analyseren, evalueren en creëren als je eerst aandacht hebt besteed aan onthouden, begrijpen en toepassen. Leerlingen kunnen direct een analyse-, evaluatie- of creëeropdracht krijgen. Uiteraard zal de leerling bij een dergelijke opdracht wel vaak een beroep moeten doen op feitelijke, conceptuele en/of procedurele kennis die behoren tot het domein van de lagere denkvaardigheden.

Wel is het zo dat onthouden, begrijpen en toepassen meestal voorwaardelijk voor de overige denkprocessen, maar dat betekent niet dat ze ook in die volgorde aangeboden moeten worden. Aan leerlingen kan bijvoorbeeld worden gevraagd een bepaald probleem op te lossen, wat een hogere denkvaardigheid veronderstelt. Gedurende deze opdracht zal in veel gevallen ook een beroep worden gedaan op lagere denkvaardigheden. Leerlingen zullen over kennis van en inzicht in de materie moeten beschikken om het probleem op te lossen.

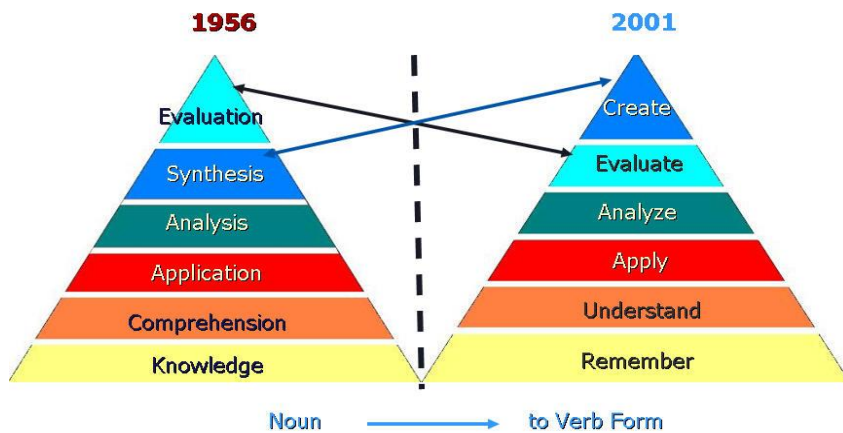
### **Taxonomie van Bloom**

Een veelgebruikte taxonomie om denkprocessen van leerlingen inzichtelijk te maken en te classificeren is de taxonomie van Bloom (1956). Deze taxonomie is in 2001 gereviseerd door Anderson en Krathwohl. In deze versie zijn de begrippen Synthesis en Evaluation omgedraaid en is het begrip Synthesis vervangen door Create.

Een andere wijziging is dat de zelfstandige naamwoorden in de gereviseerde versie zijn vervangen door werkwoorden, waarmee de actievorm wordt onderstreept.



## Changes to Bloom's



### Toelichting

1. Herinneren: Doe je boek dicht en noem de zes onderdelen van Bloom.
2. Begrijpen: Hier zijn drie vragen. Passend bij welk onderdeel van Bloom zijn de vragen gesteld?
3. Toepassen: Hoe kan kennis van Blooms taxonomie het lesgeven verbeteren?
4. Analyseren: Vergelijk Blooms taxonomie met een indeling van anderen.
5. Evalueren: Behoort Bloom een verplicht onderdeel te zijn van opleidingen?
6. Creëren: Ontwerp een nieuwe en betere indeling van denkprocessen.

### Categorieën van kennis

Kennis wordt onderscheiden in vier categorieën.

1. Feitenkennis kennis. Deze wordt onderverdeeld in:
  - a. Kennis van de terminologie
  - b. Kennis van specifieke details en elementen
2. Conceptuele kennis; dit zijn complexere begrippen, die veel relaties hebben met andere begrippen. Deze wordt onderverdeeld in:
  - a. Kennis van classificaties en categorieën
  - b. Kennis van principes en generalisaties
  - c. Kennis van theorieën, modellen en structuren
3. Procedurele kennis; is de kennis over hoe iets uit te voeren bijvoorbeeld onderzoekmanieren, algoritmes, technieken en methoden. Deze kennis wordt onderverdeeld in:
  - a. Kennis van vakspecifieke vaardigheden en algoritmes
  - b. Kennis van vakspecifieke technieken en methodes
  - c. Kennis van criteria om te bepalen welke procedure geëigend is in een bepaalde situatie.
4. Metacognitieve kennis; kennis over cognitie in het algemeen en bewustheid van eigen cognitie. Deze kennis wordt onderverdeeld in:
  - a. Kennis van strategieën

- b. Kennis over cognitieve taken, zowel wat betreft contextuele als conditionele kennis.
- c. Zelfkennis.

Samen vormen de twee dimensies – kennis (zelfstandig naamwoord) en cognitief proces (werkwoord) - een twee dimensionale tabel, zoals in onderstaande figuur te zien is<sup>1</sup>.

<b>Kennis</b>	<b>Cognitief proces</b>					
	Herinneren	Begrijpen	Toepassen	Analyseren	Evalueren	Creëren
Feitenkennis						
Conceptuele kennis						
Procedurele kennis						
Metacognitieve kennis						

### Handelingswerkwoorden

Hieronder staat een overzicht van handelingswerkwoorden die zijn ontleend aan de gereviseerde taxonomie van Bloom

#### 1. Herinneren (kennis)

Een kennisvraag is erop gericht dat je bepaalde feitelijke informatie kunt onthouden en die later kunt reproduceren.

- herkennen
- opsommen
- definiëren
- beschrijven
- aangeven
- aanduiden
- noemen

---

<sup>1</sup> Wilson (2006): Curriculum Pages. Beyond Bloom- A new version of the Cognitive Taxonomy

## 2. Begrijpen (inzichtvragen)

Over inzichtvragen moet je meestal even nadenken, je moet de relevante leerstof onderkennen en het antwoord in eigen woorden weergeven. Je moet de leerstof kunnen uitleggen.

- Uitleggen
- Toelichten
- Selecteren en samenvatten
- Een verklaring geven
- In eigen woorden weergeven
- Een tekening maken van
- Voorspel gevolgen
- Voorbeelden geven
- Uitleggen
- Grote lijnen aangeven
- Beschrijven
- Verschillen en overeenkomsten aangeven (kan ook analyseren zijn)

## 3. Toepassen

Bij toepassingsvragen moet je de leerstof in een onbekende situatie gebruiken om een probleem op te lossen.

- Een plan ontwikkelen
- Oplossingen voorstellen
- Aantonen/bewijzen dat
- Laten zien hoe
- Kennis gebruiken in een situatie
- Concrete gevallen toetsen aan abstracte definities
- Een opgave oplossen of berekening maken

## 4. Analyseren

Bij analysevragen breek je de leerstof op en breng je de onderdelen met elkaar in verband. Daarvoor moet je kritisch en diepgravend studeren. Analysevragen zijn essentieel om kennis uit een leersituatie toe te kunnen passen in de praktijk.

- In delen splitsen
- Patronen beschrijven
- Bewijzen voor conclusies aangeven
- Classificeren
- Onderzoeken
- Vergelijken

### 5. Evalueren (beoordelvragen)

Een beoordelvraag is gericht op een beargumenteerd oordeel en standpunt. Je beoordeelt een idee op zijn waarde, kiest uit verschillende oplossingen voor een probleem, of je beoordeelt een kunstwerk. Je ontwikkelt en verdedigt een opinie.

- Concluderen
- Beargumenteren
- Waarderen/waarde aangeven
- Bekritisieren
- Kiezen en de keuze rechtvaardigen
- Besluiten

### 6. Creëren (ontwerp vragen)

Ontwerp vragen zijn erop gericht onderdelen samen te brengen tot iets nieuws. Je moet creatief omgaan met kennis en inzichten. Bij ontwerp vragen zijn zeer uiteenlopende antwoorden mogelijk.

- Ontwerpen
- Scheppen
- Samenstellen
- Schrijven
- Ontwikkelen
- Voorspellen en extrapoleren
- Combineer kennis op verschillende terreinen

## **RTTI**

Een recentere taxonomie is RTTI<sup>2</sup>, dat staat voor Reproductie, Toepassing 1, Toepassing 2 en Inzicht. Bij reproductie gaat het om het zich herinneren van het geleerde, toepassing 1 heeft betrekking op het toepassen van kennis in een geoefende context, toepassing 2 op vragen in een voor de leerlingen nieuwe context. Bij inzicht, ten slotte, gaat het om kritische reflectie, verbanden en wetmatigheden aanbrengen.

## **OBIT**

Ook van recente datum is OBIT<sup>3</sup>, dat staat voor Onthouden, Begrijpen, Integreren en toepassen.

---

<sup>2</sup> <http://rtti.nl/>

<sup>3</sup> [http://www.aps.nl/in-de-onderwijsvakbladen/-/asset\\_publisher/s7OS0ceq9FZF/document/id/564768;jsessionid=BCD5967F3B4471058E716A99CB3B8B8B](http://www.aps.nl/in-de-onderwijsvakbladen/-/asset_publisher/s7OS0ceq9FZF/document/id/564768;jsessionid=BCD5967F3B4471058E716A99CB3B8B8B)

**Type 1 Onthouden (= herinneren in Bloom)**

Het gaat bij onthouden om handelingswerkwoorden als herinneren, beschrijven, benoemen.

**Type 2 Begrijpen (= begrijpen en toepassen in Bloom)**

Bij begrijpen gaat het om kennis van concepten en principes. Handelingswoorden zijn in eigen woorden weergeven, uitleggen, beredeneren, verklaren, grote lijnen schetsen.

**Type 3 Integreren (= analyseren en evalueren in Bloom)**

Bij integreren gaat het om analyseren, het aan elkaar verbinden van elementen, het activeren van voorkennis, het koppelen van nieuwe kennis aan bestaande kennis, vergelijken, voorspellen, beoordelen, aantonen, relaties zien en combineren.

**Type 4 Toepassen (= creëren in Bloom)**

Bij toepassen gaat het om het gebruik van bestaande kennis in een nieuwe context, het opstellen en toetsen van hypothesen, creëren en ontwerpen.

Voorbeelden van vragen volgens deze taxonomie bij het sprookje van Roodkapje kunnen zijn:

**Type 1 Onthouden**

- Wat wilde Roodkapje bij oma doen?
- Waar ontmoette Roodkapje de wolf?
- Waarvoor waren volgens de wolf zijn grote oren?
- 

**Type 2 Begrijpen**

- Vertel in eigen woorden wat een kapje is.
- Wat voor soort meisje was Roodkapje?
- 

**Type 3 Integreren**

- In welke opzicht lijkt de moeder van Roodkapje op jouw moeder? In welk opzicht is ze verschillend?
- Welke delen van het verhaal zouden werkelijk gebeurd kunnen zijn?

**Type 4 Toepassen**

- Maak een stripverhaal waarin één deel van het verhaal afwijkt van het sprookje en laat de anderen ontdekken of ze het kunnen vinden.
- Wat denk je dat de wolf gedaan zou hebben wanneer grootmoeder niet ziek geweest zou zijn?
- Een opvatting van sommige mensen is dat sprookjes niets met de werkelijkheid te maken hebben, dat het fantasieën zijn. Een andere opvatting is dat sprookjes een metafoor zijn voor de werkelijkheid. Wat is jouw opvatting? Beargumenteer die. Geef voorbeelden van andere

sprookjes en hun (gebrek aan) relatie met de werkelijkheid. Of pas het sprookje van Roodkapje zo aan dat het wel een afspiegeling is van de werkelijkheid.

## Kanttekeningen bij het onderscheid tussen reproductie-, toepassings- en inzichtvragen

Het is tegenwoordig in om onderscheid te maken tussen allerlei soorten toetsvragen, bijvoorbeeld reproductie-, toepassings- en inzichtsvragen. Een onderscheid dat zou helpen om te differentiëren of om onderscheid te maken tussen goede en minder goede leerlingen. Maar als je op zoek gaat naar de vraag wat onder de verschillende soorten toetsvragen wordt verstaan, kom je nogal wat verschillen tegen.

### *Bloom*

In de beroemde taxonomie van Bloom vormen reproductie, inzicht en toepassing de eerste drie denkvaardigheden (de lagere denkvaardigheden). En al is het zeker niet zo dat je je onderwijs moet beginnen met het aanbrengen van de lagere denkvaardigheden, is de volgorde van de lagere denkvaardigheden niet zonder betekenis. Je kunt over feitenkennis beschikken, zonder inzicht in de leerstof te hebben, maar je kunt kennis pas toepassen als je ook over inzicht beschikt.

De betekenis van reproductie is duidelijk, het gaat om het ophalen/ je herinneren van feitelijke kennis. Wanneer werd wie, wanneer en waar geboren, of waar en wanneer vond de slag bij Nieuwpoort plaats. Bij inzichtvragen moet je in de taxonomie van Bloom laten zien dat je de stof hebt begrepen, je moet bijvoorbeeld de relatie(s) tussen twee gebeurtenissen begrijpen en die kunnen uitleggen. Je beschikt over inzicht als je bijvoorbeeld verschillen, overeenkomsten en samenhang ziet tussen ontwikkelingen. Voor inzicht heb je nagenoeg altijd kennis nodig. Vergelijk het met een puzzel. De losse stukjes vormen de kennis en met inzicht leg je de stukjes in elkaar.

Bij toepassingsvragen, ten slotte, moet je verworven kennis en inzichten kunnen gebruiken in een nieuwe situatie.

### *RTTI*

Dit lijkt allemaal heel helder, maar als je vervolgens een blik werpt op de zogenaamde RTTI-taxonomie, komt alles weer op losse schroeven te staan. De RTTI taxonomie draait om te beginnen inzicht en toepassing om, waarbij inzichtvragen duidelijk als moeilijker worden gezien. Bij reproductie is er geen probleem, daar gaat het ook bij RTTI om het je herinneren van feitelijke informatie. Maar vervolgens maakt men bij toepassingsvragen onderscheid tussen twee soorten toepassing. Bij de eerste soort gaat het om het toepassen van kennis volgens een methode waar de leerlingen in de les mee hebben geoefend en met een vraagstelling die de leerlingen bekend voorkomt. Bij de tweede, complexere vorm van toepassing, is de situatie/context voor de leerlingen nieuw. De vraag is weliswaar te beantwoorden met eerder aangeleerde methodes, maar die moeten nu worden gecombineerd op een manier die niet eerder is geoefend. Bij inzicht, dat dus als moeilijker wordt ingeschaald dan toepassen, krijgt de leerling geen methode, context of situatie aangeboden, maar slechts een vraagstuk. Hij moet zelf bepalen hoe hij het probleem gaat aanpakken. De leerling heeft alleen kennis van de leerstof tot zijn beschikking om het vraagstuk op te lossen, waarbij hij aantoont dat hij vanuit verschillende, ook nieuwe, perspectieven kan denken.

Het classificeren van opdrachten is, lijkt mij, niet gebaat bij dit soort verschillende omschrijvingen. Laten we bijvoorbeeld eens onderzoeken welke omschrijving van inzicht de toetsdeskundigen van Cito hanteren. Vraag 16 uit het vmbo-gtl-examen geschiedenis 2019 (eerste tijdvak) wordt door hen geclassificeerd als inzichtsvraag. Wat wordt er in deze, maar ook in de hieraan voorafgaande vraag bij dezelfde bron, van leerlingen verwacht?

Volgens de syllabus moeten de leerlingen het begrip Dolkstootlegende kennen. Je mag er dan ook van uit gaan dat de docent de inhoud ervan in de klas heeft uitgelegd. De vraag is nu, is er hier sprake van reproductie, omdat de leerlingen geleerd hebben dat niet het leger, maar de politiek schuldig was aan de nederlaag van Duitsland, of van toepassing, waarbij de leerlingen hun kennis van de Dolkstootlegende moeten gebruiken om de 'anderen' te identificeren? Afgaande op de relatief lage p-waarde (niet meer dan de helft van de leerlingen hadden deze vraag goed, zou je geneigd zijn dit een toepassingsvraag te noemen, maar misschien heeft ook de slechts de helft van de leerlingen kennisgenomen van de Dolkstootlegende. Het verschil is duidelijk niet eenduidig.

De volgende vraag bestaat uit twee onderdelen, in het eerste moeten leerlingen uitleggen waarom de bron bruikbaar is voor onderzoek naar de nederlaag van de Duitsers. Ik zou zeggen dat een antwoord waarin de leerling zegt dat de bron daar informatie over geeft voldoende is. Beter nog is als hij zegt dat de generaal meent dat het leger geen schuld heeft en daarmee verwijst naar de inhoud van de bron. Merkwaardigerwijs verwijst het antwoordmodel niet naar bruikbaarheid, maar naar betrouwbaarheid: De generaal was ooggetuige of tijdgenoot of geeft zijn mening. Dit zijn allemaal aspecten die gebruikt kunnen worden voor het beoordelen van de betrouwbaarheid van een bron. Of deze aspecten een negatief of positief effect hebben op de betrouwbaarheid is overigens afhankelijk van de vraagstelling. Als de onderzoeksvraag zou luiden: Hoe dacht de top van het Duitse leger over de nederlaag van Duitsland? Krijg je een ander antwoord dan op de onderzoeksvraag: Hoe denken historici (of jij) over de oorzaak van de nederlaag van Duitsland?

Dit betekent meteen ook dat het tweede onderdeel van de vraag geen juiste vraagstelling bevat: 'Geef ook een argument waaruit blijkt dat de inhoud van de bron niet zo betrouwbaar is voor je onderzoek.' De leerlingen had hierop moeten antwoorden: 'De inhoud van de bron is niet zo betrouwbaar, omdat de (Duitse) generaal de gebeurtenissen opzettelijk anders kan hebben voorgesteld / hij de schuld (van het verlies) afschuift op anderen.'

De vraag of het hier om een toepassings- of inzichtsvraag gaat is hiermee niet eenduidig te beantwoorden. Maar ook als we de onjuiste vraagstelling laten voor wat die is, dan lijkt er toch eerder sprake te zijn van toepassing (T1 in de RTTI-taxonomie) dan van inzicht. De leerlingen hebben kennis van de afloop van de Eerste Wereldoorlog en moeten wat zij hebben geleerd over het beoordelen van bruikbaarheid en betrouwbaarheid toepassen op de bron.

**Opdracht 15 en 16 uit het vmbo-gtl examen 2019 eerste tijdvak**

Bron

Een verklaring van een Duitse generaal over de afloop van de Eerste Wereldoorlog (1919): Het Duitse leger heeft geen schuld. De prestatie van het leger is bewonderenswaardig. Waar de schuld van de nederlaag ligt, is dus duidelijk. Het is voldoende bewezen dat anderen schuldig zijn. Zelfs onze

vijanden zeggen het. Zo heeft een Engelse generaal tegen mij gezegd dat ons leger is verraden.

Vraag 1

Gebruik de bron

1p (p-waarde = 52)

De verklaring die de generaal geeft, staat bekend als de Dolkstootlegende. Volgens de generaal zijn 'anderen' schuldig aan de Duitse nederlaag. Wie zijn volgens de generaal de 'anderen' die schuldig zijn aan de Duitse nederlaag?

Vraag 2

Gebruik nogmaals de bron

2p (p-waarde = 55)

Stel: je doet onderzoek naar de nederlaag van het Duitse leger aan het eind van de Eerste Wereldoorlog en je vindt deze bron.

- Geef een argument waaruit blijkt dat de bron bruikbaar is voor je onderzoek.
- Geef ook een argument waaruit blijkt dat de inhoud van de bron niet zo betrouwbaar is voor je onderzoek.

Doe het zo: De bron is bruikbaar, omdat ... De inhoud van de bron is niet zo betrouwbaar, omdat ...

Antwoordmodel

Vraag 1 maximumscore 1

Uit het antwoord moet blijken dat de regering van de Republiek van Weimar/de socialisten/de communisten schuldig is/zijn aan de nederlaag.

Vraag 2 maximumscore 2

Voorbeeld van een juist antwoord is:

- De bron is bruikbaar, omdat de generaal een ooggetuige/tijdgenoot was/zijn mening geeft over de nederlaag 1
- De inhoud van de bron is niet zo betrouwbaar, omdat de (Duitse) generaal de gebeurtenissen opzettelijk anders kan hebben voorgesteld / hij de schuld (van het verlies) afschuift op anderen 1

*Opmerking*



*Het antwoord bij het tweede deel van de vraag dat het de mening van slechts één persoon is, mag niet goed gerekend worden.*

## Toetsmatrijs

Met een toetsmatrijs, een blauwdruk van de toets, kun je de toets heel precies afstemmen op de doelgroep en de doelstellingen van je onderwijs. In een toetsmatrijs kan het volgende worden vastgelegd:

- De leerdoelen (toetstermen)
- De verdeling van de vragen over de toetstermen
- De verdeling van de vragen volgens een taxonomie
- De verdeling van de vragen over de historische vaardigheden

De toetsvragen kunnen bijvoorbeeld worden ingedeeld volgens de taxonomie van Bloom:

		Taxonomie van Bloom					
Toetsterm	Aantal vragen	Herinneren	Begrijpen	Toepassen	Analyseren	Evalueren	Creëren

De toetsvragen kunnen ook worden ingedeeld in kennisvragen en vragen over historische vaardigheden:

		kennis	vaardigheden					
Toetsterm	Aantal vragen		vragen stellen	omgaan met bronnen	contextualiseren	continuïteit en discontinuïteit	oorzaak en gevolg	Historische empathie

<b>Totaal</b>								

## Cesuur

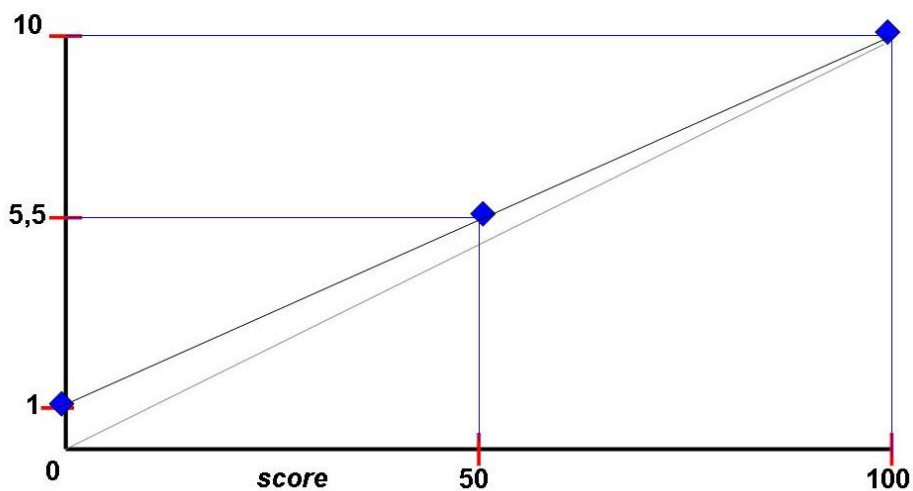
### Wat is cesuur?

De cesuur is de zak/slaaggrens bij de beoordeling van een toets. Daarbij stel je jezelf dus de vraag welke score een student tenminste moet behalen om een voldoende te halen of om te slagen. Het aantal punten wat voldoende is voor een voldoende wordt norm genoemd.

Studenten kunnen bij een tentamen of toetsing vaak een bepaald aantal punten halen (een score). Vervolgens wordt het aantal behaalde punten of de score omgezet naar een cijfer (vaak op een schaal van 1 tot 10). Dit omzetten doe je met behulp van een score-cijfer transformatieschaal (zie figuur). In een score-cijfer transformatieschaal worden de scores die behaald kunnen worden afgezet tegen het cijfer wat daarbij hoort.

De lijn van een transformatieschaal is meestal lineair. Toch zijn er verschillende mogelijkheden om deze lijn te trekken. In andere woorden er zijn verschillende cesuurmethoden die toegepast kunnen worden.

### *cijfer*



### *Voorbeeld van een Score-cijfer transformatieschaal*

#### Verschillende cesuurmethoden

Bij een absolute cesuurmethode wordt de zak/slaaggrens al bepaald voordat de studenten de toetsing hebben afgelegd. Het referentiepunt voor de cesuur ligt hierbij dus in de inhoud van het tentamen en de uiteindelijke scores van de studenten hebben daar bij deze methode dan ook geen invloed meer op.

Nadeel: er wordt geen rekening gehouden met de omstandigheden waarin de toetsing heeft plaatsgevonden met als gevolg mogelijk onterecht gezakte studenten (te strenge norm).

Bij een relatieve cesuurmethode ligt het referentiepunt voor de zak/slaaggrens niet in de inhoud van het tentamen maar in de betrouwbaarheid van de toets, de resultaten van een studentenpopulatie uit een eerder cohort of de resultaten van het huidige cohort.

Nadeel: studenten die zich onvoldoende hebben voorbereid op de toetsing hebben wel invloed op de gemiddelde toetsprestatie met als gevolg mogelijk onterecht geslaagde studenten (te coulante norm).

In de praktijk kom je vaak combinaties van zowel een absolute als relatieve cesuurmethode tegen, die worden ook wel compromismethoden genoemd. Hieronder worden de meest gebruikte cesuurmethoden beschreven.

### **Absoluut**

1. **Vast:** Bij een vaste cesuurmethode wordt vooraf aan de toets op basis van de aard van de vragen en een eventuele correctie voor gokkans een cesuur vastgesteld. Het lastige aan deze manier van cesuur bepalen is dat de norm vrij arbitrair is. Er wordt immers niet op basis van inhoudelijke argumenten gekozen voor een bepaalde cesuur.

2. **Nedelsky:** De Nedelsky methode gaat ook uit van een absolute cesuurmethode, maar hierbij wordt de norm wel beargumenteerd vastgesteld. Nedelsky (1954) stelde voor om een student die een 5,5 zou halen in gedachten te nemen. Vervolgens worden de tentamenvragen door de ogen van deze student bekeken. Daarbij stel je jezelf de vraag: welke antwoordmogelijkheden zou een (hypothetische) 5,5 student kunnen wegstrepen en tussen welke antwoordmogelijkheden zou een 5,5 student twijfelen? Op basis van het aantal antwoordmogelijkheden dat niet wordt weggestreept door een 5,5 student kan de gokkans worden berekend.

Bijvoorbeeld: wanneer van een 5,5 student verwacht wordt dat hij/zij een van de vier antwoordmogelijkheden kan wegstrepen is de gokkans 0.33, aangezien er drie mogelijke antwoorden over blijven.

Door deze gokkans voor iedere vraag te berekenen vanuit het perspectief van de 5,5 student kan een beredeneerde cesuur worden bepaald.

Bij de methode Nedelsky wordt de moeilijkheidsgraad vertaald naar de norm voor een net voldoende (5,5).

In de tabel hieronder is een verdeling weergegeven voor een toets bestaande uit 40 vierkeuzevragen.

Vier niveaus in moeilijkheidsgraad zijn onderscheiden: moeilijk, pittig, redelijk en gemakkelijk.

Een vraag is:

- **Gemakkelijk** als de net voldoende student drie alternatieven moet kunnen elimineren. Voor deze vragen moet aldus een punt worden behaald.

- **Redelijk** : als de 5,5-student twee alternatieven moet kunnen elimineren. In dit geval is de verwachte score 0,5 punt (gelijk aan 50% raadkans).
- **Pittig**: als de 5,5 student één alternatief moet kunnen elimineren. De bijbehorende score is dan  $1/3^e$  –punt.
- **Moeilijk**: vragen zijn naar verwachting te moeilijk voor een 5,5 student en daarom is de te verwachten score gelijk aan de raadkans: 0,25 punt.

De cesuurscore is nu de som van de toegemeten scorepunten.

<b>Voorbeeld:</b>	<b>“Gemakkelijk”</b>	<b>“Redelijk makkelijk”</b>	<b>“Pittig”</b>	<b>“Moeilijk”</b>	totaal
Toets 4-keuzevragen	<i>Onderscheidt onvoldoende-voldoende</i>	<i>Onderscheidt voldoende-ruimvoldoende</i>	<i>Onderscheidt ruimvoldoende van goed presteerders</i>	<i>Onderscheidt goed van zeer goed presteerders</i>	
Verhouding (bijv):	40%	20%	30%	10%	100%
Bij 40 vragen	16	8	12	4	40
Beheersingsgraad	100%	50%	33,3%	25%	
Beredeneerde score 5,5 student	1	1/2	1/3	1/4	
Beheersingsnorm voor voldoende (40 vragen)	16 punten	4 punten	4 punten	1 punt	25
Verwachte P'-waarde bij een slaag% van 75	1.00- 0.70	0.70-0.50	0.50-0.30	0.30-0.10	

Cesuur: score voor 5,5  (Toetsservice: 25)	<b>25</b>
--	-----------

3. **Angoff:** De cesuurmethode van Angoff (1971) is ook een manier om een beredeneerde absolute cesuur te bepalen. In plaats van het maken van een schatting van hoeveel antwoordmogelijkheden een 5,5 student kan wegstrepen, adviseert Angoff om per vraag te schatten welk percentage van een hypothetische groep 5,5 studenten de vraag goed beantwoordt. Door voor alle vragen de geschatte kansen op te tellen ontstaat een beredeneerde norm.

#### ***Relatief***

4. **Wijnen:** De aanname van Wijnen (1971) was dat de moeilijkheidsgraad van een toets in een specifieke context onder specifieke condities alleen afgelezen kan worden aan de scores van studenten. De redenering achter deze cesuurmethode is dan ook dat studenten op basis daarvan beoordeeld moeten worden. In de praktijk betekent dat dat de gemiddelde score van de studenten als uitgangspunt wordt genomen. Deze gemiddelde score min twee standaardmeetfouten geeft een relatief beargumenteerde cesuur. Dit wordt ook wel grating on the curve genoemd.

#### ***Compromismethode***

5. **Cohen-Schotanus:** Het uitgangspunt bij de methode die is voorgesteld door Cohen-Schotanus en collega's (1996) is in principe een absolute cesuur met een relatief referentiepunt namelijk de groep beste studenten. In plaats van het nemen van de theoretisch hoogst haalbare score kan bijvoorbeeld uitgegaan worden van de daadwerkelijk hoogst behaalde score (met name bij grote groepen studenten). Ook kan er bijvoorbeeld gekeken worden naar het gemiddelde van de studenten die scores in het 95e percentiel en hoger. In de score-cijfer transformatieschaal van figuur zou dan bijvoorbeeld niet 100 als hoogste score worden opgenomen, maar 93 omdat dat de hoogst behaalde score was.

Bij het vaststellen van de cesuur in ene toets kun je kiezen voor de absolute of de relatieve methode.

#### *Voordelen van een absolute cesuur*

- is gebaseerd op minimum eisen op basis van de leerstof
- geeft leerlingen duidelijkheid

#### *Nadelen van een absolute cesuur*

- Een absolute cesuur blijkt soms achteraf toch niet zo verantwoord (omdat deze moeilijk is vast te stellen).
- Er wordt geen rekening gehouden met de kwaliteit van de toets.
- De kwaliteit van het onderwijs beïnvloedt het percentage (on)voldoendes.

#### *Voordelen van een relatieve cesuur*

- De kwaliteit van de toets kan worden meegewogen.

- De cesuur wordt bepaald door de afnamegroep, waardoor de kwaliteit van het onderwijs minder invloed heeft.
- Deze is eenvoudig vast te stellen aan de hand van een toetsanalyse
- Het aantal leerlingen met een onvoldoende is telkens ongeveer even groot.

#### *Nadelen van een relatieve cesuur*

- Relatief zwakke leerlingen verhogen de slaagkans voor andere leerlingen
- De kwaliteit van het onderwijs en de toets wordt niet zichtbaar
- Kan niet worden gebruikt bij kleine groepen
- Kan niet worden gebruikt bij herkansers

Filmpje: <https://toetsing.sites.uu.nl/faqs/over-toetsen-7-jaap-milius/>

#### Toets- en itemanalyse

De belangrijkste toetsindices voor een toetsanalyse zijn:

- Frequentieverdeling, een overzicht van de behaalde toetsscores.
- Betrouwbaarheid; de beste betrouwbaarheidsschatting wordt berekend op basis van de onderlinge samenhang tussen de scores op de toetsvragen. Deze wordt aangeduid met coëfficiënt  $a$  voor toetsen met open vragen en met KR-20 voor toetsen met gesloten vragen. Factoren die van invloed zijn op de betrouwbaarheid zijn de kwaliteit van de afzonderlijke vragen, de lengte van de toets en de samenstelling van de groep leerlingen.
- Standaardmeetfout. De standaardmeetfout is een maat voor de gemiddelde onnauwkeurigheid van de toetsscores.
- Gemiddelde moeilijkheid. De gemiddeld behaalde score gedeeld door de maximaal te behalen score ( $P'$ -waarde).

#### Taalgebruik in opdrachten

*Dimitri Verhulst kreeg in een praatprogramma op de Nederlandse televisie ooit het verwijt dat hij wel en beetje beter zijn best had mogen doen om normaal Nederlands te schrijven.*

*Er kwamen in zijn romans allerlei woorden voor die de interviewer niet kende...*

*De Standaard 19 november 2012*

'Na afloop van een proefwerk hoor je leerlingen vaak uitroepen: "Oooh, was dat de bedoeling van de vraag. Ik snapte die vraag niet. Als ik dat had geweten, dan wist ik het wel", schrijft Marijke Katee<sup>4</sup> in een blog over toetstechniek. 'Communicatieproblemen tussen de toetsenproducerende docenten en

---

<sup>4</sup> Marijke Katee (1-10-2012), *Reactie op Toetstechniek voor docenten* (1) <http://onderzoekonderwijs.net/2012/10/01/reactie-op-toetstechniek-voor-docenten-1/#comments>

de toetsenmakende leerlingen komen regelmatig voor en zelfs bij het schriftelijk examen krijgen leerlingen verwarrende vragen voorgelegd.'

Het referentiekader van de docent is volgens haar een van de oorzaken. 'De docent weet precies wat hij bedoelt. Het aantal 'cues' dat hij nodig heeft om zelf te snappen wat zo verschrikkelijk logisch is in de vraagstelling, is meestal gering. Hij leest in veel gevallen meer dan hij schreef.' Vaak leveren minder gelukkig geformuleerde vragen ook geen onoverkomelijke problemen op. De leerlingen weten in veel gevallen namelijk wel wat de docent bedoelt. Ze zijn, bewust of onbewust getraind in het denken van de docent.

Tijdens het centraal examen geldt dat echter niet, dan zijn leerlingen ineens overgeleverd aan andere docenten, die in teamverband aan de opgaven werken. Het vijlen en schaven aan een vraag leidt er dan nogal eens toe dat leerlingen een vraag niet begrijpen. Dat bleek in dit onderzoek.

### *Begrijpelijkheid van de vragen*

Om een opdracht te kunnen maken is het uiteraard belangrijk dat je de opdracht begrijpt. Daar schort het echter bij deze opdracht en deze groep behoorlijk aan. Voor deze opdracht is het essentieel dat leerlingen geen moeite hebben met sleutelwoorden als eenzijdig, partijdig en interpretatie. Van de leerlingen die, in hun ogen, moeilijke woorden hebben omcirkeld geven vier aan het woord partijdig moeilijk te vinden, vijf noemen het woord eenzijdig en acht het woord geïnterpreteerd.

Veel leerlingen (26 van de 45 leerlingen) hebben de moeite genomen opmerkingen bij de vragen te maken en zij zeggen dat zij de vraagstelling van een of beide vragen raar of onbegrijpelijk vinden.

Eén leerling betwijfelt of dit wel een opdracht op havo-niveau is of eigenlijk weet zij zeker dat dat niet het geval is: 'Ik vind de vraagstelling heel lastig. Ik snap het verhaal niet goed en daardoor kan ik de vragen ook niet beantwoorden. De tekst kan ik in mijn hoofd niet goed samenvatten omdat het gewoonweg te moeilijk is. Dit is zeker weten geen havo-niveau'. Een andere leerling die het waarschijnlijk met deze persoon eens is dat het om een lastige opgave gaat geeft tegelijk wel aan dat het om een echt historische opdracht gaat: 'Ik vond de vragen best moeilijk om te beantwoorden want je hebt voor deze vragen historisch inzicht nodig en dat heeft niet iedereen'.

Eén leerling is heel expliciet over het moeilijke karakter. Zij begrijpt niet over welke twee partijen het gaat in vraag een en ze weet niet wat met partijdigheid wordt bedoeld. Over vraag twee schrijft ze: ik heb helemaal niets gelezen over dat Bonifatius de hoofden van ongelovigen wou afhakken, mag deze vraag wat beter worden uitgelegd'. Ook heeft ze moeite met de bron: 'Misschien moet de bron wat duidelijker, want hier staat iets over de doodstraf terwijl er in de bron iets over 'goed met kwaad bestrijden' wordt gezegd. Ik begrijp de bron en vragen niet helemaal. De bron was nog wel te volgen, maar de vragen pasten er niet bij en waren onduidelijk'

Een onbegrijpelijke vraagstelling kan leiden tot emotionele reacties. Zo schrijft een leerling 'p.s. leer Nederlands!'. Een andere leerling onderschrijft deze conclusie volledig: 'De vragen zijn onduidelijk en er staan spelfouten in. Tip: leer beter Nederlands'.

## Vragen en opdrachten

### Categoriën vragen

#### **Reproductievragen**

Leerlingen moeten een weergave geven van verworven informatie. De leerling moet informatie zoeken in zijn geheugen (**herinneren**).

#### **Identificatievragen**

Leerlingen moeten de gegeven informatie vergelijken met informatie die zij eerder geleerd hebben. De leerling haalt uit bronnen (input-informatie) het antwoord met de gevraagde kenmerken. De leerling hoeft de informatie niet toe te passen (**toepassen**) of te bewerken (**begrijpen**).

#### **Productievragen**

Productie-vragen zijn vragen waarbij de geleerde kennis of vaardigheid wordt toegepast in een nieuwe andersoortige situatie en/of bij een ongebruikelijke presentatie van het probleem.

De leerling moet zich eerder geleerde begrippen, regels of methodes herinneren en toepassen. (**toepassen**)

Tot deze categorie horen ook vragen waarbij de leerling een probleem moet identificeren en analyseren (**analyseren**) en/of een beredeneerd oordeel moet vellen (**evalueren**)

#### **Vrije verwerking**

De leerling moet zelfstandig, met een grote mate van persoonlijke vrijheid, het probleem definiëren en de bijbehorende oplosmethode kiezen. De vraag bevat nieuwe probleemsituaties en/of is ongebruikelijk wat oplosmethode betreft of er wordt gevraagd naar evaluatie van gegevens volgens niet gegeven (zelfstandig gekozen) criteria. (**analyse, evaluatie en synthese/creatie**)

#### **Gesloten vragen**

##### *Meerkeuzevragen*

Mensen als Bloom, Englert, Furst, Hill en Krathwohl (1956) waren tegenstanders van meerkeuzevragen, omdat zij meenden dat je er alleen lagere denkvaardigheden als herinneren mee kunt toetsen. Anderen, zoals Hancock (1994) zijn van mening dat je er wel degelijk ook begrip en toepassen, maar ook analyse en waardering mee kunt toetsen.

Masters, Hulsmeyer, Pike, Leichthy, Miller en Verst (2001) hebben voordelen en nadelen van meerkeuzevragen op een rijtje gezet. Zij noemen als voordelen: efficiënt, objectief, gemakkelijk na te kijken, en geschikt om allerlei inhoud te toetsen (blz. 25). Als nadelen noemen zij o.a dat meerkeuzevragen leerlingen uitnodigden om te gokken, dat ze verkeerd geïnterpreteerd kunnen worden en moeilijk om te maken zijn.

##### *Voordelen*

- Beperkte nakijktijd
- De psychometrische kwaliteit van toetsvragen (d.m.v. statistische analyses) kan geanalyseerd worden
- Betrouwbaar.



- Omdat het beantwoorden van de vragen relatief weinig tijd kost kunnen er veel vragen opgenomen worden in de toets. Dit komt de betrouwbaarheid ten goede en zorgt ervoor dat een groot deel van de leerdoelen afgedekt kan worden.
- Meerkeuzevragen kunnen ook worden ingezet om hogere denkprocessen (begrip, toepassing en evaluatie) te toetsen.
- Vragen kunnen worden bijgehouden in een databank en eenvoudig worden hergebruikt.
- Meerkeuzevragen bieden de mogelijkheid om de toets digitaal af te nemen.

### *Nadelen*

- Het formuleren van goede toetsvragen is erg tijdrovend en moeilijk. Onderzoek van Ten Cate en collega's (2008) liet zien dat bij de meeste toetsen die zij analyseerden een van de drie afleiders zo onwaarschijnlijk was dat deze nauwelijks werd gekozen door studenten. Zo'n afleider heeft dan een beperkte discriminerende waarde.
- Ook al kunnen met meerkeuzevragen ook hogere denkprocessen getoetst worden, dat geldt niet voor leeruitkomsten zoals: creativiteit, toelichting geven, schrijfvaardigheid. Het gaat immers om het herkennen van het goede antwoord en niet om het zelfstandig creëren ervan.

### **Moelijkheidsgraad van meerkeuzevragen**

De moeilijkheidsgraad van meerkeuzevragen wordt o.a. bepaald door:

- De kwaliteit van de vraag. Als de afleiders onzorgvuldig zijn geformuleerd, waardoor een of meer van de afleiders snel als onjuist kunnen worden bestempeld wordt de vraag eenvoudiger.
- Een meerkeuzevraag die beantwoord kan worden op basis van globaal aanwezige kennis zal eenvoudiger zijn dan een meerkeuzevraag die leerlingen slechts kunnen beantwoorden via een logische redenering, waarbij duidelijk slechts een van de afleiders juist kan zijn.
- Meerkeuzevragen worden moeilijker als je leerlingen vraagt een verklaring voor hun antwoord te geven. Niet alleen omdat er dan een grote begrip op het vermogen tot historisch denken en redeneren wordt gevraagd, maar ook een beroep wordt gedaan op hun taalvaardigheid.

### *Voorbeeld*

#### **Opdracht Koude Oorlog**

Voor onderstaande opdracht konden leerlingen (3 havo) maximaal vijf punten scoren. Er werd per combinatie alleen een punt toegekend als de toelichting juist was.

In een groep scoorden de leerlingen gemiddeld 1,88 punten op deze vraag ( $p'$ -waarde=.38). Als de vraag zou zijn beoordeeld zonder te vragen om een toelichting scoorden de leerlingen gemiddeld 4,0 punten ( $p'$ -waarde=.81).

In een andere groep scoorden de leerlingen gemiddeld 3,3 punten op deze vraag ( $p'$ -waarde=.66). Als de vraag zou zijn beoordeeld zonder te vragen om een toelichting lag scoorden de leerlingen gemiddeld 4,6 punten ( $p'$ -waarde=.92).

Met de vraag om de antwoorden toe te lichten was de vraag dus in één klas te moeilijk, terwijl de vraag zonder toelichting in beide klassen te makkelijk was.

5p Welke begrippen horen bij elkaar? Licht je antwoord telkens toe. Doe het zo:

1. hoort bij ..., omdat

2. hoort bij ..., omdat

Enz.

1. Berlijnse muur	a. Atoombom
2. Hiroshima	b. Warschaupact
3. Fidel Castro	c. Vietnam
4. Navo	d. Vluchten naar het westen
5. Dominotheorie	e. Cubacrisis

## Checklist gesloten vragen

Als op alle vragen met 'ja' respectievelijk 'nee' geantwoord kan worden, is de vraag correct.<sup>5</sup>

Checklist voor gesloten vragen			
criterium		Ja*	Nee*
Validiteit	1. Is de inhoud van de vraag in overeenstemming met het leerdoel?	*	
	2. Is er een probleem in de stam gesteld?	*	
	3. Is de vraag vakinhoudelijk juist?	*	
	4. Zijn de alternatieven plausibel?	*	
	5. Zijn de alternatieven in alfabetische of logische volgorde gerangschikt?	*	
	6. Sluiten de alternatieven aan op de stam?	*	
	7. Is de vraag geschikt voor het niveau van de studenten die deze cursus volgen?	*	
	8. Is de vraag met gezond verstand op te lossen?		*
	9. Is het een strikvraag?		*
	10. Sluit het juiste alternatief qua woordkeuze en stijl te sterk aan bij de stam in relatie tot de afleiders?		*
	11. Is er een overlap in de betekenis van de alternatieven?		*
	12. Zijn de alternatieven te homogeen dan wel te heterogeen?		*
Objectiviteit	1. Is de stam ondubbelzinnig (eenduidig) geformuleerd?	*	
	2. Zijn deskundigen het eens over het goede antwoord?	*	

<sup>5</sup> (Breukers, H., Doorten, M., Joosten - ten Brinke, D., Loth, F., Moerkerke, G, De Roode, F. & Schlusman, K. (2011). 2011, p. IV-16)

Checklist voor gesloten vragen			
criterium		Ja*	Nee*
	3. Bevat de stam/stelling voldoende informatie om het juiste alternatief te kunnen aanstrepen?	*	
	4. Vraagt de vraag naar de mening van de student?		*
	5. Is meer dan één alternatief juist?		*
Efficiëntie	1. Is het taalgebruik aangepast aan het niveau van de kandidaat/cursus?	*	
	2. Is de vraag grammaticaal goed en logisch geformuleerd?	*	
	3. Is de vorm van de vraag de meest geschikte toetsvorm?	*	
	4. Is de vraag duidelijk, kort en bondig geformuleerd?	*	
	5. Bevat de stam van de vraag meer informatie dan nodig is?		*
	6. Bevat de vraag een dubbele ontkenning?		*
	7. Is de vraag onnodig negatief gesteld?		*
	8. Zijn de zinnen in de vraag of in het probleem te ingewikkeld?		*
	9. Bevat de vraag onnodig vakjargon?		*
Moeilijkheidsgraad	1. Bevat de vraag meer dan één vraag of probleem?		*
	2. Is de stelling of zijn de alternatieven te moeilijk of te gemakkelijk?		*
* Sommige vragen moeten beantwoord worden met 'ja' en sommige met 'nee'. De vakken met de wenselijke antwoorden zijn grijs en bevatten een *.			

Voorbeelden van goede meerkeuzevragen

Voorbeeld van een meerkeuzevraag

### Bron

Vier teksten uit verschillende verdragen.

tekst 1

Alle mensen zijn vrij en gelijk geboren. Zij hebben een verstand en een geweten. Zij behoren zich tegenover elkaar in de geest van broederschap te gedragen.

tekst 2

De partijen komen overeen, dat een gewapende aanval tegen één of meer van de lidstaten als een aanval tegen haar allen zal worden beschouwd.

tekst 3

Het verwijderen van de bestaande belemmeringen vereist eensgezind optreden. Op deze manier garanderen we onze groei, het evenwicht in het handelsverkeer en de eerlijkheid in de concurrentie.

tekst 4

Wij hebben besloten onze krachten te verenigen om de internationale vrede en veiligheid te bewaren. Wij doen een beroep op de internationale instellingen om de economische en sociale vooruitgang van alle volken in de wereld te begunstigen.

*Gebruik de bron*

1p Welke tekst is afkomstig uit het oprichtingsverdrag van de Europese Economische Gemeenschap in 1957?

- a. tekst 1
- b. tekst 2
- c. tekst 3
- d. tekst 4

In onderstaand schema<sup>6</sup> wordt aangegeven aan welke eisen een goede meerkeuzevraag moet voldoen. Voor elk criterium wordt aangegeven of dit betrekking heeft op efficiëntie, specificiteit, relevantie of objectiviteit

E= Efficiëntie

S= Specificiteit

R= Relevantie

O= Objectiviteit

#### **De vraag**

moet aansluiten bij de inhoud van de toetsterm	R
moet aansluiten bij het niveau van de toetsterm zoals aangegeven in de toetsmatrijs	R
mag geen bias (vertekend beeld) door leeftijd, sekse, cultuur, geloof en dergelijke opleveren	S
mag geen dubbele ontkenning bevatten	E

<sup>6</sup> <http://www.teelen.nl/goedetoetsen/>

mag geen strikvraag zijn	O/E
--------------------------	-----

### De stam

moet een duidelijke probleemstelling of vraag bevatten	O/E
mag maar één probleemstelling of vraag bevatten	E/S
moet bij voorkeur positief geformuleerd zijn	E
mag geen absolute of vage formuleringen bevatten	S
mag geen (grammaticale) aanwijzingen in de richting van het juiste antwoord bevatten	S
moet kort en bondig geformuleerd zijn en mag geen overbodige informatie bevatten die de kandidaat kan afleiden	E
mag maar op één manier kunnen worden geïnterpreteerd	O/E
Vermijd zoveel mogelijk de lijdende vorm	E
vermijd zoveel mogelijk samengestelde zinnen	E
Vermijd tangconstructies	E

### De sleutel

mag niet tot discussie tussen deskundige leiden	O
mag geen absolute of vage formuleringen bevatten	S
mag niet het meest gespecificeerde of langste alternatief zijn	S
mag niet als 'alle van bovenstaande' of 'geen van bovenstaande' geformuleerd zijn	R/E

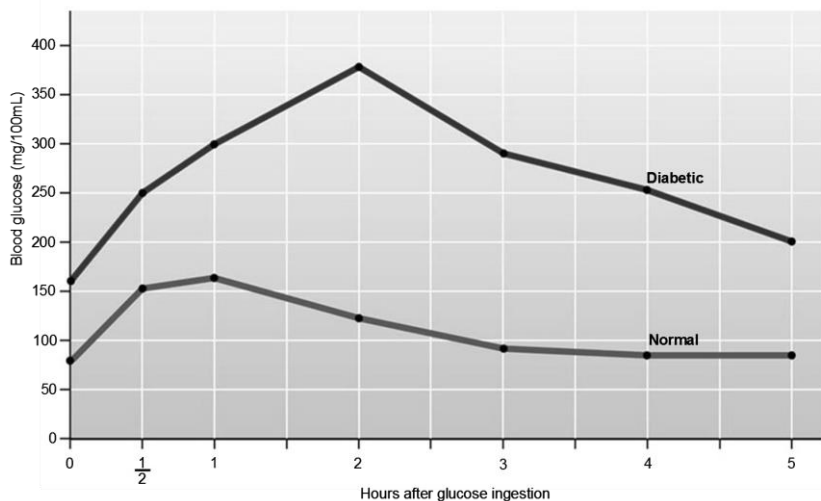
### De afleiders

moeten elkaar uitsluiten (mogen niet onderling afhankelijk zijn)	S
moeten qua lengte, woordgebruik en formulering met elkaar en met de sleutel vergelijkbaar zijn (homogeniteit)	S
moeten plausibel zijn voor degenen die de stof niet goed hebben bestudeerd	S
moeten inhoudelijk en grammaticaal goed aansluiten op de stam	S
mogen geen woordherhalingen bevatten	E
mogen geen absolute of vage formuleringen bevatten	S
mogen niet als 'alle van bovenstaande' of 'geen van bovenstaande' geformuleerd zijn	R/E
moeten in een logische volgorde gerangschikt worden	E/S
mogen geen letterlijke teksten of voorbeelden uit de leerstof bevatten	S

### Voorbeelden van goede meerkeuzevragen

#### Redeneer vraag

The graph below shows the result of blood glucose test from a diabetes patient.



When tested 3 hours after having a carbohydrate rich meal, the blood glucose level of this patient was 3 times higher than that of a normal individual. However, there was no difference in the level of insulin in the blood between the two individuals.

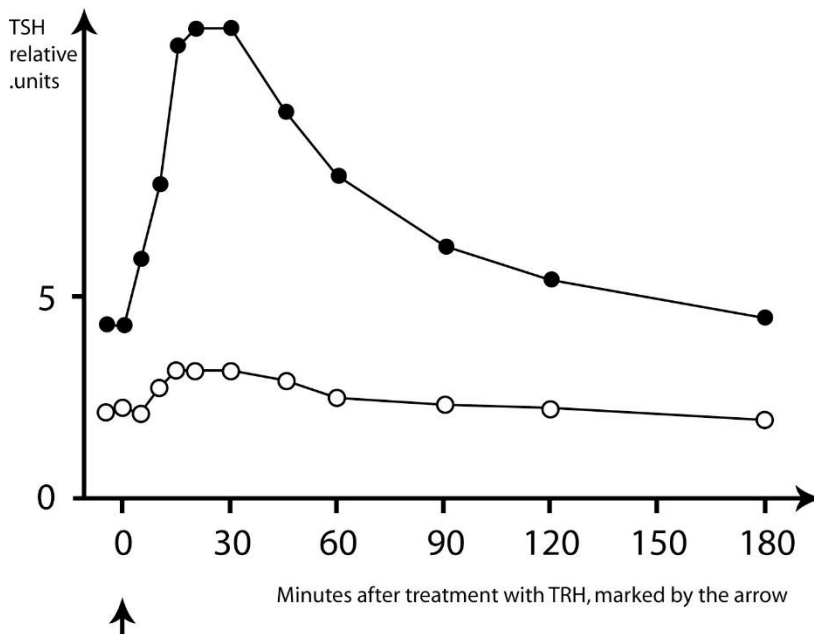
Which of the following could be the reason for diabetic symptoms in this patient?

1. Degradation of pancreatic beta-islets cells.
2. Degradation of pancreatic alpha-islets cells.

3. Abnormal proliferation of pancreatic beta-islet cells.
4. Reduced sensitivity of insulin-receptor mediated signal transduction.
5. Increased sensitivity of insulin-receptor mediated signal transduction.

*Analysevraag*

The following graph shows the concentration of thyroid-stimulating hormone (TSH) in human blood during the 3 hours following an injection of TSH-releasing hormone (TRH) in two groups of people (A and B). One group was treated with thyroxine daily for a week prior to the experiment. (2 marks)



STATEMENT	True (1) or False (2)
a. Thyroxine treatment stimulated TSH release in Group A	
b. Group A has been treated with thyroxine daily before treatment with TRH	
c. Group B has been treated with thyroxine daily before treatment with TRH	
d. Thyroxine treatment inhibited TSH release in Group B	

*Reproductievraag*

I have got three keys here but ..... of them fit.

(Ze passen allemaal niet.)



- a. neither
- b. either
- c. both
- d. none

Dit is in principe een goede vraag omdat twee van de drie afleiders dicht bij het goede antwoord liggen, either and neither (wordt gebruikt bij een keuze uit twee). Slechts both is een zeer onlogische afleider.

#### *Voorbeelden van slechte meerkeuzevragen*

##### *Voorbeeld 1*

Bij dit type vraag is er meestal slechts in schijn sprake van een vraag met vier afleiders. Vaak is er namelijk sprake van een combinatie van twee goed/fout vragen. Dergelijke vragen zijn eigenlijk alleen bruikbaar als de twee beweringen direct met elkaar verbonden zijn. De praktijk leert dat leerlingen met dit type vraag vaak grote problemen hebben, die dan niet worden veroorzaakt door een onvoldoende beheersing van de stof, maar door de lees- en denktijd die het kost om de formulering van het probleem te begrijpen en het probleem op te lossen.<sup>7</sup>

Hieronder volgen twee uitspraken:

I Vóór 1848 mocht elke geloofsrichting zelf bepalen hoe het onderwijs werd ingericht.

II Vrijheid van onderwijs stelt in de praktijk weinig voor als slechts het openbaar onderwijs volledig door de overheid betaald wordt.

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

##### *Voorbeeld 2*

Onderstaande vraag lijkt een goede, op kennis gerichte, meerkeuzevraag. Het is bij deze vraag echter mogelijk om via een redenering die niets te maken heeft met je kennis van het Engelse vocabulair de vraag te reduceren tot een tweekeuze vraag.

Er worden namelijk twee, erg op elkaar lijkende, afleiders gegeven, pretty penny en pretty pound. Het ligt daarom voor de hand dat een van beide antwoorden juist is. Als je vervolgens redeneert dat een pound meer waard is dan een petty en pound dus het, te voor de handliggende, antwoord is, kom je tot het, goede antwoord, pretty penny. Je toetst dan dus eerder de redeneervaardigheid van de leerling, dan zijn kennis van het Engelse vocabulair. En misschien is dat overigens ook wel een belangrijkere vaardigheid.

---

<sup>7</sup> Sanders, P. (red) (2011). *Toetsen op school*, blz.101. Arnhem: Cito

(Uitdrukking: dit is een duur ding.)

This item costs a ..... .

1. heavy penny
2. whole rib
3. pretty penny
4. pretty pound

#### *Voorbeeld 3*

Een vergelijkbare redenering kun je toepassen bij deze vraag. Bij de afleiders staan twee namen van vissen, waardoor het min of meer voor de hand ligt dat een van beide vissen het goede antwoord is.

Uitdrukking: als iets afleidt van het eigenlijke punt/probleem.

This is a red ..... .

1. flag
2. boat
3. salmon
4. herring

#### *Voorbeeld 4*

Eenzelfde redenering kun je volgen bij deze vraag. Twee van de afleiders hebben betrekking op een saus. De kans is daarmee groot dat een van beide het goede antwoord is. Omdat ketchup het modernste woord is en het hier om een uitdrukking gaat is de kans groot dat het antwoord mustard is. Dit blijkt te kloppen.

(Uitdrukking: hij is enorm enthousiast)

He is keen as ...

1. Ketchup
2. Mustard
3. Peas
4. Fried eggs

#### *Voorbeeld 5*

Ik ken de Engelse uitdrukking niet. Desondanks kwam ik via logisch redeneren tot het het juiste antwoord. Italië ligt te veel voor de hand (Rome is tenslotte de hoofdstad van Italië), het bereiden van een maaltijd en het bouwen van een huis zijn vergelijkbare zaken. Het antwoord zal dus wel het villen van een kat zijn, wat juist blijkt te zijn.

Uitdrukking: vele wegen leiden naar Rome.

There are many ways to ..... .

1. prepare a meal
2. build a house
3. Italy
4. skin a cat

*Voorbeeld 5*

Deze vraag is om een andere reden ongelukkig. Hij zet je namelijk op het verkeerde been. Het antwoord, golden, ligt zo voor de hand dat je geneigd bent voor silver te kiezen.

(Uitdrukking: gouden handdruk.)

He has received a ..... handshake.

1. silver
2. blind
3. sweaty
4. golden

*Voorbeeld 6*

In onderstaande vraag staan de afleiders zo ver af van het goede antwoord dat je geen historisch redeneren toetst.



Deze poster uit W.O.II maakte onderdeel uit van een campagne om Amerikanen aan te moedigen om:

1. de economie te ondersteunen door een auto te kopen
2. zuinig te zijn met zaken als brandstof die belangrijk waren voor het voeren van oorlog
3. verkeersproblemen te helpen oplossen door minder gebruik te maken van de auto
4. een autoreis door Duitsland te maken.

#### *Voorbeeld 7*

Vermijd vragen waarin gevraagd wordt naar een fout antwoord (de betrouwbaarheid is gering).

Welke van de volgende beweringen is FOUT?

- A. Een toename van hormoon B veroorzaakt een afname van de hormonen C en D.
- B. Een sterke toename van hormoon C stimuleert de productie van de hormonen A en B
- C. Een lage concentratie van hormoon C belemmert de productie van de hormonen A en B
- D. Een hoge concentratie van hormonen C en D belemmert de afscheiding van de hormonen A en B

#### *Voorbeeld 8*

Een betekenis van het Engelse woord 'box-office' is .....

1. archiefkantoor
2. advertentiebureau
3. theaterkassa
4. buxuskwekerij

Alternatief 4 is een, voor iedereen duidelijk, onzin antwoord.

#### Matching-vragen

Bij matchingvragen moeten leerlingen begrippen/gebeurtenissen/jaartallen uit één rij koppelen aan begrippen/gebeurtenissen/jaartallen uit een andere rij. Dit type opdrachten kan moeilijker worden gemaakt door van leerlingen te vragen hun keuze te verantwoorden. Daarmee wordt de vraag ook meer open.

Voorbeelden van matchingvragen

#### *Voorbeeld 1*

Enkele gebeurtenissen:

- 1 de Russische Revolutie
- 2 oprichting van de NAVO
- 3 uitvinding van de atoombom in de Verenigde Staten

enkele gevolgen:

- a de oprichting van het Warschaupact
- b het begin van een jarenlange wedloop in kernwapens
- c westers wantrouwen ten opzichte van de Sovjetunie

2p Welk gevolg hoort bij welke gebeurtenis? Let op! Elke gebeurtenis en elk gevolg mag maar één keer gebruikt worden.

*Voorbeeld 2*

Het interbellum

Enkele bronnen over het Duitse nationaal-socialisme.

Bron 1

In 1935 werd in Neurenberg bepaald dat joden niet meer met personen van Duits bloed mochten trouwen en geen Duitse staatsburgers waren. Daarna begon de discriminatie van de joden steeds drastischer vormen aan te nemen. Joden in overheidsdienst verloren hun baan, joodse winkels werden geboycot, joodse kinderen werden van school gestuurd.

Bron 2

Hitler had al in zijn boek 'Mein Kampf' (1924) geschreven dat het belangrijkste doel van zijn buitenlandse politiek het verkrijgen van meer gebied voor het Duitse volk was. Oekraïne was bijvoorbeeld voorbestemd om de graanschuur van Duitsland te worden.

Bron 3

De Hitler-Jugend nam de plaats in van andere jongerenorganisaties en de pers werd grotendeels onder controle van de nazi's gesteld, evenals de radio- en filmwereld. Universiteiten werden 'gezuiverd' en kunstenaars moesten lid worden van de Reichskulturkammer.

*Gebruik bron 1, 2 en 3*

2p Vijf kenmerken van Hitlers nationaal-socialisme in Duitsland:

- a de bouw van vernietigingskampen
- b de gelijkschakeling
- c de herbewapening
- d de invoering van rassenwetten
- e de noodzaak tot het scheppen van Lebensraum

Geef per bron aan welk kenmerk erbij hoort. Let op! Elk kenmerk mag slechts één keer gebruikt worden en er blijven twee kenmerken over.

### Chronologie-vragen

Veel chronologievragen kunnen door leerlingen beantwoord worden als zij jaartallen die bij gebeurtenissen horen uit hun hoofd hebben geleerd. Het geniet echter de voorkeur wanneer leerlingen via historisch redeneren tot een antwoord kunnen/moeten komen.

Voorbeelden van chronologievragen

#### *Voorbeeld 1*

Deze opdracht vraagt van leerlingen gebeurtenissen die in tijd dicht bij elkaar liggen in chronologische volgorde te plaatsen.

Plaats de volgende gebeurtenissen in de juiste chronologische volgorde. Schrijf alleen de letters op.

- a. de bestorming van de Bastille
- b. Lodewijk XVI probeert naar het buitenland te vluchten
- c. Napoleon vestigt een dictatuur
- d. de Nationale Vergadering wijkt uit naar een kaatsbaan
- e. de Jacobijnen o.l.v. Robespierre nemen de macht over
- f. de koning roept de Staten Generaal bijeen.

#### *Voorbeeld 2*

Deze opdracht vraagt van leerlingen gebeurtenissen die in tijd verder uit elkaar liggen in chronologische volgorde te plaatsen. Deze opdracht is vooral geschikt als men cumulatief toetst.

De volgende gebeurtenissen hebben met verschillende godsdiensten te maken en staan in willekeurige volgorde:

1. Jezus Christus wordt geboren.
2. De profeet Mohammed trekt van Mekka naar Medina.
3. De Romeinse keizer Constantijn bekeert zich tot het christendom.
4. Jagers en verzamelaars maken grotschilderingen, die volgens onderzoekers uitingen zijn van het geloof dat er goden wonen in planten, dieren, rivieren, bergen enz.
5. Maarten Luther, die de katholieke kerk wil hervormen, verbrandt in het openbaar een officiële brief van de paus, waarmee hij uit de kerk werd verbannen.
6. De monotheïstische joodse godsdienst ontstaat in het Midden-Oosten.

2p Zet deze zes gebeurtenissen in de juiste volgorde van vroeger naar later. Noteer alleen de nummers.

Meer van dergelijke chronologie-vragen zijn te vinden op: <http://histoforum.net/chronologie>

#### *Voorbeeld 3*

In het vwo examen 2011 tweede tijdvak werd voor het eerst de volgende vorm van chronologievraag gebruikt:

Bron 1 tot en met 4 horen bij elkaar en gaan over onderwijs:

### **Bron 1**

*Dit opschrift staat boven de deur van het gebouw uit 1578 van de voormalige Latijnse School te Gouda:*

Praesidium atque decus quae sunt et gaudia vitae; formant hic animos Graeca Latina rudes.

Vertaling:

Om eens tot steun, sieraad en bron van vreugde in het leven te zijn, vormen Grieks en Latijn hier de nog ruwe geest.

### **Bron 2**

*Naar aanleiding van de nieuwe schoolwet van 1806 wordt een schoolreglement opgesteld. Artikel 22 stelt:*

Alle Schoolonderwijs zal zodanig moeten worden ingericht, dat het onder het aanleren van gepaste en nuttige kundigheden, de verstandelijke vermogens der kinderen ontwikkelt, en zijzelf opgeleid worden tot alle maatschappelijke en christelijke deugden.

### **Bron 3**

*Een fragment uit een ingezonden stuk in de Schager Courant van 2 juni 1870:*

Vroeg men, een kwart eeuw geleden, aan een flinke elf à twaalfjarige jongen, die getrouwd de lagere school bezocht: wat doet ge dagelijks wel bij uw meester, of wat leert gij bij hem, dan was 't antwoord: lezen, schrijven en rekenen. Wanneer men thans bij de niet bedeesde leerling onderzoek doet naar de werkzaamheden op school, zo bepalen deze zich niet enkel meer bij lezen, schrijven en rekenen, maar hij voegt er vormleer, geschiedenis, aardrijkskunde, dier-, plant-, natuurkunde en ik weet niet wat al kunde meer aan toe.

### **Bron 4**

*In 1998 wordt de Tweede Fase ingevoerd. In de Leeuwarder Courant van 15 september 1998 verschijnt hierover een artikel:*

Met het studiehuis wordt een onderwijskundig klimaat bedoeld waarin de leerling zelfstandig taken uitvoert. (...) De vernieuwing in de Tweede Fase omvat echter meer dan alleen de invoering van het studiehuis. (...) Er is een aantal nieuwe vakken in het lespakket opgenomen zoals algemene natuurwetenschappen, culturele en kunstzinnige vorming, en management en organisatie. Ook is bij de meeste vakken de inhoud grondig herzien en gemoderniseerd. Zo is het gebruik van computers en grafische rekenmachines (kleine wiskundige zakcomputers) geïntegreerd in de lesmethoden. Oriëntatie op studie en beroep komt ook door middel van opdrachten in de vaklessen aan de orde.

*Gebruik bron 1, 2, 3 en 4.*

Bij deze vier fragmenten doe je de volgende bewering:

Het onderwijs in een bepaalde tijd past bij kenmerkende ontwikkelingen van die tijd.

4p Toon dit voor elk van deze vier fragmenten aan.

### Open vragen

Open vragen kunnen verschillende vormen hebben:

- In- en aanvulvragen
- Korte-antwoordvragen
- Lange-antwoordvragen
- Opstelvragen/betoogvragen

### *Voordelen van open vragen*

- De antwoorden op open vragen geven de docent inzicht in de mate waarin de student zich correct met gebruikmaking van vaktermen kan uitdrukken.
- Open vragen lenen zich uitstekend voor het meten van bijvoorbeeld creativiteit en schrijfvaardigheid.
- Met de open vraag kan een onderwerp meer in de diepte bevraagd worden dan met een gesloten vraag.
- Het antwoord kan niet worden gevonden door herkenning, maar moet worden uitgeschreven.
- De student heeft vrijheid in het beantwoorden van de vraag.
- De antwoorden op een open vraag geven de docent gedetailleerde informatie over de mate waarin studenten de leerstof beheersen en waar problemen zich (eventueel) precies voordoen.

### *Nadelen van open vragen*

- Het maken en nakijken van open vragen vereisen een grote mate van inhoudsdeskundigheid.
- Vrijheid en creativiteit kunnen door studenten worden gebruikt om te bluffen.
- Het is moeilijk open vragen taalkundig zo te formuleren dat er geen misverstand bestaat over de aard en lengte van het te geven antwoord.
- Het nakijken is tijdrovend.
- Het nakijken van open vragen is vaak minder objectief dan het nakijken van gesloten vragen.
- Doordat de beantwoording van een open vraag veel tijd kost, kan slechts een beperkt aantal vragen gesteld worden.
- Voor taalzwakke studenten leveren open vragen extra problemen op. Het is niet altijd duidelijk of een 'foute' beantwoording van de vraag uitsluitend te wijten is aan onvoldoende kennis dan wel (mede) het gevolg is van een onvoldoende taalbeheersing. Met andere woorden: is de taal of de vakinhoud de bottleneck?



Mogelijke problemen bij het nakijken van open vragen

### **Normverschuiving**

De beoordelingsnorm vervlakt of verscherpt ten opzichte van het gemiddelde prestatieniveau van de beoordeelden.

*Bijvoorbeeld:* een beoordelaar wordt meer coulant als het percentage onvoldoendes (te) hoog dreigt te worden, of strenger als de gemiddelde kwaliteit tijdens het beoordelen hoger blijkt dan verwacht.

*Strategie ter voorkoming:* Om ervoor te zorgen dat niet steeds de studenten waarvan de vragen steeds als eerste worden nagekeken, de dupe worden van de normverschuiving bij beoordelaars, is het raadzaam de volgorde van antwoordvellen van studenten regelmatig wijzigen.

### **Sequentie-effect**

Er is een nawerking van voorafgaande beoordelingen van dezelfde schrijfpodracht.

*Bijvoorbeeld:* een beoordelaar heeft de neiging om na enkele zwakke schrijfproducten een volgend redelijk product een relatief hoger cijfer te geven dan wanneer dit product volgt op enkele zeer goede schrijfproducten (zie ook Meuffels, 1992).

*Strategie ter voorkoming:* Het sequentie-effect wordt beperkt door het beoordelen van de antwoorden per vraag (in plaats van per student). Ook het wijzigen van de volgorde van antwoordvellen tijdens de beoordeling is hiervoor een oplossing.

### **Halo-effect**

Een eigenschap van de tekst of van de schrijver heeft een storende uitstraling (halo) op de beoordeling, ook als die eigenschap niet relevant is.

*Bijvoorbeeld:* een goede tekst wordt laag beoordeeld omdat hij er slordig uitziet, of een matige tekst krijgt een hoog cijfer omdat de schrijver bekend staat als hardwerkend en slim (zie ook Nisbett & DeCamp Wilson, 1977).

*Strategie ter voorkoming:* Het halo-effect wordt beperkt door de toets anoniem na te kijken. Studenten wordt gevraagd niet hun naam op de toets te zetten, maar hun ID-nummer of het inschakelen van een tweede beoordelaar, onafhankelijk van de eigen docent.

### **Contaminatie-effect**

De beoordelaar gebruikt de beoordeling voor oneigenlijke doelen.

*Bijvoorbeeld:* een laag cijfer als 'straf' voor iemand die altijd te laat komt, een hoog cijfer voor een student die aanmoediging behoeft, of voor een student die met zijn onderzoeksresultaten aansluit bij de theorie van zijn beoordelaar (zie ook Meuffels, 1991).

*Strategie ter voorkoming:* Het contaminatie-effect is moeilijk oplosbaar zolang er geen scheiding is aangebracht tussen doceren en beoordelen. Intercollegiaal uitwisselen van beoordelingsactiviteiten (docent A beoordeelt de toets van studenten van docent B, en andersom) biedt een oplossing.

### **Persoonlijke tendens**

De individuele normen en gewoontes van de beoordelaar hebben een verhullende uitwerking op de beoordeling.

*Bijvoorbeeld:* een docent geeft op een schaal van tien nooit hoger dan acht of nooit lager dan een vier.

Het is bijzonder lastig om strategieën te ontwikkelen ter voorkoming van een persoonlijke tendens, omdat meestal persoonlijk (impliciete) overtuigingen ten grondslag liggen aan de tendens.

### De rol van attitude bij het beantwoorden van vragen

Leerlingen van havo 3 vonden het in eerste instantie een prima toets, maar achteraf bleek dat ze er te makkelijk over hadden gedacht. Er worden uitvoerige antwoorden met toelichting (gebruik de bron!) gevraagd en dat zijn ze niet gewend. Ze laten daar dan ook veel punten liggen. Dit is ook terug te zien in de werkboeken, waarin leerlingen laten zien dat ze te snel tevreden zijn met een antwoord.

### **Checklist open vragen<sup>8</sup>**

<b>Checklist voor open vragen</b>			
criterium		Ja*	Nee*
Validiteit	1. Is de inhoud van de vraag in overeenstemming met het leerdoel?	*	
	2. Is de vraag vakinhoudelijk juist?	*	
	3. Is de vraagvorm geschikt voor de te toetsen doelstelling?	*	
	4. Is er een duidelijk probleem in de vraag gesteld?	*	
	5. Is de vraag met gezond verstand op te lossen?		*
Objectiviteit	1. Zijn deskundigen het eens over het goede antwoord?	*	
	2. Bevat de vraag voldoende informatie om het juiste antwoord te kunnen aanstrepen?	*	
	3. Is er een antwoordmodel?	*	
	4. Geeft de vraag voldoende informatie over de lengte en vorm van het gewenste antwoord?	*	

<sup>8</sup> Breukers, H., Doorten, M., Joosten - ten Brinke, D., Loth, F., Moerkerke, G, De Roode, F. & Schlusman, K. (2011). 2011, p. V-14.

Checklist voor open vragen			
criterium		Ja*	Nee*
	5. Vraagt de vraag naar de mening van de student?		*

Efficiëntie	1. Is de vraag duidelijk en ondubbelzinnig geformuleerd?	*	
	2. Is de vraag grammaticaal goed en logisch geformuleerd?	*	
	3. Zijn eventuele figuren bij e.d. functioneel?	*	
	4. Zijn eventuele figuren bij de vraag duidelijk?	*	
	5. Zijn eventuele figuren bij de vraag correct?	*	
	6. Is het referentiekader van de vraag duidelijk?	*	
	7. Is de vorm van de vraag de meest geschikte vraagvorm?	*	
	8. Bevat de vraag meer informatie dan nodig is?		*
	9. Bevat de vraag een dubbele ontkenning?		*
	10. Is de vraag onnodig negatief gesteld?		*
	11. Is de formulering van de vraag te ingewikkeld?		*
	12. Bevat de vraag onnodig vakjargon of onnodig moeilijke woorden?		*
	13. Omvat de vraag een te omvangrijk probleem?		*
Moeilijkheidsgraad	1. Heeft de vraag een aanvaardbare moeilijkheidsgraad voor het bedoelde niveau?	*	
	2. Wordt de moeilijkheidsgraad verhoogd door irrelevante gegevens?		*
* Sommige vragen moeten beantwoord worden met 'ja' en sommige met 'nee'. De vakken met de wenselijke antwoorden zijn grijs en bevatten een *.			

#### Kortantwoord vragen (structured questions)

Dit zijn bijvoorbeeld vragen waarin leerlingen een korte omschrijving van een begrip of gebeurtenis moeten geven. Bij dit type vragen is het belangrijk dat leerlingen precies weten wat er van hen verwacht wordt, Bijvoorbeeld: Omschrijf in 200-300 woorden, waarin je de essentie/het wezenlijke weergeeft, de volgende begrippen:

1. Marshall-hulp

## 2. Truman-doctrine

### Voordelen

- Je kunt specifieke kennis toetsen
- Je kunt vaardigheden toetsen zonder dat je ook schrijven toetst
- Je biedt leerlingen structuur waarmee je taalzwakkere leerlingen tegemoet komt

### Nadelen

- Betere leerlingen kunnen niet laten zien wat ze kennen en kunnen
- Normeringsmodel maken moeilijker
- Consistent en nauwkeurig beoordelen moeilijker

### Essayvragen

Bij essay-vragen kun je onderscheid maken tussen 'restricted-response' vragen en 'extended-response'vragen:

- Bij 'restricted-response' essays ligt de nadruk op basiskennis die de leerling moet tonen door betrekkelijk korte en voorgestructureerde antwoorden. Voorbeeld: Leg in drie stappen uit hoe gebeurtenis A, gebeurtenis B beïnvloedde.
- Bij 'extended-response' essays moet de leerling een interpretatie of verklaring geven op basis van een relatief grote hoeveelheid informatie en/of bronnen.

### Voordelen

- Relatief gemakkelijk te maken
- Valide manier om hogere denkvaardigheden te toetsen
- Leerling heeft invloed op de omvang van zijn antwoord

### Nadelen

- Je kunt kennis hiermee minder direct toetsen
- Mogelijk toets je ook schrijfvaardigheden (construct-irrelevant variance)
- Beoordelen is meer subjectief; invloed op betrouwbaarheid

### Instructiemodel essay vragen

#### **Kennisniveau**

1. Je hebt de vraag niet alleen gelezen, maar deze ook begrepen. Voor je de inleiding gaat schrijven, moet je precies weten wat je in de conclusie gaat zeggen. Dat betekent dat je een schema moet maken voor de wijze waarop je gaat antwoorden.
2. In de inleiding schrijf je wat de belangrijkste factoren zijn. In de volgende paragrafen onderbouw je deze factoren met argumenten/bewijzen.
3. In de conclusie geef je een onderbouwd oordeel over het belangrijkste punt in de vraag.

#### **Vaardighedenniveau**

1. Je geeft een volledig, to the point en uitgebalanceerd, antwoord. Je gebruikt de context, inhoud en de herkomst van bronnen om een oordeel te vellen over bruikbaarheid, betrouwbaarheid en nauwkeurigheid. Je gebruikt je eigen kennis om de bron te contextualiseren en de vraag te beantwoorden.
2. Bij vragen over bruikbaarheid geef je de beperkingen van het materiaal aan en geef je een beoordeling met verwijzing naar inhoud en herkomst.
3. Bij het beoordelen van de betrouwbaarheid besteed je aandacht aan de herkomst van de bronnen en gebruik je je eigen kennis om uitspraken te doen over inhoud en context.
4. Je kunt een oordeel geven over een interpretatie door de doelen van de auteur te analyseren en de inhoud/aard van de bron.
5. Je schrijft een conclusie met een definitief, goed onderbouwd, antwoord, gebruikmakend van de bron of bronnen.

## Mondelinge toets/performances

### *Voordelen*

- Relatief gemakkelijk te maken
- Valide manier van toetsen voor sommige vakken

### *Nadelen*

- Moeilijk om een beoordelingsmodel te maken
- Moeilijk om nauwkeurig, consistent en betrouwbaar te beoordelen
- Kost vaak veel tijd

## Opdrachten met bronnen

### **Inleiding**

Opdrachten met bronnen vragen veel van leerlingen, zeker als daarin een beroep gedaan wordt op hogere denkvaardigheden. Leerlingen moeten algemene vaardigheden bezitten, maar ook historische vaardigheden en metacognitieve vaardigheden. Uiteraard moeten ze ook beschikken over (context) kennis.

### **Algemene vaardigheden**

Informatie kunnen lezen en begrijpen:

- a. De leerling begrijpt de vragen en de bronnen.
2. Informatie selecteren uit gegeven contextrijke bronnen:
  - a. De leerling zoekt in de bronnen relevante elementen die het mogelijk maken een antwoord te geven op de vragen.
3. Informatie ordenen uit een veelheid van gegevens:

- a. De leerling rubriceert elementen uit de bronnen die informatie geven over respectievelijk de ene en de andere partij.
  - b. De leerling vergelijkt informatie uit de vraag (vraag 2) met informatie uit de bronnen.
4. Redeneervaardigheden:
- a. De leerling bouwt, als antwoord op de vraag (vraag 2) een redentatie op met argumenten ontleend aan de door hem geselecteerde elementen uit de bronnen.

### **Historische vaardigheden**

1. De leerling redeneert over en met bronnen
2. De leerling herkent in de bron (en de toelichting) een interpretatie.
3. De leerling geeft zelf een andere mogelijke interpretatie op basis van gegevens uit de vraag en eigen kennis.

### **Metacognitieve vaardigheden**

1. Oriënteren op de opdracht:
  - a. De leerling vraagt zich af waarop hij moet letten bij het analyseren van de bronnen om antwoord te kunnen geven op de vragen
  - b. De leerling weet wat er van hem verwacht wordt als hij een beredeneerd antwoord moet geven.
2. Proces bewaken
  - a. De leerling houdt tijdens de uitvoering van de taak in de gaten of het proces volgens plan verloopt (monitoring)
3. Bijstellen
  - a. De leerling vraagt zich, werkende aan de opdracht, geregeld af of de door hem geselecteerde elementen inderdaad informatie die relevant is voor het beantwoorden van de vragen.
  - b. De leerling stelt, werkend aan de opdracht, zo nodig zijn criteria bij.
4. Evalueren
  - a. De leerling vraagt zich af of het door hem gegeven antwoord voldoet aan de eisen die in de vragen gesteld zijn.
5. Reflecteren
  - a. De leerling vraagt zich af of de door hem gekozen werkwijze heeft geleid tot een bevredigend antwoord.

### **Kennis**

1. De leerling heeft contextkennis van de kerstening van West-Europa, in casu de Friezen (niet geheel duidelijk is in hoeverre deze kennis daadwerkelijk voorwaardelijk is om de vragen te kunnen beantwoorden).
2. De leerling kent begrippen die voor essentieel zijn voor het beantwoorden van de vragen: heiligenleven, missionaris, monnik, bekeerde, bijgeloof, goddeloze, vrome, beulsknechten, tierende, Heilige Schrift, martelaarskroon, heidenen, heilbrengende, interpretatie, partijdigheid, beredeneren.

3. De leerling weet dat een historisch verhaal (narratio) altijd een interpretatie is.

#### Voorbeelden van opdrachten met bronnen

Bij opdrachten met bronnen wordt soms gezegd 'lees de bron' of 'bekijk de bron'. Beter is het om 'gebruik de bron' te gebruiken. Deze formulering maakt beter duidelijk wat er van de leerlingen wordt verwacht. Toch is ook deze formulering niet altijd even eenduidig en is een nadere specificatie gewenst. Moet de leerling in zijn antwoord bijvoorbeeld duidelijk maken dat hij de bron heeft begrepen of moet hij een letterlijk citaat gebruiken?

#### *Voorbeeld 1: Tekstbronnen*

##### **Bron**

*De Engelse beroepsmilitair Walter Morgan was van 1572 tot 1574 kapitein in de legers van Willem van Oranje. Hij hield een dagboek bij. Daarin beschrijft hij hoe de prins van Oranje met zijn leger bij de stad Roermond aankomt en probeert de burgers van de stad over te halen om zich bij hem aan te sluiten*

Aan het slot van de dramatische redevoering (van Willem van Oranje) moesten de burgers van Roermond besluiten wat zij zouden doen. Zij kwamen tot het standpunt dat de prins volkomen terecht de wapens tegen de koning had opgenomen om te vechten voor de privileges van het land en voor zijn leven en zijn bezit. Maar zij moesten toch zijn eis naast zich neerleggen, want zij waren bang voor de wreedheden die het grote leger van de koning zou begaan, zoals dat ook in Napels en andere plaatsen was gebeurd. Bovendien betwijfelden zij of de prins en de opstandelingen sterk genoeg zouden zijn om tegen het kolossale leger van de koning weerstand te bieden. Na zijn vergeefse verzoek beval de prins dat er geweld moest worden gebruikt. Enkele kolonels uit zijn regiment zorgden ervoor dat direct ladders en andere benodigde zaken werden aangevoerd en zij brachten enkele stukken geschut in stelling.

Toen stormden de soldaten onder dekking van het geschutsvuur met de kreet 'alle hens aan de ladders' op de muren aan. De arme burgers hadden helemaal geen ervaring met het gebruik van vuurwapens in de strijd en vluchtten van de muren weg. Omdat zij bovendien niet over ervaren soldaten beschikten, die hen aanwijzingen konden geven voor de verdediging, bleek het niet moeilijk de stad zonder veel bloedvergieten te veroveren. Onder de burgerij waren bij de verovering 24 slachtoffers gevallen.

#### *Gebruik de bron*

Sommige historici noemen het conflict in de Nederlanden dat resulteerde in de stichting van een Nederlandse staat (de Tachtigjarige Oorlog):

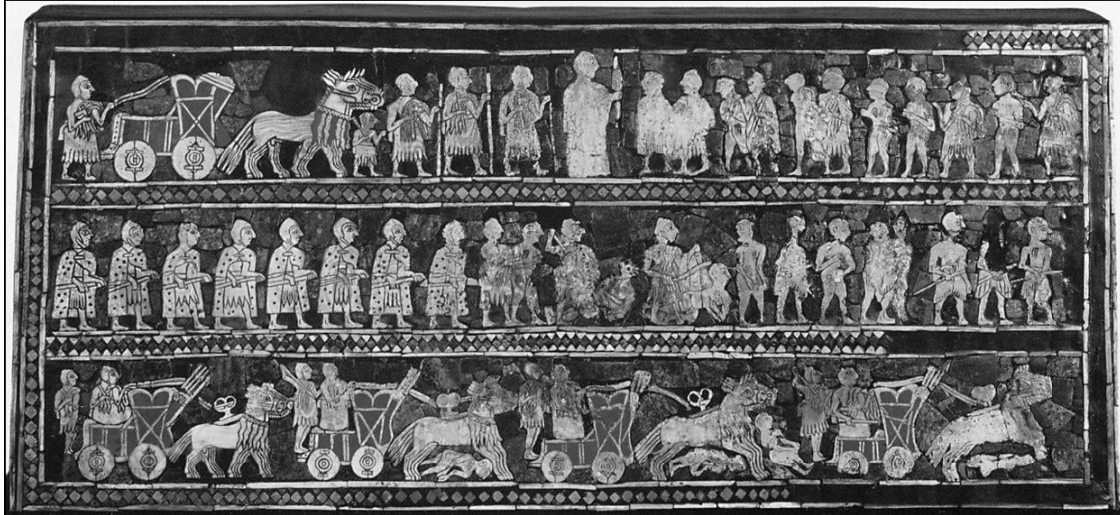
- een burgeroorlog of
- een opstand tegen de politiek van centralisatie van de koning van Spanje in de Nederlanden.

4p Ondersteun beide opvattingen met een argument uit de bron.

## Voorbeeld 2: Afbeeldingen

### Bron

Afbeelding van een Soemerisch leger op een houten voorwerp gevonden in een graf bij Oer in het tegenwoordige Irak, omstreeks 2600 v. Chr.



### Toelichting

Van boven naar beneden is te zien dat aan de koning een parade voorbijtrekt van zijn leger met krijgsgevangenen, speerdragers en andere soldaten in (met koperen plaatjes) gepantserde mantels en door paarden of ezels getrokken strijdwagens. Op de grond liggen gesneuvelde soldaten.

Uit deze bron kun je afleiden dat de bevolking van Oer omstreeks 2600 v. Chr. in een stedelijke gemeenschap leeft.

4p Leg uit, telkens met een gegeven uit de bron, dat er in die tijd rond Oer een stedelijke gemeenschap bestaat

- in economisch en
- in sociaal opzicht.

## Voorbeeld 3: Tabellen

### Bron

Een tabel:



Overzicht van katoenproductie en aantal slaven in de Verenigde Staten 1790 – 1840

<i>Jaar</i>	<i>1790</i>	<i>1820</i>	<i>1840</i>
<i>Aantal balen katoen</i>	<i>4.000</i>	<i>73.000</i>	<i>1.400.000</i>
<i>Aantal slaven</i>	<i>700.000</i>	<i>1.500.000</i>	<i>2.500.000</i>

*Gebruik de bron*

Een bewering:

Met deze gegevens kun je het verband aangeven tussen de groei van het aantal slaven in de Verenigde Staten en de Industriële Revolutie.

2p Leg dit verband uit.

*Voorbeeld 4: Cartoon en tekst*

**Bron 1**

*Een gedeelte uit een beroemde toespraak van president Truman tot het Amerikaanse Congres op 12 maart 1947, waarin hij de Marshallhulp (financiële en economische steun voor de wederopbouw) aan West-Europa aankondigt*

Ik geloof dat onze hulp in eerste instantie economisch en financieel moet zijn, omdat dit essentieel is voor economische stabiliteit en ordelijke politieke processen. (...)

De zaden van totalitaire regimes worden gevoed door ellende en gebrek. Zij verbreiden zich en groeien in de slechte grond van armoede en strijd. Zij komen tot volle wasdom als de hoop van een volk op een beter leven gestorven is.

Wij moeten de hoop levend houden. De vrije volkeren van de wereld kijken naar ons voor steun bij het handhaven van hun vrijheden.

**Bron 2**

Een politieke prent uit de Amerikaanse Minneapolis Star uit 1947 met als ondertekening: "Gas geven dokter!"



Gebruik bron 1 en 2

In de prent geeft de tekenaar zijn mening over de toespraak van president Truman.

5p Wat is de mening van de tekenaar? Ondersteun je antwoord door twee elementen uit de prent te verbinden met twee elementen uit de toespraak.

### Examenvragen

Een overzicht van vragen die zijn ontleend aan het centraal examen:

#### Beredeneer

- Neem een standpunt in en geef argumenten voor je standpunt. Ga, indien mogelijk, ook in op tegenargumenten.

#### Gebruik de bron

- ondersteun de conclusie/stelling/opmerking met (een voorbeeld uit) de bron
- leg met de bron uit (wat wil je precies)
- Leg uit, telkens met een gegeven uit de bron,
- ontleen aan de bron een voorbeeld
- geef een redenering met een verwijzing naar de bron
- Aan deze bron zijn verschillende voorbeelden van romanisering te ontleen. Noem hiervan drie voorbeelden.
- Geef aan hoe dit uit de bron blijkt.

Het examen bestaat verschillende soorten vragen:

1. Vragen met bronnen
  - a. tekstbronnen
    - i. bruikbaarheid van bronnen

- ii. betrouwbaarheid van bronnen
  - iii. argumenten geven
  - iv. uitleg geven
  - v. kennis toepassen op bronnen
  - vi. conclusies onderbouwen
  - vii. aantonen met behulp van de bron
  - viii. lezen van een tabel
  - ix. representativiteit beoordelen
  - x. continuïteit/discontinuïteit
  - xi. oorzaak/gevolg
- b. beeldbronnen
- i. foto's/tekeningen/schilderijen
    - 1. uitleggen met gebruik van elementen uit de foto
    - 2. aan de foto argumenten ontleen
  - ii. cartoons (zie ook bijlage 2)
    - 1. combineren van eigen kennis met informatie uit cartoon
    - 2. interpretatie van tekenaar herkennen
  - iii. tabellen
2. Vragen zonder bronnen
- a. vragen naar oorzaken/gevolgen
  - b. veranderingen
  - c. verbanden leggen
  - d. beredeneren
3. Chronologie vragen (een of twee)

## Formulering van opdrachten

### Schrijven van een vraag

- Bepaal welk onderdeel, kennis, begrip of vaardigheden je wilt toetsen
- Bepaal op welk toetsdoel de vraag betrekking heeft
- Kies een toepasselijk handelingswerkwoord
- Formuleer duidelijk en eenduidig
- Gebruik bij voorkeur geen woorden als sterk, nooit of altijd.
- Maak, zo nodig, een keuze uit bronnen en/of tabellen en grafieken
- Schrijf een eerste versie
- Formuleer een antwoord (eventueel ook mogelijke foute antwoorden)

Het schrijven van vragen is een iteratief proces

Vraag je het volgende af:

- Zal de leerling de vraag begrijpen?

- Geeft het handelingswerkwoord precies weer wat je van de leerling verwacht? Zie voor een overzicht van handelingswerkwoorden bij de taxonomie van Bloom bijlage 1.
- Heb je alle informatie gegeven die nodig is voor het kunnen beantwoorden van de vraag?
- Is de formulering helder en eenduidig?
- Weet de leerling wat voor soort antwoord hij moet geven?
- Kan niets de leerling afleiden van een goed antwoord?
- Is het beoordelingsmodel nauwkeurig genoeg?

### Onbedoelde effecten (bias)

Twee zaken kunnen een onbedoeld effect hebben op het kunnen beantwoorden van een vraag:

- Item content bias (soms kan iets van invloed zijn op het kunnen beantwoorden van een vraag, terwijl dat niet de bedoeling is)
- Statistical item bias (zaken als de sekse van iemand kan van invloed zijn op de prestaties; bijvoorbeeld het onderwerp van een leestekst, of bij geschiedenis het geloof van iemand, of de culturele achtergrond)

### *Oorzaken van onbedoelde effecten:*

- De vorm (jongens hebben meer met gesloten vragen dan meisjes)
- Inhoud
- De sociale context (jongens houden minder van bepaalde onderwerpen, zoals relaties of zorgen voor)
- Verwachtingen (verwachting van een leraar kunnen werken als self-fulfilling prophecies)
- Beoordelen (de volgorde waarin je het werk nakijkt kan invloed hebben)

Vermijd:

- opdrachten die iets anders toetsen dan je wilt toetsen (construct-irrelevant variance), bijvoorbeeld kennis van een vreemde taal of lees- en schrijfvaardigheden bij geschiedenis.
- dat opdrachten onvoldoende representatief zijn

### *Taalkundige aspecten*

- Zorg voor een eenduidige vraagstelling
- Vermijd waar mogelijk de lijdende vorm
- Vermijd waar mogelijk samengestelde zinnen
- Vermijd dubbele ontkenningen
- Gebruik geen overbodige woorden
- Vermijd tangconstructies

### Checklist voor het evalueren van toetsvragen

1. Wat wil je met de vraag meten? (feitenkennis, analyse vaardigheden, probleemoplossende vaardigheden enz.)
2. Wat wordt er feitelijk gemeten? (ligt de nadruk bijvoorbeeld te veel op feitenkennis?)
3. Wordt er gemeten wat we wilden meten? (Is de toets in lijn met wat we in de lessen hebben onderwezen, met de inhoud en met de beoogde leerresultaten?)
4. Zijn de vragen in lijn met de leerdoelen/eindtermen?
5. Is de taal eenvoudig, duidelijk en eenduidig?
6. Zijn de gebruikte handelingswerkwoorden duidelijk en eenduidig?
7. Is de taal die wordt gebruikt gemakkelijk te begrijpen, ook voor leerlingen voor wie het Nederlands niet de eerste taal is? (Gebruik gewone omgangstaal)
8. Controleer interpunctie en grammatica, want dit kan de betekenis van zinnen behoorlijk veranderen.
9. Worden bepaalde leerlingen door de vragen bevoordeeld of benadeeld (bijvoorbeeld als gevolg van godsdienstige overtuiging of etnische afkomst)
10. Kunnen de antwoorden betrouwbaar worden beoordeeld?
11. Zijn de vragen evenwichtig verdeeld over de leerstof?
12. Kan de vraag worden afgerond in de beschikbare tijd, ook door leerlingen voor wie het Nederlands niet de moedertaal is?
13. Maakt de toets voldoende onderscheid tussen goede en slechte leerlingen? Met andere woorden kunnen leerlingen laten zien wat ze kennen en kunnen?

### De psychologie van het beantwoorden van vragen

Als je weet wat er zich in de hersenen van leerlingen afspeelt bij het beantwoorden van vragen kun je daarmee rekening houden bij het schrijven van vragen.

#### *Fasering*

- Fase 0 het bestuderen van de leerstof
- Fase 1 het lezen van de vraag
- Fase 2 het verkennen van het geheugen
- Fase 3 het combineren van opgeroepen kennis met de vraag
- Fase 4 het bedenken van een antwoord
- Fase 5 het schrijven van een antwoord

#### Lezen

Lezen is een actief proces dat deels plaats vindt op een onbewust niveau. Factoren die invloed hebben op 'leesspanning':

- Het al dan niet begrijpen van een tekst.
- Leerlingen zien woorden als aanwijzingen, niet als bouwstenen
- Leerlingen maken gevolgtrekkingen en interpretaties
- Leerlingen zien/lezen wat zij verwachten te zien (voorkennis is dominant over nieuwe informatie)

De belangrijkste woorden of begrippen in vragen kunnen veel verschillende ideeën activeren. Sommige hiervan zullen van belang (relevant) zijn, andere niet. Leerlingen moeten de relevante begrippen selecteren en daar iets meedoen en de andere links laten liggen.

## Beoordelingsmodellen

Er zijn drie soorten beoordelingsmodellen:

- Voorschrijvend (goed of fout)
- Richtinggevend (een voorbeeld van een goed antwoord is ....)
- Beschrijvend (stap voor stap beschrijving van de criteria waaraan het antwoord moet voldoen)

### *Tips voor beoordelingsmodellen*

- Maak het beoordelingsmodel gelijktijdig met de vragen.
- Beoordeel goede prestaties.
- Geef in de toets aan hoeveel punten leerlingen kunnen krijgen voor bepaalde onderdelen.
- Vermijd stapelvragen.
- Geef aan wat goede en foute antwoorden zijn.
- Model moet aansluiten bij het vereiste beheersingsniveau.

## Verbeteren van toetsen

- Vraag je telkens af wat het doel van de toets is.
- Vraag je af telkens af of deze vorm van toetsen geschikt is voor het doel.
- Vraag je af wat de gevolgen van de toets zouden moeten zijn voor onderwijzen en leren.

## Standaardisatie

Standaardiseren gaat over het vaststellen van een normeringsmodel dat door iedere beoordelaar wordt begrepen en op gelijke wijze wordt toegepast.

### Standaard procedure

- Beoordelaars beoordelen een deel van het werk van de leerlingen
- Vervolgens komen de beoordelaars bijeen om het definitieve normeringsmodel vast te stellen.
- Pas hierna wordt het werk van de leerlingen beoordeeld.

## Moderation

- Moderation betekent dat het werk van de beoordelaar beoordeeld door een extern persoon (in Nederland bijvoorbeeld de tweede corrector).
- Gebeurt meestal op basis van steekproeven.
- Bij vakken waarbij er sprake is van bijvoorbeeld optredens van leerlingen is moderation op locatie gewenst.
- Je kunt ook statistieken gebruik voor moderation

- Moderation moet gebruikt worden om cijfers aan te passen, om het leerproces en de toets te evalueren

Figure 8 The Five Phases of Moderation

