



René Westra

Een inkijkje in de toets

De ins and outs van het maken, gebruiken, beoordelen en evalueren van toetsen in het onderwijs, met een aantal 'veelkleurige vragen' voor de onderbouw en bovenbouw vmbo

Inhoud

Inleiding

Hoofdstuk 1 Het samenstellen van een toets

Hoofdstuk 2 Het maken van vragen

Hoofdstuk 3 Digitale toetsvormen

Hoofdstuk 4 Adaptieve toetsvormen

Hoofdstuk 5 Toetscontrole en afronding

Hoofdstuk 6 Toetsafname en correctie

Hoofdstuk 7 Toetsanalyse

Hoofdstuk 8 Cesuurbepaling en cijfers

Hoofdstuk 9 Feedback op de toets

Hoofdstuk 10 Kritiek op toetsen

Hoofdstuk 11 Een serie veelkleurige vragen voor de onderbouw
vmbo/havo/vwo

“De schrijver moet zijn beloning zoeken in zijn plezier in het werk en in het zich bevrijden van de last van zijn gedachten en moet zich, onverschillig tegenover al het andere, niets aantrekken van lof of blaam, mislukking of welslagen.”

(uit ‘The moon and sixpence’ van William Somerset Maugham)

Inleiding

Een aantal jaren geleden wilden enkele leerlingen graag “een inkijkje in de toets” hebben van het centraal schriftelijke eindexamen Frans en pleegden een inbraak in de kluis van hun school.

Maar ik bedoel met “een inkijkje in de toets” iets anders. In mijn lange carrière in het voortgezet onderwijs heb ik me heel vaak beziggehouden met het verschijnsel toetsen: vragen maken, toetsen samenstellen, surveilleren en daarna analyseren en nabespreken van toetsen.

Vroeger als leerling moest ik natuurlijk ook zelf regelmatig een toets (toen proefwerk genoemd) maken. Zo'n proefwerk bestond in die tijd altijd uit een serie open vragen. De eerste keer dat ik een meerkeuzetoets maakte was in 1968, in het tweede jaar van mijn biologiestudie. Het resultaat was bar slecht: 3,95. Daarna kwam ik veel vaker meerkeuzevragen tegen en las ook het boek “Vijven en zessen. cijfers en beslissingen: het selectieproces in ons onderwijs” uit 1966, van de wiskundige en psycholoog professor A.D. de Groot. Daarin legt hij onder andere uit dat meerkeuzevragen bij uitstek geschikt zijn om objectief kennisonderdelen te meten. Hij was betrokken bij de oprichting van het Cito en de ontwikkeling van de Citotoets, die aan het eind van het onderwijs op de basisschool heel frequent wordt gebruikt.

Dat betekende overigens niet dat een toets sinds zijn boek alleen uit meerkeuzevragen moest bestaan, omdat bijvoorbeeld het uitvoeren van een berekening of een argumentatie vaak beter in een open vraag kan worden getoetst. Toch bestonden in mijn vak, de biologie, de centraal schriftelijke examens van 1975 tot 1989 uitsluitend uit een serie van rond de veertig meerkeuzevragen. Op de school waar ik werkte als docent biologie maakten we, vaak samen, toetsen voor gemeenschappelijk gebruik bij een proefwerk in bijvoorbeeld alle tweede klassen. Ik hield ervan om vragen te ontwikkelen, terwijl een van mijn collega's juist heel precies was in de controle van al mijn vragen. Zo'n duo is heel belangrijk, vragen maken moet je nooit alleen doen. Iedereen heeft een dode hoek of blinde vlek. Je tegenlezer bewaart je voor een wat onduidelijke formulering of een vaktechnisch foutje.

We probeerden ook die toetsen te evalueren. Zijn alle vragen goed of leveren sommige ondanks onze controle toch verwarring op bij de leerlingen? Moeten er misschien vragen achteraf verwijderd worden uit de toets? En hoe komen we uit de verzamelde resultaten tot een cijfer? Ik vond het boeiend om te kijken hoe de verschillende vragen scoorden. Ik leerde al snel dat er technieken waren om dat te onderzoeken en legde de resultaten vast in een steeds verder uitdijend kaartsysteem en later in een database op mijn computer.

Twee jaar nadat ik begonnen was als leraar, werd mij gevraagd naast mijn werk een cursus te geven voor een derdegraads opleiding voor mavo/lbo-docenten biologie. Per thema werd daarbij een meerkeuzetoets afgenomen van vijftig vragen. De cursusleiders werd gevraagd ieder twintig vragen te maken, uit de verzamelde vragen werd de toets samengesteld. Het bleek dat in iedere toets een flink aantal van de door mij gemaakte vragen was geselecteerd. Zo kwam ik erachter dat ik op dit terrein een redelijke vaardigheid bezat.

Ik werkte later in mijn carrière een aantal jaren bij het Cito, waar ik me in een team bezig hield met het maken van vragen voor het centraal schriftelijke examen biologie voor havo en voor de eindtoets van de basisschool voor het onderdeel Natuur en Techniek.

Mijn vaardigheid als vragen- en toetsconstructeur werd nog verder ontwikkeld, toen ik daarna ging helpen bij het maken van vragen voor de Biologie Olympiade en bij de uitgevers Malmberg en Noordhoff werd aangesteld als toetsspecialist voor de methodes Biologie voor Jou en Nectar en bij de Vereniging van Hogescholen aan de slag ging bij de constructie van de Kennisbasistoets voor biologie. In het derde jaar van de tweedegraads lerarenopleiding doen alle studenten een digitale toets die bestaat uit honderd vragen, voor het overgrote deel meerkeuze. Zij moeten die toets met een voldoende afronden om verder te kunnen met hun opleiding. Bij de staatsexamens voortgezet onderwijs werkte ik in een team aan de ontwikkeling van schriftelijke toetsen voor twee vakken waarvoor geen schriftelijke eindexamen bestaat: ANW (Algemene Natuurwetenschappen) en NLT (Natuur, Leven en Technologie). Uiteindelijk heb ik in mijn werkzame leven tot nu toe zo'n 25.000 vragen ontwikkeld. Die verzameling bestaat

niet uitsluitend uit zelf bedachte vragen, maar ook uit door mij vertaalde vragen die ik op het internet vond uit landen zoals het Verenigd Koninkrijk, België, Canada, Australië, Denemarken, Zweden, Italië en Turkije. Leerlingen in Nederland zoeken vaak wel het internet af naar oude examenvragen die ook regelmatig door leraren worden gebruikt, maar kijken niet naar buitenlandse bronnen.

Ik kan me voorstellen dat lang niet elke docent zo'n toetsfanaat is als ik ben en dat werk liever (geheel of gedeeltelijk) aan experts overlaat. Maar het gaat bij het maken van vragen en toetsen niet alleen om techniek en creativiteit.

Er dienen door elke docent ook meer wezenlijke vragen beantwoord te worden. Wat is het doel van de afgenomen toets? Gaat het om het bepalen van een cijfer (summatieve toets), om te checken of de leerlingen de stof begrijpen (formatieve of diagnostische toets) of om op te sporen en vervolgens te helpen verbeteren waar individuele leerlingen te kort schieten (adaptieve toets)? En als de score van een summatieve toets eenmaal duidelijk is, kijk je dan per leerling of per groep? In het laatste geval ontstaat een soort ranglijst, waarbij je een cesuur moet bepalen. Welke score hoort bij de grens tussen voldoende en onvoldoende?

Wil je dat de leerlingen parate kennis demonstreren, geleerde zaken kunnen toepassen of inzichtvragen kunnen oplossen, of een combinatie van die drie categorieën? En in dat laatste geval, in welke verhouding dan?

Veel vraagcategorieën die in het onderwijs gebruikt worden om vragen te kenschetsen zijn gebaseerd op het werk van de Amerikaanse onderwijspsycholoog Benjamin Bloom. In zijn boek 'The Taxonomy of Educational Objectives: The Classification of Educational Goals' uit 1956 categoriseerde hij leerdoelen in een schema dat bekend is geworden als Bloom's Taxonomy. Het was een van de eerste pogingen om niveaus van cognitief functioneren systematisch te classificeren.

Vind je een mooi geformuleerd antwoord met een correcte zinsbouw bij een open vraag van belang, of accepteer je ook een slechte zin waar de kern van het gevraagde duidelijk uit wordt?

Kun je, nu de computer in het onderwijs steeds meer gebruikt wordt, digitaal toetsen en zo, ja wat zijn daarvan voor- en nadelen?

Moet je op veel of weinig momenten toetsen, met als consequentie weinig of veel stof per toetsmoment?

Kun je alles wat je behandelt eigenlijk wel toetsen en zo nee, hoe ga je dan daarmee om?

Veel vragen die in dit boek allemaal aan de orde komen.

Hoofdstuk 1 Het samenstellen van een toets

Als je een toets over een bepaald onderwerp gaat maken, is het verstandig om je eerst bepaalde zaken af te vragen:

1. Welke leerdoelen wil ik toetsen?
2. Zijn alle leerdoelen even belangrijk of moeten bepaalde doelen extra gewicht krijgen?
3. Gaat het om een toets waarin alleen pure kennis aan de orde komt of betreft het ook bijvoorbeeld toepassing van geleerde kennis en inzicht in de bestudeerde stof?
4. Welke vraagvormen wil ik gebruiken; alleen open vragen, alleen gesloten vragen (bijvoorbeeld juist-onjuist, meerkeuze, matching) of een combinatie van beide?
5. Is de toets bedoeld om het bereikte niveau van je leerlingen of studenten te bepalen (summatief), om hen te laten onderzoeken of ze de stof voldoende beheersen (formatief, diagnostisch), of om hen met individuele feedback bij te sturen in hun leerproces (adaptief)?

Als je deze vragen hebt beantwoord, maak je eerst een matrijs waarin je een overzicht krijgt over je toets, zodat je kunt zien wat je in je vragen allemaal wilt onderbrengen. Hieronder zie je een voorbeeld van een matrijs voor een toets die bestaat uit 30 vragen, waarvan 7 open vragen, 5 juist-onjuistvragen en 18 meerkeuzevragen.

Er zijn 10 leerdoelen, waarvan leerdoel 5 en 7 twee keer zo belangrijk zijn dan de andere doelen en daarom extra punten krijgen. Dat kan door over die twee doelen vragen te maken met een zwaarder gewicht en of er meer vragen over op te nemen.

In totaal zijn er 60 punten te behalen.

De verdeling kennis : inzicht : toepassing in punten is hier 10% : 30%: 60%.

Vraagnummer	Leerdoel	Type vraag	K:I:T	Punten
1	1	juist-onjuist	K	1

2	2	juist-onjuist	K	1
3	8	juist-onjuist	T	1
4	9	juist-onjuist	T	1
5	3	juist-onjuist	T	1
6	10	open	I	2
7	5	meerkeuze	K	2
8	9	meerkeuze	T	2
9	7	meerkeuze	T	2
10	1	meerkeuze	I	2
11	7	meerkeuze	T	2
12	1	meerkeuze	T	2
13	6	meerkeuze	T	2
14	6	meerkeuze	T	2
15	7	meerkeuze	T	2
16	9	meerkeuze	I	2
17	4	meerkeuze	T	2
18	8	meerkeuze	T	2
19	4	meerkeuze	T	2
20	4	meerkeuze	I	2
21	3	meerkeuze	I	2
22	2	meerkeuze	T	2
23	2	meerkeuze	T	2
24	3	meerkeuze	2	2
25	5	open	I	2
26	6	open	T	2
27	5	open	T	3
28	8	open	I	2
29	10	open	I	3
30	10	open	T	4
				Totaal 60

Zo'n matrijs kan ook gelden voor een diagnostische toets. Dat is nuttig, doordat leerlingen bij het maken van die toets al enig idee krijgen van wat hen te wachten staat bij de toets waarop

ze beoordeeld worden. Het is sowieso goed om leerlingen van tevoren in te lichten over de samenstelling van de toets die ze gaan maken. Het is voor hen belangrijk dat ze weet hebben van de vorm: zijn er alleen open vragen, alleen meerkeuzevragen of is er een combinatie van vraagtypen? Het is verstandig om als eerste vraag altijd met iets redelijk eenvoudigs te beginnen. Sommige leerlingen blijven namelijk lang hangen bij de eerste vraag als ze die niet direct weten, worden zenuwachtig en verspelen veel kostbare tijd. Dat voorkom je door eenvoudig te beginnen, dat werkt voor veel gespannen leerlingen geruststellend. Het is ook verstandig om de leerlingen aan te raden om eerst te kijken hoeveel vragen er precies zijn en eens door de hele toets heen te scrollen. Ook is het goed om hen te vertellen dat ze vrij zijn in de volgorde waarin ze de vragen maken, als ze maar goed bijhouden welke vragen ze al gedaan hebben.

Hoofdstuk 2 Het maken van vragen

Na het vastleggen van de contouren van de toets met behulp van de matrijs komt nu de constructie van de vragen. Dat kost meestal meer tijd dan je denkt, als je je werk tenminste serieus aanpakt. Soms spreekt men hier van de wet van π : de benodigde tijd voor toetsconstructie is π x de tijd die je dacht nodig te hebben!

Bij mij begint die constructie vaak met iets wat ik zie of lees en raakt aan het onderwerp waarover

ik een vraag ga maken. Zo zat ik eens te kijken naar een aflevering uit een serie van Govert Schilling over de ontstaansgeschiedenis van de mens, terwijl op mijn takenlijst het onderwerp evolutie stond. Ik zag de jonge Nederlandse biologe Laura van Holstein, die werkt aan de universiteit van Cambridge UK, en hoorde haar vertellen over haar originele onderzoek. Dan wordt dat



onderzoek het skelet van een vraag. Maar daarmee begint het werk pas. Allereerst kijk je dan bij welk leerdoel de vraag die je in je hoofd hebt past. Vervolgens kies je voor het type vraag: bijvoorbeeld meerkeuze, juist/onjuist, matching of open. En je kiest voor de categorie: kennis, inzicht of toepassing. Er zijn vele varianten bij de indeling in categorieën. Bij het Cito gebruikten we er in mijn tijd twee: R (Reproductie) en P (Productie). In het hedendaagse voortgezet onderwijs is er een veel gebruikte variant met vier categorieën: R (Reproductie), T_1 (Toepassing in een voor de leerlingen bekende context), T_2 (Toepassing in een onbekende context) en I (Inzicht). Dat kan echter alleen als je de toets maakt voor de jou bekende leerlingen. Dan weet je immers welke contexten zij kennen. Werk je met onbekende kandidaten, zoals bij het afnemen van mondelinge staatsexamens of bij het maken van vragen voor

het centraal schriftelijk examen, dan weet je niet welke contexten de leerling kent, doordat er meerdere van elkaar verschillende lesmethodes door hen kunnen worden gebruikt. Voor alle gemaakte vragen geldt de zogenaamde Rosé Kwaliteits Matrijs, waarbij op alle vragen Ja dient te worden geantwoord.

Kwaliteitscriterium	Vraag	Ja	Nee
Relevant	Is het zinvol om de vraag te stellen?		
	Heeft de vraag een relatie met het toetsdoel?		
	Sluit de vraag aan op het taxonomie-niveau?		
Objectief	Zijn deskundigen het eens over het goede antwoord?		
	Wordt er, zonder dat daar een duidelijke reden voor is, gevraagd naar de mening van een leerling?		
Specifiek	Kan alleen een leerling die de stof beheerst de vraag beantwoorden?		
	Bevat de vraag <u>geen</u> (grammaticale) aanwijzingen in de richting van het juiste antwoord?		
Efficiënt	Bevat de vraag <u>geen</u> overbodige informatie/informatie die afleidt bij het kiezen van het antwoord?		
	Is de vraag kort en bondig geformuleerd?		

Meerkeuzevragen en juist-onjuistvragen

Als het een meerkeuzevraag moet worden, begin je met een korte tekst (de stam). Daarna volgt een vraagzin en vervolgens plaats je behalve het eigenlijke antwoord op je vraag twee of drie plausibele zogenaamde afleiders. Dat kan heel lastig zijn. Leerlingen zijn soms heel handig in het herkennen van afleiders, zij bezitten ‘testwijsheid’. Zo is uit onderzoek van een

groot aantal vragen gebleken dat het voor een leerling die geen flauw idee heeft (doordat hij of zij het leerboek niet van buiten kent, maar alleen van buiten kent) handig is om bij een vierkeuzevraag C te kiezen. Om tegen te gaan dat C significant vaak juist is, kun je de drie of vier mogelijkheden in alfabetische of chronologische volgorde opvoeren. Wat je zeker moet vermijden is een afleider als “Geen van de antwoorden is juist.” Ook ontkennende vraagzinnen kun je beter vermijden, door het gebruik van het woord ‘niet’ ontstaat vaak verwarring. Soms gebruikte ik veel voorkomende foute antwoorden (misconcepten) van open vragen als afleiders voor een meerkeuzevraag over hetzelfde onderwerp.

Als je echt geen derde afleider kunt verzinnen zonder te vervallen in aperte onzin, kun je beter een driekeuzevraag kiezen.

Gebruik ook geen begrippen als ‘altijd’ en ‘nooit’. Zulke stellige beweringen zullen door een slimme leerling niet als juist gekozen worden en dat is vaak terecht.

Voor de duidelijkheid is het goed om na de stam een vraagzin te formuleren, ook al kost dat iets meer woorden. De leerling kan vast nadenken over het probleem en hoeft niet eerst alle antwoorden door te lezen.

Dus niet:

Antoni van Leeuwenhoek is een beroemde onderzoeker uit de 17^e eeuw.

2p Hij

- A** ontdekte bacteriën in een peperoplossing.
- B** ontwikkelde de eerste microscoop.
- C** verbeterde het anatomisch onderzoek aan de mens.
- D** zorgde voor een doorbraak in de genetica.

Beter is :

Antoni van Leeuwenhoek is een beroemde onderzoeker uit de 17^e eeuw.

2p Waardoor werd hij zo beroemd?

- A** Hij ontdekte bacteriën in een peperoplossing.
- B** Hij ontwikkelde de eerste microscoop.

- C** Hij verbeterde het anatomisch onderzoek aan de mens.
- D** Hij zorgde voor een doorbraak in de genetica.

Soms zie je een vraag met twee stellingen of beweringen, bijvoorbeeld:

Tarik maakt een profielwerkstuk over skeletspieren en hun kracht.

Hij hangt een gewichtje via een touwtje aan een vinger van zijn proefpersonen en vraagt hen de vinger gestrekt te houden.

Vervolgens hangt hij steeds zwaardere gewichtjes aan het touwtje en stelt vast bij welk gewicht de vinger gaat buigen.

Volgens Yvette, een van zijn proefpersonen, trekken de spiervezels die de vinger gestrekt houden bij toenemend gewicht steeds sterker samen.

Volgens Marjo, een andere proefpersoon, zijn steeds meer motorische eenheden actief om de vinger gestrekt te houden bij toenemend gewicht.

2p Wie heeft of wie hebben gelijk?

- A** geen van beiden
- B** alleen Yvette
- C** alleen Marjo
- D** zowel Yvette als Marjo

Het bezwaar tegen deze formulering is dat een leerling die een van beide beweringen goed heeft beantwoord, geen punten krijgt.

Beter is in zo'n geval een splitsing in twee juist/onjuistvragen.

Tarik maakt een profielwerkstuk over skeletspieren en hun kracht.

Hij hangt een gewichtje via een touwtje aan een vinger van zijn proefpersonen en vraagt hen de vinger gestrekt te houden.

Vervolgens hangt hij steeds zwaardere gewichtjes aan het touwtje en stelt vast bij welk gewicht de vinger gaat buigen.

1p Volgens Yvette, een van zijn proefpersonen, trekken de spiervezels die de vinger gestrekt houden bij toenemend gewicht steeds sterker samen.

Geef aan juist of onjuist.

1p Volgens Marjo, een andere proefpersoon, zijn steeds meer motorische eenheden actief om de vinger gestrekt te houden bij toenemend gewicht.
Geef aan juist of onjuist.

De tekst moet zo transparant mogelijk zijn, zodat er geen misverstanden kunnen ontstaan. De vraag moet relevant zijn, het bijbehorende leerdoel moet echt aan de orde komen. Over de lengte van de tekst lopen de meningen erg uiteen. In veel vakliteratuur wordt de nadruk gelegd op compactheid, dus een zo kort mogelijke tekst.

Maar bijvoorbeeld in de natuurwetenschappelijke vakken in het voortgezet onderwijs werkt men sinds 2013 met de concept-contextmethode, die men dan ook in het examen tot uitdrukking wil brengen. Dit leidt vaak tot vrij lange teksten die dan wel bij meerdere vragen (in clustervorm) aan de orde komen.

Persoonlijk prefereer ik korte teksten, zodat de wat minder taalvaardige leerlingen geen problemen ondervinden. Wel gebruik ik vaak een afbeelding. Mensen zijn vaak visueel ingesteld, dat kan hen mijns inziens helpen om zich in het probleem in te leven. Een afbeelding zegt ook vaak meer dan een lap tekst. Ik herinner me ook dat ik brugklassen les gaf waarin een aantal kinderen geen idee hadden hoe een paardenbloem of een meerkoet eruit zag. Door een afbeelding aan een vraag over deze organismen toe te voegen, werd dit voor hen duidelijk.

Meerkeuzevragen krijgen vaak standaard 2 punten en juist-onjuistvragen 1 punt.

Juist-onjuistvragen zijn eigenlijk ook meerkeuzevragen, met slechts één afleider.

Overigens is het niet zo dat je met meerkeuzevragen alleen de categorie kennis kunt afvragen, zoals veel mensen denken.

Met enige creativiteit kun je prima ook toepassingsvaardigheden en inzicht toetsen met meerkeuzevragen.

Een nadeel ten opzichte van open vragen is dat het maken van goede afleiders relatief veel tijd kost. Daar staat tegenover dat het nakijken sneller gaat en dat je gemakkelijker veel leerdoelen kunt toetsen, doordat de beantwoording van een meerkeuzevraag de leerlingen minder tijd kost. Je kunt dus in een toets meer vragen van dit type stellen.

Matchingvragen

Bij matchingvragen krijgt de leerling twee rijtjes in alfabetische of chronologische volgorde aangeboden, met de opdracht de juiste begrippen uit de beide rijtjes te koppelen.

Het is handig om in een van beide rijtjes een extra begrip op te nemen. Dat voorkomt dat een fout automatisch leidt tot een tweede fout.

Een voorbeeld uit de anatomie van de mens.

2p In het geraamte komen veel beenderen voor. Zet de vier beenderen bij het juiste lichaamsdeel. Er blijft een lichaamsdeel over!

Ellepijp	Arm
Lendenwervel	Bovenbeen
Scheenbeen	Hoofd
Slaapbeen	Onderbeen
	Romp

En een voorbeeld uit de geschiedenis.

2p Op verschillende momenten hebben in de Nederlandse geschiedenis belangrijke gebeurtenissen plaatsgevonden. Zet vier van de vijf jaartallen bij de juiste gebeurtenis. Er blijft een jaartal over!

Jaartallen	Gebeurtenissen
1568	Einde van de tweede wereldoorlog
1672	Moord op de gebroeders de Witt
1739	Nederland wordt een koninkrijk
1815	Start van de 80-jarige oorlog
1945	

Open vragen

Bij open vragen (ook wel essayvragen genoemd) vraagt de formulering een nog grotere nauwkeurigheid dan bij de bovenstaande vraagtypes. Maak heel duidelijk wat je verwacht in het antwoord. Als ja of nee niet volstaat, geef dan aan dat je meer wilt zien, bijvoorbeeld door de formulering “Leg je antwoord uit” of “Geef twee argumenten voor je keuze”. Wijs leerlingen er van tevoren op dat ze goed moeten lezen wat er precies gevraagd wordt. Dat voorkomt dat ze hele verhalen opschrijven en van jou verwachten dat je het goede eruit haalt. Je moet ze bijvoorbeeld leren dat als de vraag luidt “Noem drie zoogdieren uit het ecosysteem loofbos”, je niet van plan bent om uit een lijst van tien door hen genoemde dieren de zoogdieren eruit te zoeken.

Maak ook meteen een voorbeeldantwoord, met een duidelijke verdeling in punten. De leerling moet erop voorbereid zijn dat er een relatie is tussen het aantal punten en de stappen die in het antwoord aan de orde moeten komen. Vragen als “Schrijf alles wat je weet over ... op” met 4 punten als beloning leiden tot een enorme discussie bij de nabespreking.

Gebruik in vakken als biologie of scheikunde liefst zo weinig mogelijk formuleringen met het woord ‘waarom’ erin. Ik zag eens onderstaande vraag: “Waarom is een tomaat rood?”

Antwoorden als: “God heeft ze zo geschapen” zijn dan niet meteen fout. Het woord waarom vraagt naar een *reden*, terwijl de auteur van deze vraag waarschijnlijk geïnteresseerd was in de *oorzaak* (aanwezigheid van rode korrels in de cel) of in de *functie* (aantrekken van vogels in verband met zaadverspreiding). Dus moet je het vraagwoord ‘waarom’ en het voegwoord ‘omdat’ alleen gebruiken als je de reden van iets wilt weten. Gaat je interesse uit naar de oorzaak van iets (dat komt in de biologie veel vaker voor) gebruik dan ‘waardoor’ en ‘doordat’. Zelfs de om zijn denken bekend staande mens doet trouwens regelmatig dingen zonder dat daar een reden bij komt kijken.

Als voorbeeld de volgende 'foute vraag' die ik tegenkwam in een biologiemethode:

"Een man loost per zaaduitstorting miljoenen zaadcellen. Waarom doet hij dat, denk je?"

Dit veronderstelt dat een man bedenkt dat de kans op bevruchting niet zo groot is, dus produceert hij maar een flink aantal zaadcellen. Dat is natuurlijk nonsens, die man denkt hier helemaal niet over na.

De toevoeging 'denk je?' is trouwens helemaal dodelijk.

Als een leerling opschrijft: "Ik denk dat mannen nogal verspillend zijn", moet je het hele antwoord goed rekenen. "Dat dacht ik echt, meneer/mevrouw!"

Ook het volgende vraagtype dat ik ooit tegenkwam, leidt tot vreemde antwoorden die strikt genomen niet fout gerekend kunnen worden. Eerst werd een verschijnsel beschreven.

Daarna luidde de vraag als volgt:

2p Kun je voor bovengenoemd verschijnsel een verklaring geven? Beargumenteer je antwoord.

Stel dat een leerling nu schrijft:

"Dat kan ik niet. Het argument is dat ik daar te dom voor ben."

Nu heb je een probleem als corrector.

Ik heb jarenlang bij de staatsexamens gewerkt, waar wij onder andere autistische leerlingen moesten examineren, die bepaalde zinsneden vaak letterlijk opvatten. Op een cursus leerde ik dat de volgende vraag bij een examen wiskunde daardoor bij hen problemen kan opleveren.

2p Benoem de hieronder afgebeelde figuren.



Een leerling noteerde: "Roos, Bert, Inge, Hans."

Hoofdstuk 3 Digitale toetsvormen

Toetsvragen maken vond ik altijd leuk, maar een toets nakijken was minder prettig. Het kostte heel veel tijd en soms was het werk ook nog eens moeilijk leesbaar. Bij meerkeuzevragen was soms ook niet duidelijk of de leerling B of D had gekozen. En dan had je ook nog het probleem dat leerlingen soms bij de nabespreking dingen veranderden, waarbij er een discussie ontstond of jij iets verkeerd had gezien of dat zij bij de nabespreking zo'n verandering hadden aangebracht. Er waren op school collega's die daarom alle toetsen kopieerden. Toen de computer zijn intrede deed op school en er digitale toetsprogramma's als Wintoets (de voorloper van Quayn) ontwikkeld werden, ging ik aan de slag met het digitaliseren van alle vragen in mijn kaartsysteem. Dat was veel werk, maar die tijd werd teruggewonnen doordat het corrigeren veel sneller ging. In een programma als Wintoets worden meerkeuzevragen automatisch nagekeken, zodat je alleen de antwoorden op de open vragen (die nu goed leesbaar getypt waren) hoefde na te kijken. Ik voorzag iedere vraag van metadata zoals het niveau (bijvoorbeeld brugklas, 4 havo), onderwerp (bijvoorbeeld gedrag, evolutie, genetica) categorie (kennis, inzicht of toepassing) en oorsprong (bijvoorbeeld eigen werk, examenvraag, vertaalde vraag). Na hard werken had ik een grote verzameling, waarmee ik via een gerichte zoekopdracht een toets over een bepaald onderwerp (bijvoorbeeld voortplanting) voor een bepaald niveau (bijvoorbeeld 5vwo) met een gewenste verdeling kennis : inzicht : toepassing kon samenstellen.

Ook een herkansing, waar mijn collega's vaak over zuchtten vanwege het extra werk, was snel gefabriceerd. Ik herinner me dat ik een herkanser soms, terwijl ze naar huis fietste, via haar mobieltje de uitslag kon doorgeven, zo snel gaat de correctie! Als extra kon je in de toets nu ook foto's in kleur, geluidsfragmenten, videofragmenten, animaties en computermodellen opnemen.

Zo maakte ik toetsen voor de brugklas waarbij ik de soortenkennis van planten testte met dertig kleurenfoto's en die van vogels, waarbij de leerlingen foto's moesten matchen met bij die vogels passende zangfragmenten.

Mijn werk kreeg een boost toen het Ministerie van Onderwijs scholen zocht voor een digitaal pilotproject genaamd Compex. Zowel op havo (vanaf 2002) als op vmbo en vwo (vanaf 2003) werd die pilot voor verschillende vakken georganiseerd. Mijn collega en ik meldden ons aan voor het Compex-examen biologie. De rector van onze school zorgde voor een flink aantal prima computers, onze afdeling systeembeheer regelde met veel zorg alle benodigde instellingen, onze klusjesman zorgde voor schotten tussen de computers tegen afkijken en medewerkers van het Cito kwamen op bezoek om te kijken of alles goed functioneerde. De leerlingen van de aan die pilot deelnemende scholen kregen op hun centraal examen eerst een gedeelte met vragen die hetzelfde waren als die van de 'gewone' examens. Maar het tweede gedeelte van hun examen bestond uit vragen die alleen met behulp van de computer konden worden opgelost. Dat betrof vragen met aangeboden videofragmenten (bijvoorbeeld over diergedrag), animaties (bijvoorbeeld van de hartslag of van een reflex) en dynamische modellen. Met die modellen konden allerlei complexe berekeningen over de veranderingen die in het lichaam optreden tijdens inspanning of over de populatieontwikkelingen van vogels op een eiland worden uitgevoerd. Alle vragen werden door de leerlingen schriftelijk beantwoord. Negen jaar lang draaide die pilot prima. Er waren alleen kleine problemen. Zo liep een keer een computer vast, toen een leerling een deling door nul uitvoerde en was er een keer geen geluid bij een videofragment, maar met een uitgetypte tekst die dat geluidsfragment verving losten we dat op. De leerlingen vonden het leuk (minder tekst, meer beeld). Vooral de videofragmenten over bijvoorbeeld het gedrag van leeuwen, wippende en hokkende scholeksters met heel enthousiast commentaar van Bruno Ens en chimpansees in actie uit Bert Haanstra's film 'Chimps onder elkaar' zorgden voor blije reacties na afloop. De bedoeling was dat in 2008 Compex voor biologie (en enkele andere vakken) landelijk zou worden ingevoerd. Veel collega's zagen echter beren op de weg. "Stel dat de stroom uitvalt of dat een animatie niet draait, wat dan?" Zoals wel vaker kiest de overheid dan niet voor de enthousiaste voorlopers, maar voor de grote meute, die wil dat alles blijft zoals het altijd al geweest is. Zonder enige toelichting

werd de pilot plotseling gestopt, een grote teleurstelling en fnuikend voor je enthousiasme om nogmaals mee te doen aan een of ander volgend vernieuwingsproject. Overigens zie je dit terugtrekgedrag van de overheid veel vaker bij vakgerichte dan bij onderwijskundige vernieuwingen. Daar wordt vaak juist wel doorgedrukt, ook als veel leraren die vernieuwingen helemaal niet willen. Gelukkig is er voor de leerlingen van kader en basis op het vmbo wel doorgezet met het computerexamen. Het complete centrale examen bestaat voor hen nu uit het zogenaamde FACET-computerexamen, soms al vóór de schriftelijke examens (in april) georganiseerd voor meerdere vakken. Er worden verschillende versies ontwikkeld, waarbij de scholen het meest geschikte moment kunnen bepalen waarop het examen wordt afgenomen. De leerlingen zijn blij, het lezen van grote lappen tekst zoals bij hun collega's van GL, TL, HAVO en VWO, blijft hen bespaard! In mijn jaren bij het staatsexamen heb ik die FACET-examens een aantal keren bijgewoond. Voor het eerst van mijn leven heb ik kandidaten van basis en kader, zittend achter de computer, zien grinniken tijdens een examen. Toen ik na afloop vroeg naar hun ervaringen, vertelden ze dat ze sommige fragmenten, met name over diergedrag, heel leuk vonden. Ook waren ze blij met het gebruik van beeldfragmenten in plaats van lappen tekst om een bepaalde vraagcontext aan de orde te brengen.

Hoofdstuk 4 Adaptieve toetsvormen

Toetsen die meetellen voor de beoordeling (summatieve toetsen) kunnen worden voorbereid door leerlingen met eenzelfde type toetsen te confronteren, zodat ze kunnen vaststellen of ze voldoende voorbereid zijn. Zulke voorbereidingstoetsen worden diagnostische toetsen genoemd. Een andere vorm van voorbereiding kwam ik tegen in mijn werk bij Biologie voor Jou. Een groep biologiedocenten had onderzoek gedaan naar enkele voor veel leerlingen lastige onderwerpen en binnen elk van die onderwerpen moeilijke onderdelen, waarover leerlingen ook vaak misconcepten hadden, aangedragen. Mijn taak was om over die onderdelen een grote serie vragen in allerlei vormen (meerkeuze, open, invul, matching, juist-onjuist) te maken. Leerlingen konden die vragen dan zelfstandig maken vanaf de computer en kregen steeds feedback te zien op hun scherm. Daarbij werd uitgelegd wat goed was gegaan en wat fout, steeds met een duidelijke toelichting.

Afhankelijk van hun score waren ze hierna klaar met het behandelde onderwerp of gingen ze indien nodig voor een herkansing, met een nieuwe serie vragen. Vaak bleek de feedback hen goed te helpen bij die herkansing. Deze vorm is een voorbeeld van CAT (computergestuurd adaptief toetsen), geschikt voor zelfstandig werken door de leerling en eventueel docent-onafhankelijk uit te voeren.

Jarenlang heb ik mondelinge examens biologie, ANW en NLT afgenomen bij de staatsexamens. Daar doen de examenkandidaten wel een centraal schriftelijk examen, maar hebben ze geen schoolonderzoek. De reden daarvan kan zijn dat de kandidaten niet (meer) op een school zitten, of hun opleiding volgen op een school voor vso (voortgezet speciaal onderwijs). Vso-scholen hebben geen examenbevoegdheid, vaak zijn de docenten op die scholen ook niet bevoegd voor alle verschillende vakken die ze geven. Ook zo'n mondeling examen is een voorbeeld van een adaptieve toets. Anders dan bij een schriftelijke toets kun je de leerlingen die een verkeerde kant opgaan in hun redenering, weer op het rechte pad zetten en onderzoeken of ze het gevraagde antwoord eigenlijk wel weten. Leidraad bij de examinatoren was altijd: "We zijn op

zoek naar wat ze wel en niet naar wat ze niet weten.” We kregen als examinatoren ook trainingen om niet in allerlei valkuilen te trappen. Zo hebben sommige leerlingen wat langer tijd nodig om hun antwoord te formuleren. Wij leerden om rustig te wachten en niet meteen door te gaan naar een volgende vraag. Ook leerden we om zeer duidelijk vragen te formuleren. Vooral autistische leerlingen vatten alles wat je zegt nogal letterlijk op, dat vraagt een transparante communicatie. Uit het boek “Hoe we onszelf voor de gek houden” van Suzanne Weusten uit 2013 leerden we wat er allemaal mis kan gaan bij zo’n mondeling examen. Zo is er de beschikbaarheidsfout. Als je net een kandidaat hebt gehad die het er nogal matig afbracht, blijft die informatie in je geheugen hangen. Als de volgende kandidaat dan een fout maakt, komt die informatie snel beschikbaar: “O, weer zo’n zwakke kandidaat.”

Of je maakt bij de nabespreking met je collega die het examen heeft geprotocolleerd de confirmation bias-fout. Je vond het een prettig verlopend gesprek en hebt je onbewust geconcentreerd op de goede antwoorden van je kandidaat. Je collega is niet overtuigd dat jouw oordeel klopt en dat begrijp je pas, als je in het gemaakte protocol de door de kandidaat gemaakte fouten ziet beschreven. Of er komt een kandidaat binnen die heel zenuwachtig is, stottert en lang moet nadenken over het antwoord. Nu ligt de fundamentele attributiefout op de loer. “Ze weet haar zaakjes niet voldoende, daarom is ze zo zenuwachtig.” De kunst is je te concentreren op haar antwoorden, die bepalen als het goed is haar cijfer. Haar houding doet er eigenlijk niet zoveel toe.

Hoofdstuk 5 Toetscontrole en afronding

Als de toets is gemaakt, is het goed om te checken of alles klopt met de matrijs en om de tijd in te schatten die de leerlingen nodig zullen hebben. Controleer ook grondig of er geen vragen tussen zitten die je kunt beantwoorden met informatie uit een andere vraag. De leerling die dit ontdekt is een goede en slimme lezer, maar wordt dan niet echt getest op kennis. Als de toets in concept gemaakt is, is het hierna verstandig (eigenlijk noodzakelijk) om dit concept, het bijbehorende antwoordmodel en de matrijs door één en liefst meer vakcollega's te laten bestuderen. Het grote voordeel daarvan bij de vraagconstructie is de bewaakfunctie. Kill your darlings! Soms vind je zelf een vraag misschien wel prachtig, maar maken je collega's duidelijk dat hij dat niet is. Dan moet die vraag weg of misschien wordt hij in onderling overleg grondig aangepast.

Toetscontrole zorgt ervoor dat vakinhoudelijke fouten, onduidelikheden en niet passende relaties tussen de vragen en het antwoordmodel eruit gehaald worden. Ik herinner me dat mijn collega op school de vragen in de door mij gemaakte toetsen altijd eerst zelf maakte en daarna pas naar mijn antwoordmodel keek. Soms belde hij me dan en zei dat zijn antwoord bij een bepaalde vraag anders was dan in mijn antwoordmodel stond. Dus werd de vraag dan aangepast. Ook kunnen de collega's de tijd noteren die zij nodig hadden om de toets te maken.

Als vuistregel heeft een leerling 2-2,5 keer zoveel tijd nodig. Eventueel kan ook nog een spellingcontrole worden uitgevoerd.

Hoofdstuk 6 Toetsafname en correctie

Bij het afnemen van de toets in de klas is het verstandig om aan het begin de leerlingen te zeggen hoeveel vragen er zijn en hen erop te wijzen aan het eind te controleren of ze alle vragen gedaan hebben. En dat het verstandig is om bij meerkeuzevragen in ieder geval iets in te vullen, met de kans dat ze goed gokken. Wijs hen er ook op dat ze volledig vrij zijn om bij een andere vraag te beginnen als ze de eerste vraag moeilijk vinden. Controleer bij het inleveren ook of ze niet per abuis bepaalde vragen niet hebben gemaakt. Ik heb in mijn carrière menigmaal meegemaakt dat een leerling de laatste vragen, op de achterkant van een vel, niet had ontdekt! Geef dit ook door als iemand anders jouw toets surveilleert. Maak dat surveilleren gemakkelijker door de vragen in meerdere versies uit te delen, waardoor fraude hopelijk wordt voorkomen. Ook zou het goed zijn als de toets niet 's morgens vroeg wordt afgenomen. Uit onderzoek is gebleken dat veel leerlingen in de vroege ochtend nog niet maximaal alert zijn. Ook moet er rekening mee worden gehouden dat ongeveer 50% van de leerlingenpopulatie te maken kan hebben met periodieke problemen, namelijk tijdens de menstruatie. Als het nodig is, zou dat naar mijn mening aanleiding kunnen zijn om de toets voor de betreffende meisjes op een ander moment dan afgesproken af te nemen. Ik moest hier kortgeleden aan denken toen de Chinese tennisspeelster Zheng Qinwen de eerste set won in haar partij tegen de nummer 1 op de ranglijst, Iga Swiatek, maar daarna twee sets verloor, volgens haar mede door menstruatiepijn. Zoveel invloed kan dat soms blijkbaar hebben bij fysieke inspanning. Dat zou ook kunnen gelden voor geestelijke inspanning.

Met behulp van het antwoordmodel kun je na afloop het werk nakijken. Bij meerkeuzevragen is dat eenvoudig, bij open vragen niet. Leerlingen schrijven zelden precies het antwoord op dat in het antwoordmodel staat. In mijn Citotijd probeerden we te achterhalen wat leerlingen zoal zouden kunnen opschrijven bij een door ons voor een examen bedachte vraag. In een pretest werden die vragen, met een verplichting tot geheimhouding voor de docent, gemengd met diens eigen

vragen bij een toets. In sommige gevallen leidde dat tot een aanpassing van het antwoordmodel. Maar het blijft vaak een probleem om het door leerlingen gegeven antwoord te waarderen. Staat het er nu wel? En zo ja, staat het er volledig? Soms is het meer een kwestie van taal dan van begrip. Bij een vraag die drie punten oplevert, wordt vaak een redeneerlijn verwacht, zoals A, waardoor B, zodat C. Veel docenten trainen hun leerlingen daarop, maar er blijven leerlingen die nogal impliciet formuleren, terwijl ze de kern van de zaak begrepen hebben. De beoordeling doet een beetje denken aan het systeem van beoordelen bij het schansspringen. Je meet objectief de prestatie en daarnaast subjectief de stijl. Bij de centraal schriftelijke examens werkt men daarom met een intersubjectieve beoordeling. Naast de eigen docent kijkt ook een docent van een andere school het werk van de leerlingen na. Vervolgens overleggen beiden en komen tot een besluit. Soms worden ze het eens, soms wordt er gemiddeld. Bij grote afwijkingen tussen beiden kan er een derde corrector worden ingeschakeld. Vaak vergaderen vakgenoten in een regionale kring, waarbij ze elkaar de antwoorden van hun leerlingen voorleggen en uitzoeken wat nog aanvaardbaar is. Bij vragen die apert fout zijn of tot verwarring leiden, komt het Cito na enige tijd met een aanpassing van het antwoordmodel.

Hoofdstuk 7 Toetsanalyse

Of de leraar nu een eigen toets heeft nagekeken, de beoordelingen van alle docenten bij het centraal examen naar het Cito zijn opgestuurd, of als een toets digitaal wordt nagekeken, van belang is daarna de analyse van de toets, de bepaling van de cesuur en vervolgens van de cijfers van de leerlingen.

Door de ontwikkeling van de computer is de analyse veel eenvoudiger dan vroeger. Daartoe moeten bij een niet digitaal afgenomen toets eerst de scores in een soort Excelbestand worden gezet. Bij een digitaal afgenomen toets is dat alleen nodig indien er open vragen tussen zitten. Om te beginnen wordt eerst van elke vraag de moeilijkheid bepaald. Dit gebeurt met behulp van de p-waarde. Dat is bij een meerkeuzevraag een getal tussen 0 en 1 dat aangeeft welke fractie van de leerlingen die vraag goed heeft beantwoord. Bij een vierkeuzevraag zijn er ook nog drie a-waarden, dat zijn de getallen tussen 0 en 1 die aangeven welke fractie de bijbehorende afleider heeft gekozen.

Dus $p + a_1 + a_2 + a_3 = 1$

(Voor driekeuzevragen geldt $p + a_1 + a_2 = 1$)

Een p-waarde van 1 betekent dat iedereen de vraag goed heeft, een p-waarde van 0 dat niemand de vraag goed heeft. In het eerste geval is de vraag te gemakkelijk, in het tweede geval te moeilijk.

Je zou kunnen tegenwerpen dat de bedoeling van onderwijs is dat je leerlingen de stof volledig beheersen. Dus dan zou je blij zijn als iedereen de vraag goed heeft. Het probleem is dat er twee categorieën leerlingen zijn die dit droombeeld verstoren: leerlingen die ondanks alle moeite die leraren doen tijdens het onderwijs, te weinig capaciteiten hebben op dit terrein en leerlingen die onvoldoende gestudeerd hebben op de stof. Dus een paar vragen met $p = 1$ is leuk, dat bevordert zeker aan het begin van een toets de motivatie, maar het moeten er niet te veel worden.

Aangezien de toets dus onderscheid wil maken tussen kandidaten die de stof goed beheersen en zwakkere kandidaten, heeft men een p-waarde van 0,67 (bij vierkeuzevragen) en van 0,70 (bij driekeuzevragen) gekozen als ideaal.

Bij open vragen is de analyse wat ingewikkelder. Dat komt doordat sommige vragen meer punten opleveren dan andere. Waar voor meerkeuzevragen vaak een vast puntentotaal wordt gebruikt (bijvoorbeeld altijd 1 punt of altijd 2 punten), zie je bij open vragen een flinke variatie, die kan lopen van 1 tot 5 punten. Als een vraag die 5 punten waard is een p-waarde van 0,4 heeft, kan dat van alles betekenen. Bijvoorbeeld 40% heeft die vraag helemaal goed en 60% helemaal fout, maar ook alles daartussen!

De volgende stap is de berekening van de Rit-waarde of correlatiecoëfficiënt, die een indicatie geeft van de correlatie tussen de betreffende vraag en de andere vragen in de toets. Je wilt graag weten of de score van de onderzochte vraag past in het beeld dat tevoorschijn komt uit de resultaten van de andere vragen, met als gezamenlijk doel de kennis en inzichten over de verschillende leerdoelen te meten. Dat betekent dat je ervan uit moet gaan dat leerlingen die een bepaalde vraag goed hebben ook meer kans hebben om andere vragen in je toets goed te hebben en dat het niet een soort geluksspel is.

De Rit-waarde van een vraag heeft een bereik tussen -1 en +1. Als de Rit-waarde 1 is, betekent dit dat de betreffende vraag door alle goed scorende leerlingen goed en door alle slecht scorende leerlingen fout wordt beantwoord. Bij een Rit-waarde van -1 is het omgekeerde het geval, dat wordt beschouwd als een aanwijzing dat er iets niet klopt aan die vraag.

Er is natuurlijk een relatie tussen p-waarde en Rit-waarde. Als $p = 1$, heeft iedereen de vraag goed en is de Rit-waarde automatisch 0. Vooral als de Rit-waarde laag is (0,15 wordt vaak als grens aangehouden) en de p-waarde ook (een moeilijke vraag), kan dat aanleiding zijn om de vraag te verwijderen.

Een veel gebruikte variant is de Rir-waarde, dat is de correlatie tussen de vraag en de rest (alle overige vragen). Dus bij de berekening van de correlatie wordt de totaalscore bekeken minus de eigen bijdrage van de vraag. Bij een toets met veel of uitsluitend meerkeuzevragen is dat niet zo belangrijk (alle vragen hebben dezelfde waarde). Maar als er open vragen tussen zitten met een flinke variatie in waarde, kan dat wel een rol spelen.

Er zijn meerdere manieren om de betrouwbaarheid van de hele toets te bepalen. Het gaat hier om de interne betrouwbaarheid van de afgenomen toets. Met externe betrouwbaarheid wordt bedoeld dat als de toets meerdere keren op verschillende momenten met dezelfde leerlingen wordt uitgevoerd, de resultaten niet wezenlijk anders worden. In de praktijk wordt dit echter nauwelijks toegepast.

Een veel gebruikte manier om de interne betrouwbaarheid te bepalen is de Cronbach's alfa-methode, ontwikkeld door Lee Cronbach in het artikel 'Coefficient alpha and the intern structure of tests' in het tijdschrift Psychometrika 16, p. 297-334 uit 1951.

Daarin gebruikt hij de volgende formule:

$$\alpha = \frac{n}{n - 1} \left(1 - \frac{\sum s^2(X_i)}{s^2(Y)} \right)$$

Daarin staat het symbool α voor de interne betrouwbaarheid van de toets. Verder is n het aantal vragen, staat $\sum s^2(X_i)$ voor de som van de varianties van alle vragen en $s^2(Y)$ voor de variantie van de totale scores van de toets.

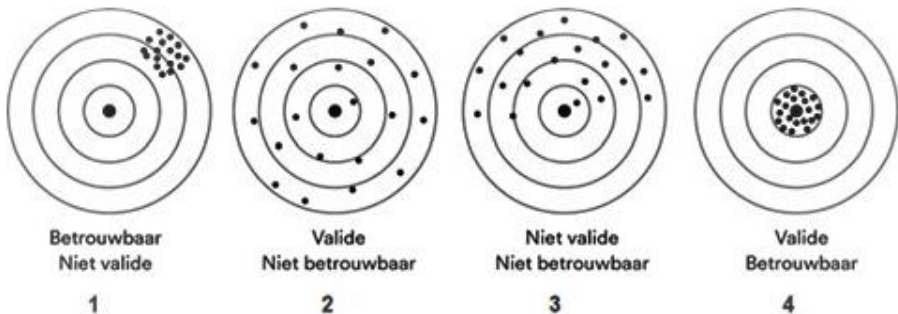
Cronbach's alfa kan waarden aannemen tussen 0 en 1.

Waarden tussen 0,7 en 1,0 worden beschouwd als een indicatie dat de toets betrouwbaar is.

Behalve betrouwbaarheid speelt ook validiteit een rol. Toetst de toets echt de leerdoelen die we willen testen? Daartoe is geen

formule voorhanden. De matrijs speelt hier een belangrijke rol. Zijn alle leerdoelen aan de orde gekomen en op een evenredige en adequate manier? Kunnen we geldige conclusies trekken over de prestaties van de leerlingen op basis van hun score op de toets?

Hieronder een afbeelding ter verduidelijking om het belang van de begrippen betrouwbaarheid en validiteit in beeld te brengen.



Stel dat je bij de 100 meter sprint de tijden van de verschillende lopers wilt meten. Als je zes juryleden hebt die ieder een stopwatch gebruiken die ze moeten indrukken bij start en finish, heb je in principe wel een valide instrument. Je kunt er de tijd mee meten. Maar niet iedereen drukt precies op dezelfde wijze en bovendien is een stopwatch niet nauwkeurig genoeg, dus onbetrouwbaar (afbeelding 2). De tijden worden tegenwoordig in duizendsten gemeten met behulp van elektronische apparatuur, pas dan is de uitkomst van de sprint valide en betrouwbaar (afbeelding 4).

Als je bij het schaatsen heel nauwkeurig de tijden op de 500 meter hebt gemeten, heb je nog geen enkele indicatie voor de uitslag na alle vier afstanden bij het allround-toernooi voor de vrouwen (500, 1500, 3000 en 5000 meter). Die ene meting is dus wel betrouwbaar, maar niet valide voor het bepalen van de beste allrounder (afbeelding 1).

Hoofdstuk 8 Cesuurbepaling en cijfers

Hoe bepaal je de cesuur: de grens tussen voldoende en onvoldoende?

Een veel gebruikte methode is te werken met de volgende formule:

$$\frac{\text{(score in punten / maximale score in punten)} \times 9 + 1}{1}$$

Die 1 wordt toegevoegd omdat we vaak werken met een omrekening

naar een cijfer in het tientallig stelsel, waarin 1 dus de ondergrens is.

Toen ik lang geleden zelf nog op school zat, werd soms een 0 uitgedeeld. Rekenkundig is dat onjuist. Als een leerling een leeg proefwerkblaadje inlevert, hoor je als cijfer een 1 in te vullen.

Standaard wordt een score van 55% gebruikt als grens tussen voldoende en onvoldoende: de cesuur.

Een voorbeeld. Er zijn bij een toets 40 punten te behalen. Een leerling heeft 55% van de punten gehaald, dus 22 punten.

$$22/40 \times 9 + 1 = 5,95 \text{ (voldoende).}$$

Cesuurbepaling houdt echter veel meer in dan afspreken dat een score van 55% een voldoende oplevert.

Ten eerste komt het vaak voor dat in de toets naast open vragen ook meerkeuzevragen zijn opgenomen, waarbij je rekening moet houden met de gokkans. Dan wordt de bovenstaande berekening, die is gebaseerd op uitsluitend open vragen, ingewikkelder.

Bij tweekeuzevragen (bijvoorbeeld juist-onjuist) is de gokkans 50%, bij driekeuzevragen 33%, bij vierkeuzevragen 25% enzovoort.

Dus ligt de cesuur bij uitsluitend tweekeuzevragen op $55 + \frac{1}{2} \times 45 = 77,5\%$, bij uitsluitend driekeuzevragen op $55 + \frac{1}{3} \times 45 =$

70% en bij uitsluitend vierkeuzevragen op $55 + \frac{1}{4} \times 45 = 66,3\%$.

Bij een combinatie wordt de cesuurbepaling dus een uitgebreid stuk rekenwerk. Gelukkig kan dit werk uitbesteed worden aan speciale computerprogramma's.

Weer een voorbeeld. Stel dat een toets bestaat uit 20 vierkeuzevragen (score per vraag 2p), 10 juist-onjuistvragen (score per vraag 1p) en 5 open vragen (score totaal 12p). Dan is de berekening als volgt:

$$20 \times 2 \times 0,663 + 10 \times 1 \times 0,775 + 12 = 26,52 + 7,75 + 12 = 46,27.$$

Dus bij een maximale score van 62 is de cesuur dan 46,27.

Ten tweede zijn er niet slechts een, maar twee methodes om de cesuur te bepalen. In het bovenstaande geval, waarbij vooraf wordt bepaald wat de cesuur is, is sprake van een *absolute cesuur*. Maar je kunt er ook voor kiezen om de cesuur achteraf te bepalen, op basis van de scores van alle studenten: een *relatieve cesuur*. Beide methodes kun je ook combineren.

De absolute cesuur wordt het meest gebruikt in de praktijk en is in de meeste gevallen ook aan te raden. Maar als die cesuur ertoe leidt dat vrijwel iedereen een onvoldoende haalt, terwijl uit de vragenanalyse blijkt dat zich daar geen duidelijke problemen voordoen, ga je misschien toch de norm bijstellen: het werk was blijkbaar erg moeilijk. Zo zie je dat bij het Cito per vak via een aanpassing van de zogenaamde N-term de examencijfers kunnen worden bijgesteld. Dan ontstaat een relatieve cesuur.

Een tussenvorm vind je bij de cesuurmethode van Cohen-Schotanus, van der Vleuten en Bender, gepubliceerd in een artikel in het tijdschrift 'Onderzoek van Onderwijs' in 1996. Bij die methode bepaal je vooraf een absolute cesuur. Normaal gesproken krijgt een leerling de maximale score (een 10) als hij of zij de (theoretisch) hoogst haalbare score heeft behaald (100% van de punten). De methode van Cohen-Schotanus gaat er echter vanuit dat niet 100% de hoogst haalbare score is, maar de score van de beste (vijf of tien) leerling(en). De

cesuur wordt vervolgens berekend aan de hand van die hoogst haalbare score.

Een toets bestaat bijvoorbeeld uit 40 open vragen. De absolute cesuur is 55%. Elke vraag kan één punt opleveren. Het is theoretisch dus mogelijk om 40 punten te halen. Bij een normale cesuur krijgt een leerling pas een 10 als hij of zij alle 40 punten haalt. Bij 22 punten krijgt hij of zij een voldoende (5,95). Na afname van de toets bij 100 leerlingen blijkt dat de beste vijf leerlingen gemiddeld 36 punten haalden. Volgens de methode van Cohen-Schotanus et al moet de cesuur nu zo aangepast worden, dat 36 punten de maximale score oplevert. Als gevolg daarvan hoeft een leerling niet 55% van 40 punten te halen voor een voldoende, maar 55% van 36 punten. De daadwerkelijke cesuur ligt dan op 20 punten (in plaats van de oorspronkelijke 22). Zo wordt bij deze cesuurmethode rekening gehouden met de moeilijkheid van de toets.

In het proefschrift 'Setting Standards in Small Samples' uit 2022 beschrijft Monika Vaheoja dat een goede cesurbepaling extra lastig is bij afname van een toets in een nogal kleine populatie, zoals in het voortgezet onderwijs bij door een school zelf afgenomen toetsen vaak het geval is. Verschillende toetsen over hetzelfde onderwerp kunnen verschillen in moeilijkheidsgraad en bij kleine groepen is de betrouwbaarheid die bepaald wordt via psychometrische analyse minder nauwkeurig dan bij grote groepen. Dan wordt het lastig om vast te stellen of toets 1 moeilijker is dan toets 2 of dat de bij die twee toetsen geteste leerlingen verschillen in kennis en of vaardigheden. Een probleem dat voor veel docenten niet gemakkelijk is op te lossen. Een mogelijkheid ligt hier in het gebruik van zogenaamde ankervragen, die in de twee verschillende toetsen over hetzelfde onderwerp gebruikt worden en waarbij je de psychometrische waarden voor de twee geteste groepen kunt vergelijken.

Hoofdstuk 9 Feedback op de toets

Als je de toetsresultaten bekend maakt, is het belangrijk dat je als leraar met je leerlingen niet alleen de fouten bespreekt, maar ook benadrukt wat ze goed hebben gedaan. Eigenlijk begint dat al bij de correctie. Beperk je niet tot het met de beroemde rode pen aanstrepen van fouten en het vermelden van het totale aantal fouten. Geef in het werk van je leerlingen dus ook aan wat ze goed hebben gedaan en noteer het totale aantal behaalde punten, omgezet naar een cijfer. Feedback bestaat uit een combinatie van positieve en negatieve opmerkingen. Bedenk dat mensen meer gevoelig zijn voor afbrekende dan voor opbouwende kritiek. Een oud Russisch spreekwoord zegt: "Eén lepel teer kan een vat honing bederven, maar één lepel honing in een vat teer betekent niets." Een verzameling negatieve opmerkingen leidt bij leerlingen tot de 'mosselreflex': zij sluiten zich af en reageren daarna niet meer op aanmoedigingen die bedoeld zijn voor de toekomst. Sta als leraar ook open voor kritiek van je leerlingen. Iedere leraar maakt wel eens een rekenfout, vergist zich bij het optellen van alle punten of ziet niet altijd dat sommige van het antwoordmodel afwijkende antwoorden toch goed zijn.

Hoofdstuk 10 Kritiek op toetsen

De laatste jaren is er veel kritiek op toetsen. Volgens sommigen staan toetsen het eigenlijke leerproces in de weg en ontstaat een 'teaching to the test'-cultuur. Leerlingen gaan pas serieus aan het werk vlak voor de toets, stampen soms tot diep in de nacht, maken de toets en zijn kort daarna alles weer kwijt: zweten, weten, vergeten. Dat is een reëel probleem.

Uiteindelijk gaat het bij leren niet om het even scoren van een goed cijfer, maar om een doorlopende ontwikkeling, waarbij kennis, vaardigheden en inzicht geleidelijk toenemen. Daarom is er, vooral aangezwengeld door een boek als "Toetsrevolutie. Naar een feedbackcultuur in het voortgezet onderwijs" van Dominique Sluijsmans en René Kneijber uit 2016, toenemende belangstelling voor formatief handelen, waarbij het leren wordt gestuurd door voortdurende feedback. Daarin is de toets als afsluiting van het leerproces (de summatieve toets) niet verdwenen, maar een eindstap in een continu leerproces. Wel vermindert het aantal summatieve toetsen in het leertraject. Dat heeft onvermijdelijk tot gevolg dat de hoeveelheid stof per toets toeneemt.

Ook is het niet mogelijk alles wat in het onderwijs aan de orde komt, in toetsen onder te brengen. Dat kan betekenen dat je ervoor kiest naar andere vormen van beoordeling te zoeken, bijvoorbeeld met behulp van rubrics, portfolio's, verslagen of gesprekken. In dit boek ga ik hier niet verder op in, behalve dan door op te merken dat hier minder mogelijkheden liggen voor valide en betrouwbare meting.

Daarnaast is er kritiek op toetsen van meer psychologische aard. Een aantal mensen ervaart een summatieve toets als stress bevorderend en vindt vooral een momentopname met grote invloed op het eindresultaat, zoals bijvoorbeeld het centraal schriftelijk examen in het voortgezet onderwijs, ongewenst. Zij pleiten voor het zogenaamde programmatische toetsen. In hun opvatting zijn docenten daarbij prima in staat via allerlei tussentijdse beoordelingen met voortdurende

feedback tot een goed oordeel over hun leerlingen te komen. Diep in hun hart willen zij eigenlijk af van een 'stressvolle' summatieve toets aan het eind. Mij overtuigen deze mensen niet. Ten eerste denk ik niet dat stress ten allen tijde vermeden moet worden. Bijvoorbeeld in de sport of bij een muzikuitvoering accepteren we zonder veel protesten een juniorensprintwedkamp op de 100 meter sprint of een uitvoering met je band of jeugdorkest. Bij wedstrijden accepteren we ook dat er winnaars en verliezers zijn en dat er ranglijsten worden opgesteld na bijvoorbeeld een hardloopwedstrijd over 3 km. Niemand klaagt over het feit dat je al op jonge leeftijd in een schriftelijk gedeelte en een praktijkexamen moet laten zien dat je kunt autorijden. Bovendien denk ik dat het goed is om aan het eind van je loopbaan als leerling in het voortgezet onderwijs een centrale toets af te leggen, zodat je vervolgopleiding een zekere garantie heeft over je niveau. Niet voor niets stellen we in het onderwijs, zoals eerder beschreven in dit boek, kwaliteitseisen zoals betrouwbaarheid en validiteit aan zo'n toets. Daarmee proberen we een objectieve meting te waarborgen. Hoe goed een docent ook is, hij of zij is toch in zekere mate subjectief. Dat kan voor leerlingen gunstig of ongunstig uitpakken. In 2020 was er geen Citotoets op de basisschool en geen centraal schriftelijk examen in het voortgezet onderwijs. Het resultaat was dat op de basisschool twee groepen een gemiddeld lager advies over het vervolg van hun onderwijs van hun school meekregen dan in de voorafgaande jaren, waarin wel een Citotoets werd afgenomen. Dat betrof kinderen uit lagere sociale klassen en meisjes. En in het voortgezet onderwijs was, ondanks het feit dat veel lessen door corona niet konden worden gegeven, het percentage geslaagden hoger dan in de daaraan voorafgaande jaren. Er zijn meer mensen die vinden dat de kritiek op toetsen overdreven wordt. In het artikel 'De mythe van de afrekencultuur' in Didactief online van 13 mei 2022 van Jaap Scheerens en Ton van Haperen keren de auteurs zich tegen

het beeld dat toetsen op school leiden tot een afrekencultuur, waarin 'meten is weten' het enige is dat telt. Volgens hen is die kritiek niet gebaseerd op technische argumenten, maar meer op politieke, organisatorisch-technische en pedagogische argumenten.

Het blijkt moeilijk om binnen het onderwijs tot een gezamenlijk standpunt hierover te komen.

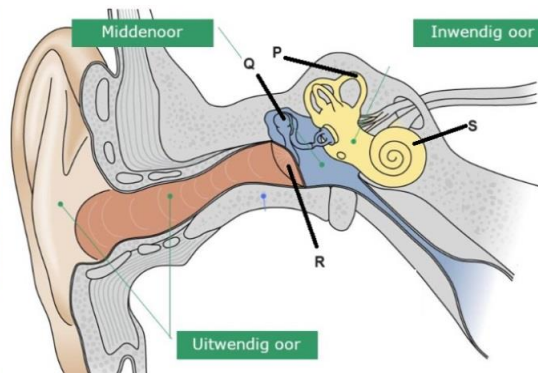
Hoofdstuk 11 Een serie veelkleurige vragen voor de onderbouw vmbo/havo/vwo

OPGAVEN

Nibiconferentie mei 2024
RWE

1 Concert

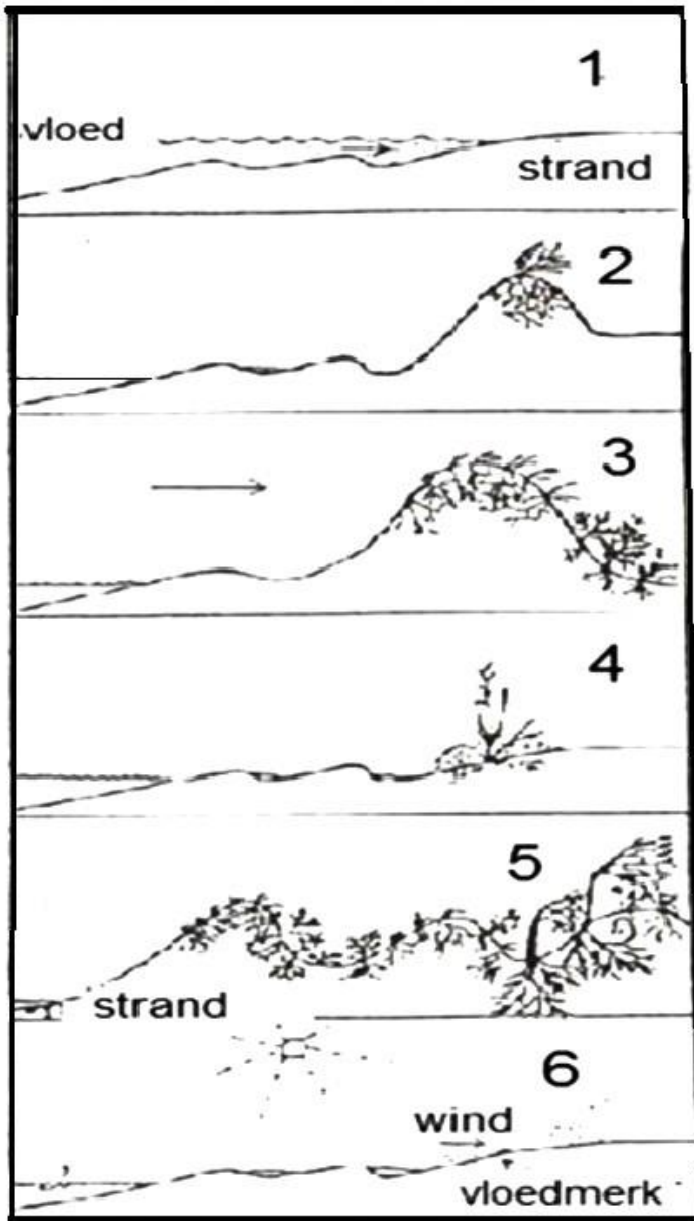
Josien Bakker trad op met haar band in het Groningse Veendam. Haar zang bereikten de oren van de luisteraars in de vorm van geluidstrillingen. Met welke letter in de afbeelding wordt het onderdeel van het oor aangegeven dat geluidstrillingen worden omgezet in impulsen?



- A met P
- B met Q
- C met R
- D met S

2 Duinvorming 1/3

Zet de zes verschillende stadia van de duinvorming in de afbeelding in de g van begin tot eind.

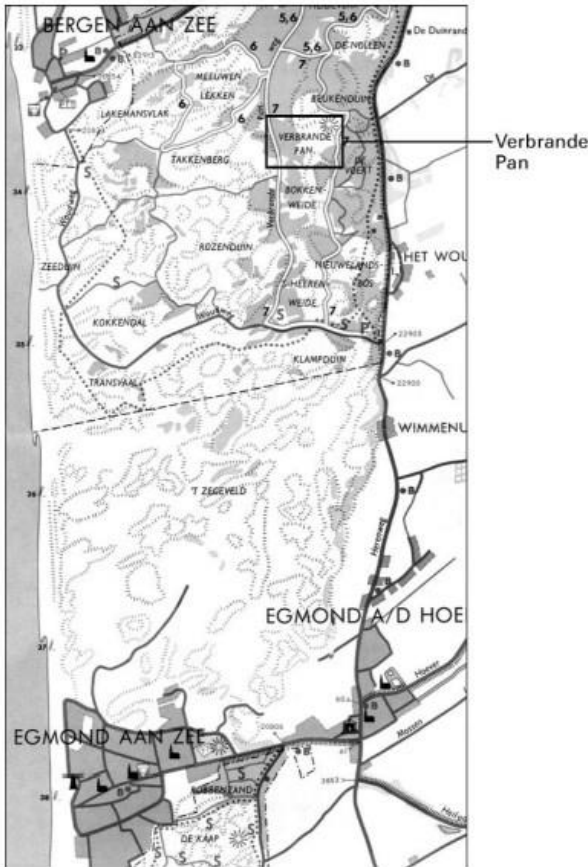


3 Duinvorming 2/3

Hoe noemen we de hele ontwikkeling van de eerste soort naar een compleet duingebied?

4 Duinvorming 3/3

Kardinaalsmutsen zijn planten die vooral in de kalkrijke duinen voorkomen: in Zeeland tot Bergen in Noord-Holland. In het kalkarme deel van het duingebied tot Den Helder en op de Waddeneilanden komen deze planten niet voor. Voorts geldt het omgekeerde: deze plantensoort vind je vooral in het noordelijk duingebied ten zuiden van Bergen. In het duingebied verandert bij het duinreservaat 'de Verbrande Pan' tussen Egmond en Bergen zowel het kalkgehalte als de vegetatie heel sterk. Door verschillende zeestromingen komt er meer kalk in het zuidelijk duingebied. Waar komt die kalk vandaan?



5 Longvolume

Hieronder staan enkele gegevens over de ademhaling van persoon P.
Wat is het maximale longvolume van deze persoon P?

Vitale capaciteit 4 liter. Ademvolume 0,5 liter. Restvolume* 1,5 liter.

*Restvolume = de lucht die na diepe uitademing nog in de longen aanwezig is.

- A 3,5 liter
- B 4 liter
- C 4,5 liter
- D 5,5 liter
- E 6 liter

6 Grafieken maken

René wil een grafiek maken van de 106 marathons die hij tussen 1982 en 2018 heeft gelopen. Hij wil hierbij de tijd per marathon uitzetten tegen de jaartallen. Hoe zou hij dit het beste kunnen doen? Wat is de juiste manier is om dit te doen.

Welke variabele moet er op de X-as komen en waarom?



- A De tijd per marathon, want dit is de gemeten variabele.
- B De tijd per marathon, want dit is de gekozen variabele.
- C De jaartallen, want dit is de gemeten variabele.
- D De jaartallen, want dit is de gekozen variabele.

7 WK torenspringen 1/2

Else Guurtje Praasterink haalde op het WK 2022 in Boedapest de finale bij het torenspringen. Daarbij moet je vanaf een platform op 10 meter hoogte in het water springen en onderweg combinaties van salto's en schroeven maken.

Als Else Guurtje boven op het platform op 10 meter boven het water staat, vult ze haar hoofd met iets aan haar hormoonspiegel.

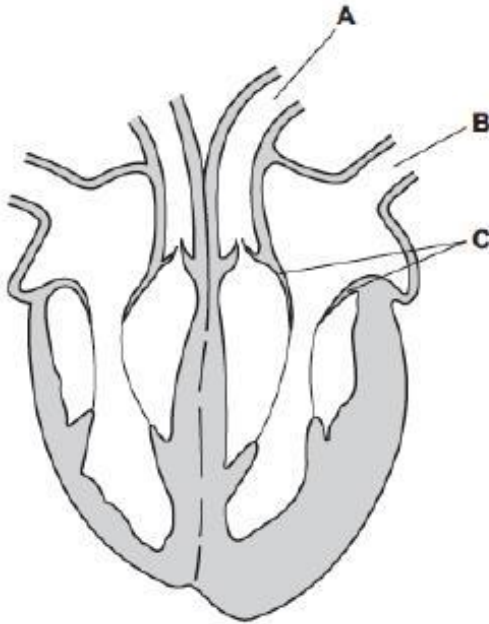
Welke uitspraak is juist?



- A Alleen de productie van adrenaline neemt toe.
- B Alleen de productie van insuline neemt toe.
- C De productie van zowel adrenaline als insuline neemt toe.

8 WK torenspringen 2/2

Terwijl Else Guurtje op het platform staat, neemt haar hartslag ook toe. Je ziet hieronder een schematische doorsnede van het hart.
Welk bloedvat wordt aangegeven met A?



- A aorta
- B holle ader
- C longader
- D longslagader

9 Pupillen

Hier zie je Thekla met grote pupillen.
Waardoor zijn die pupillen zo groot?



- A Ze is boos.
- B Ze is net opgestaan.
- C Ze staat in de schemering.

10 Een proefje met suiker 1/2

Bekijk de tabel hieronder.

Sterkte suikeroplossing	Wel of niet zoet
1 mg/L	Niet zoet
10 mg/L	Wel zoet
100 mg/L	Wel zoet

Joris en Jolien doen een practicum om de drempelwaarde te bepalen voor suiker in een oplossing. Ze gebruiken hiervoor drie verschillende suikeroplossingen. Ze brengen met wattenstaafjes van elke oplossing enkele druppels op de tong en proeven van elke oplossing aan of die wel of niet zoet smaakt. De resultaten zijn in de tabel hierboven weergegeven. Waar ligt bij Jolien de drempelwaarde voor het proeven van een suikeroplossing?

- A Beneden 1 mg/L
- B Tussen 1 en 10 mg/L
- C Tussen 10 en 100 mg/L
- D Boven 100 mg/L

11 Een proefje met suiker 2/2

Wat is bij dit practicum de adequate prikkel voor de smaakzintuigcellen in de Jolien?

- A De opgeloste stof.
- B Het aanraken met het wattenstaafje.
- C Het water van de oplossing.

12 Gedrag van een salamander

Als het mannetje van de salamander een vrouwtje ontmoet, zie je bepaalde gedragselementen, zoals 'staart wapperen', 'staart wapperen met kromme r met de kop'.

Tot welk gedrag horen die elementen? .



- A Tot het conflictgedrag.
- B Tot het voedingsgedrag.
- C Tot het voortplantingsgedrag

13 De ooglenzen van Jane

Jane tuurt naar een paar eenden aan de andere kant van het water.

Zijn haar ooglenzen nu bol of plat?

Die zijn



14 Kraanvogels

Op de afbeelding zie je twee kraanvogels die samen een dans uitvoeren. Het is bedoeld om elkaar te verleiden.

Van welk soort gedrag is hier sprake?



- A Balts
- B Broedzorg
- C Paringsgedrag
- D Territoriumgedrag

15 Bevalling

Tijdens de bevalling ontstaan weeën. Door weeën wordt de hals van de baar wijder.

Hoe heet het gedeelte van de bevalling waarin dit gebeurt?



- A Indaling
- B Nageboorte
- C Ontsluiting
- D Uitdrijving



Bij bonobo's komt seksueel gedrag voor tussen mannetje en vrouwtje, maar ook tussen twee mannetjes en tussen twee vrouwtjes

Seksualiteit heeft drie functies: 1. Voortplanting; 2. Relaties versterken; 3. G. Welke functie of functies heeft seksualiteit tussen twee mannetjes?

- A Alleen 1
- B Alleen 2
- C Alleen 3
- D 1 en 2
- E 1 en 3
- F 2 en 3

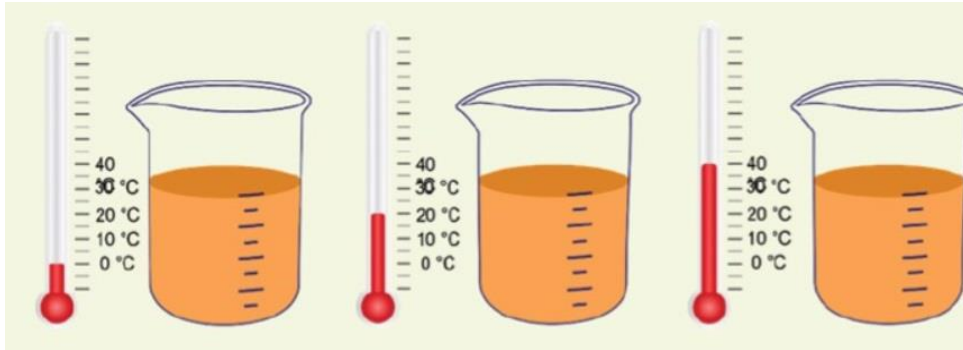
17 Menstruatie

De mens is een van de weinige zoogdieren waarbij menstruatie plaatsvindt. Bij een deel van de mens is een deel van het baarmoederslijmvlies het lichaam. Bij veel andere zoogdieren wordt een deel afgebroken en worden de vrijkomende stoffen opgenomen in het bloed en hergebruikt.

Leg uit wat de functie is van het regelmatig verwijderen van een deel van het baarmoederslijmvlies.

18 Proef met water

Je ziet in de opstelling hiernaast drie bekers met water. Niels doet twee vingers linkerhand in ijswater in het linker bekersglas en twee vingers van zijn rechterhand in het rechter bekersglas. Na twee minuten doet hij alle vier vingers in het middelste bekersglas. Wat neemt hij waar?



- A Alle vier vingers voelen koud aan.
- B Alle vier vingers voelen warm aan.
- C De vingers van zijn linkerhand voelen koud aan en de vingers van zijn rechterhand voelen warm aan.
- D De vingers van zijn linkerhand voelen warm aan en de vingers van zijn rechterhand voelen koud aan.

19 Ballet

Topballetdanseres Igone de Jongh toont hier een moeilijke oefening. In welke toestand zijn haar rompspieren hier?



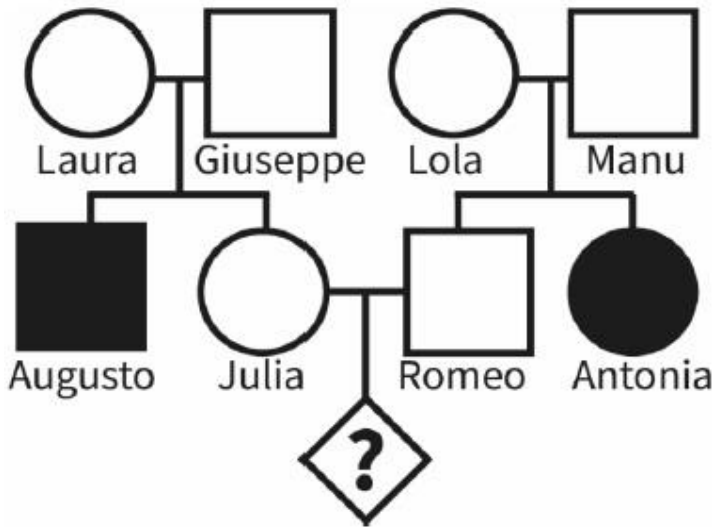
- A Buikspieren en rugspieren samengetrokken
- B Buikspieren samengetrokken en rugspieren uitgerekt
- C Buikspieren uitgerekt en rugspieren samengetrokken
- D Buikspieren en rugspieren uitgerekt

20 De ziekte van Tay-Sachs

De ziekte van Tay-Sachs erft recessief over. De mutatie bevindt zich niet in het geslachtschromosoom.

In de stamboom zijn gezonde mensen aangeduid met witte vormen, en mensen met de ziekte van Tay-Sachs zijn aangeduid met zwarte vormen.

Hoe groot is de kans dat een kind van Romeo en Julia de ziekte van Tay-Sachs hebben? Geef je antwoord als een breuk.



21 Stoeplantjes

In mei 2020 bereikte vanuit Frankrijk een nieuwe rage ons land: het botanische stoeplantje. Mathilde maakt met stoepkrijt een cirkel rond een stoeplantje, krijgt de naam van het plantje eronder en zet hiervan foto's op instagram. Veel Nederlanders volgen haar voorbeeld. Mathilde roept iedereen op om stoeplantjes verder te onderzoeken en te beschermen, want ze zijn nuttig.

Naar aanleiding van deze oproep worden de volgende beweringen gedaan. Welke beweringen hebben welke mogelijke effecten van stoeplantjes op het ecosysteem.

- I. I. Ze helpen de stad te verkoelen.
- II. II. Ze leveren voedsel voor bijen en andere bestuivers.
- III. III. Ze vergroten de biodiversiteit.
- IV. IV. Ze zorgen voor minder wateroverlast bij zware regenval.

Welke bewering is of welke zijn juist?



- A Bewering I
- B Bewering II
- C Bewering III
- D Bewering IV

22 Linnaeus

Op de afbeelding zie je de Zweedse barones Hedda Stiernstedt met en bloemkrans op haar hoofd, Zo'n krans wordt veel gedragen op het Midzomerfeest.

Haar landgenoot ridder Carl von Linné (Linnaeus) (1707-1778) was de eerste die plantensoorten indeelde op grond van hun bloembouw.

Welke eigenschap van bloemen is het meest geschikt om te gebruiken voor de indeling van plantensoorten?



- A De dikte van de bloemsteel.
- B De grootte van de bloem.
- C De kleur van de bloem.
- D Het aantal meeldraden van de bloem.

23 Gastroparese

Zangeres Emma Kok heeft de zeldzame chronische ziekte gastroparese, ook maagverlamming genoemd. De diagnose kwam vijf jaar geleden, maar de symptomen zijn er altijd al. Bij gastroparese werken de maagspieren slecht of helemaal niet, waardoor de maag niet normaal kan worden gelegeerd.

Leg uit dat hierdoor problemen optreden bij de spijsvertering.

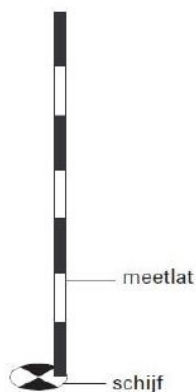


24 Onderzoek bij een sloot

Een groep leerlingen onderzoekt de helderheid van slootwater in een diepe behulp van een speciale meetlat. Op de meetlat staan afwisselende zwarte strepen van 10 centimeter lang en onderaan de lat zit een zwart-witte schijf. laten de meetlat in het water zakken. Hoe diep je de schijf kunt zien, geeft a het water is.

De leerlingen willen de meting nauwkeuriger maken.

Welke aanpassing aan de meetlat kan de meting nauwkeuriger maken?



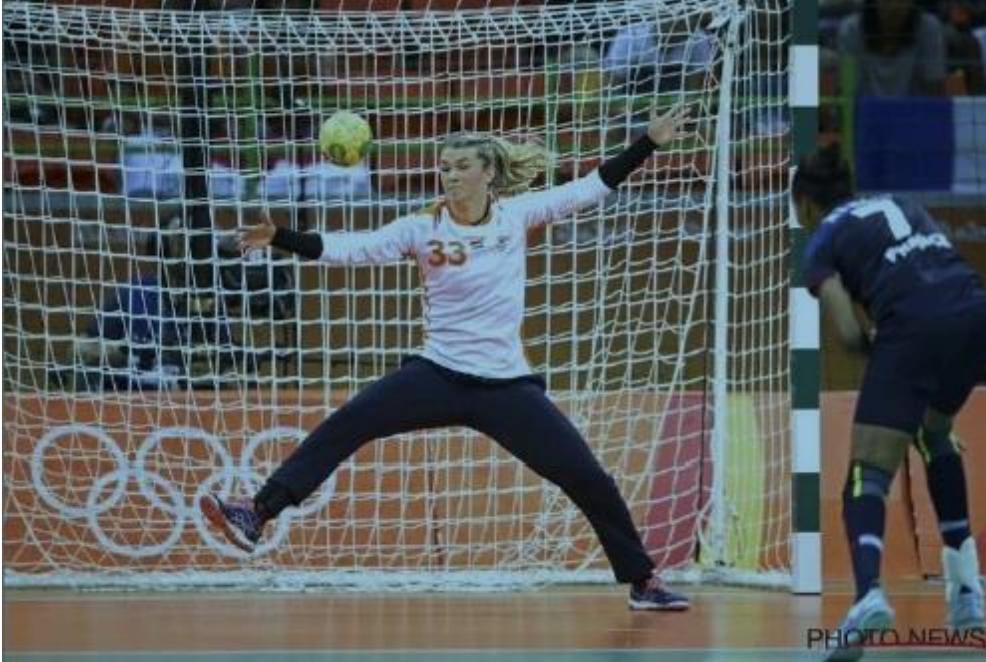
- A De meetlat krijgt strepen van elk 5 centimeter lang.
- B De leerlingen gebruiken een langere lat.
- C De schijf aan de lat wordt groter gemaakt.
- D De schijf aan de lat wordt helemaal zwart gemaakt.

25 Handbal 1/2

Tess Wester was keepster van het Nederlandse handbalteam. Op de foto is Tess te zien die de bal uit de lucht vangt met de sprongtechniek 'Hampelmann' uit, een sprongtechniek waarbij ze de kans op een doelpunt door zich extra breed te maken.

Hieronder staan vier beweringen over Tess' rechterbeen op het moment dat de foto genomen werd.

Welke bewering is juist?



- A De buigspieren in het bovenbeen en onderbeen zijn beide aangespannen.
- B De buigspier in het bovenbeen is aangespannen en de buigspier in het onderbeen is uitgerekt.
- C De buigspier in het bovenbeen is uitgerekt en de buigspier in het onderbeen is aangespannen.
- D De buigspieren in het bovenbeen en onderbeen zijn beide uitgerekt.

26 Handbal 2/2

Tess gebruikt de sprongtechniek, omdat ze waarnam dat de tegenstandster de bal richting doel probeerde te gooien.

Waar in Tess' ogen worden de eerste impulsen opgewekt door de lichtprikkels in de omgeving?



- A In de iris
- B In het hoornvlies
- C In het netvlies
- D In het vaatvlies

27 Onderzoek aan bruinvissen 1/2

Lonneke IJsseldijk van de faculteit diergeneeskunde van de Universiteit U onderzoek aan walvissen. Met haar team zocht zij uit waardoor er steeds bruinvissen worden gevonden met ernstige verwondingen. Uit dit onderzoek dat de verwondingen niet door scheepsschroeven of netten werden veroorzaakt door een dier (de verwondingen waren dus bijtsporen). Er waren bruinvissen met verwondingen aan hun zijkant en bruinvissen met verwondingen aan de kop. Mardik Leopold (van Wageningen Marine Research) ontdekte botjes van de maag van sommige verwonde bruinvissen. Deze botjes waren van verschillende soorten. Sommige vissen leven in het oppervlaktewater van de Noordzee terwijl andere vlak boven de zeebodem van de Noordzee leven.

Welk type verwonding is te verwachten bij bruinvissen waarbij in de maag botjes gevonden van vissen die vlak bij de bodem leven? En welk type verwonding

verwachten bij bruinvissen met botjes in de maag van vissen uit oppervlaktewater



- A Vlak bij de bodem en in het oppervlaktewater wonden aan de keel.
- B Vlak bij de bodem wonden aan de keel en in het oppervlaktewater wonden aan de zijkant.
- C Vlak bij de bodem wonden aan de zijkant en in het oppervlaktewater wonden aan de keel.
- D Vlak bij de bodem en in het oppervlaktewater wonden aan de zijkant.

28 Onderzoek aan bruinvissen 2/2

Bij de bruinvissen werden nog andere bijtsporen ontdekt, die veroorzaakt worden door hoektanden. Na DNA-onderzoek bleek de grijze zeehond de dader. Het bijten van zeehonden is een recente trend, misschien door een toevallige ontdekking van een zeehond die succes had met het aanvallen van bruinvissen.

Door welk leerproces heeft dit gedrag zich verspreid van één zeehond naar andere soortgenoten?

- A Conditionering (aanpassen aan de omgeving via herhaaldelijke prikkel)
- B Gewenning
- C Imitatie
- D Trial and error (proefondervindelijk leren)

29 Fossielen

Op het Zweedse eiland Öland krijg je als finisher van de marathon geen medaille, maar een stuk geslepen kalksteen met een fossiele inktvis.

Waaruit is het fossiel van de inktvis ontstaan?



- A Uit de harde schelp
- B Uit de versteende binnenkant
- C Uit een afdruk in de kalkrijke zeebodem

30 Reactiesnelheid

In de sport zijn reactietijd en reactiesnelheid belangrijke begrippen. Bij de atletiekbaan geldt de regel dat er een reactietijd van minimaal 100 milliseconde. Een snellere reactietijd is er sprake van een valse start.

Om de reactiesnelheid via het zenuwstelsel te meten, deed leraar René een test met zijn hele klas. De klas vormde een cirkel met alle leerlingen, waarbij iedereen twee burens een hand gaf. De armen in de cirkel waren daarbij gestrekt.

René mat de omtrek van de cirkel (meters): die was 60 meter. Julie kreeg het kneepje met haar rechterhand een kneepje in de linkerhand van haar rechterhand en die gaf het kneepje weer door aan de volgende, totdat Julie een kneepje in de linkerhand voelde. Zij moest bij de eerste kneep meteen 'start' zeggen en bij de laatste kneep 'stop'. René startte zijn stopwatch bij 'start' en drukte weer in bij 'stop'. Hij kreeg een tijd van 12 seconden.

Bereken de reactiesnelheid in m/s (meters/seconde).



31 Een knieblessure

Knieblessures komen vaak voor bij topsporters. Zo brak bij wielrenster Anne Vleuten tijdens de wereldkampioenschappen op de weg in 2018 een stuk van de bovenkant van het dijbeen achter de knieschijf af bij een val. Ze moest een operatie ondergaan om de knie te herstellen voor ze weer kon fietsen. Op de röntgenfoto zie je een breuk bij de knie van Annemiek. Een stukje bot is afgebroken. Het herstel duurde maanden. Pas daarna kon ze voorzichtig gaan bewegen.

Welk type gewricht vind je bij de knie?



- A Kogelgewricht
- B Rolgewricht
- C Scharniergewricht

32 Erfelijke variatie

Bij welk van de onderstaande huwelijken zal de variatie in erfelijk materiaal bij kinderen het kleinst zijn? De mannen en vrouwen in deze huwelijken komen uit families die al heel lang in de genoemde steden wonen.

- A Irma uit Groningen trouwt met haar neef Geert uit Maastricht.
- B Katja uit Assen trouwt met haar stadgenoot Bram.
- C Stine uit Amsterdam trouwt met Simon uit Kopenhagen.

33 Hardlopen bij dieren

Welke bewering over hardlopen bij een hert en een cheeta (jachtluipaard)

- A Een hert kan harder lopen dan een cheeta, maar is eerder uitgeput.
- B Een hert kan minder hard lopen dan een cheeta en is eerder uitgeput.
- C Een hert kan harder lopen dan een cheeta en is minder snel uitgeput.

D Een hert kan minder hard lopen dan een cheeta, maar is minder snel u

34 Smaakknopjes

De smaakknopjes in de tong kunnen vier smaken onderscheiden. Welke hoort daar **niet** bij?

- A Bitter.
- B Puur.
- C Zoet.
- D Zout.
- E Zuur.

35 Hartslagmeter

Kamiel traint met een hartslagmeter. Hij is een heel goed getrainde duursport- 'sporthart'.

Wat is de slagfrequentie van Kamiel in rust, voordat hij begint met trainen?



- A Ongeveer 40 slagen per minuut
- B Ongeveer 70 slagen per minuut
- C Ongeveer 100 slagen per minuut
- D Ongeveer 130 slagen per minuut

36 Vrije trap

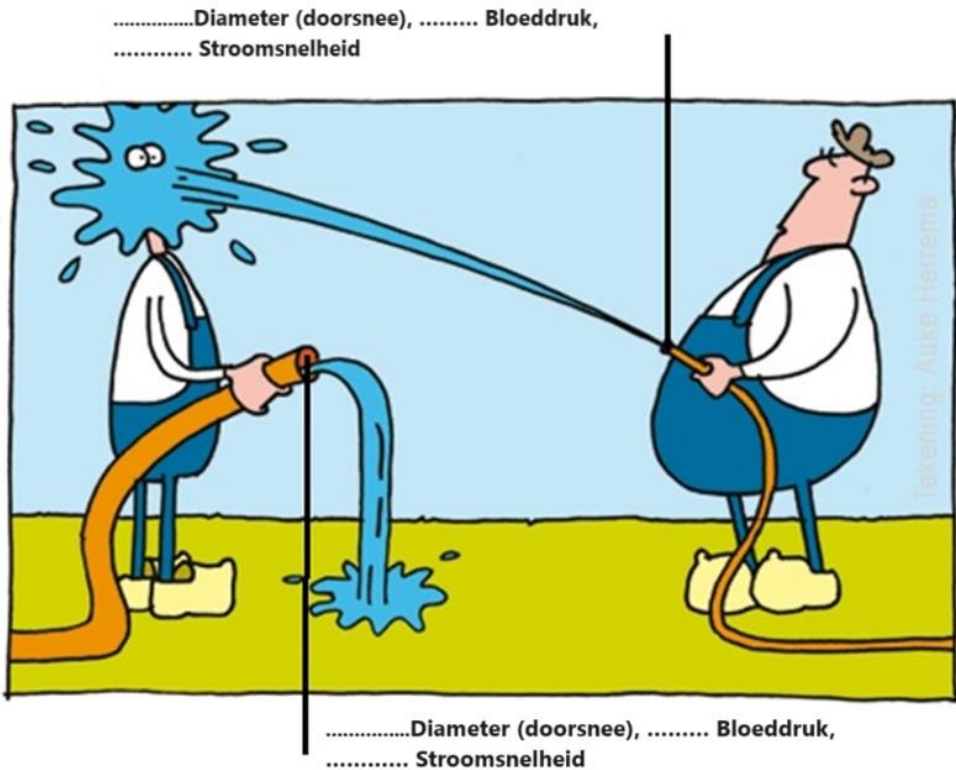
Als een voetballer een vrije trap neemt, verandert er iets aan de spieren en

van zijn been.
Wat verandert er?



- A De spieren trekken samen en de pezen worden korter.
- B De spieren worden korter en de pezen trekken samen.
- C De spieren worden korter en de pezen blijven even lang.
- D De pezen trekken samen en de spieren blijven even lang.

- 37** Bekijk de afbeelding, die gaat over de bloedstroom in een smal of breed vat, wat er moet staan op de stippelijntjes.
Boven: bij diameter (doorsnee van het vat): groot of klein, bij bloeddruk hoog of laag, bij stroomsnelheid: hoog of laag.
Onder: bij diameter (doorsnee) van het vat: groot of klein, bij bloeddruk hoog of laag, bij stroomsnelheid: hoog of laag.





38 Oogkleur 1/5


Bij de volgende vragen mag je voor de oogkleur uitgaan van twee genen, met twee mogelijkheden (allelen).


Er is een gen voor bruine (B) of blauwe (b) ogen. En er is een gen voor groen (G) of blauw (g) ogen. Lastig is dat het dominante bruine allel van het eerste gen het dominante groene allel uit het tweede gen overheerst, zodat groen alleen tot het fenotype bij een homozygoot recessief genotype voor het eerste gen (bb) komt. Dit geeft de mogelijkheden die je ziet in de bron.


 **BB GG** : bruine ogen


 **BB Gg** : bruine ogen


 **BB gg** : bruine ogen


 **Bb GG** : bruine ogen

 **Bb Gg** : bruine ogen

 **Bb gg** : bruine ogen

 **bb GG** : groene ogen

 **bb Gg** : groene ogen

 **bb gg** : blauwe ogen

39 Oogkleur 2/5

Bij de onderstaande kruising is alleen het fenotype van de ouders bekend genotype:

  x   (groen x bruin)

Welke bewering kan of welke beweringen kunnen juist zijn?

- A Uit deze kruising kunnen kinderen met blauwe ogen ontstaan, die homozygoot zijn voor beide genen.
- B Uit deze kruising kunnen kinderen met bruine ogen ontstaan, die homozygoot zijn voor beide genen.

- C Uit deze kruising kunnen kinderen met groene ogen ontstaan, die homozygoot zijn voor beide genen.

40 Oogkleur 3/5

Bij de onderstaande kruising is het fenotype van de ouders bekend (en dus genotype), maar wel dat beide ouders homozygoot zijn voor beide genen:



Welke bewering is juist?

- A Uit deze kruising kunnen kinderen met blauwe ogen ontstaan, die heterozygoot zijn voor beide genen.
- B Uit deze kruising kunnen kinderen met bruine ogen ontstaan, die heterozygoot zijn voor beide genen.
- C Uit deze kruising kunnen kinderen met groene ogen ontstaan, die heterozygoot zijn voor beide genen.

41 Oogkleur 4/5

Bij de onderstaande kruising is het genotype van de ouders bekend:

Bb GG x Bb gg

Hoe groot is de kans op kinderen met bruine ogen die homozygoot zijn voor beide genen?

- A 0%
- B 25%
- C 50%
- D 75%

42 Oogkleur 5/5

Bij de onderstaande kruising is het genotype van de ouders bekend:

Bb Gg x Bb Gg

Hoe groot is de kans op kinderen met blauwe ogen?

- A 0%
- B 6% (1/16)

C	19% (3/16)
D	25% (1/4)
E	50% (1/2)
F	56% (9/16)

Antwoorden

- | | |
|-----------|--|
| 1 | D (2 punten) |
| 2 | 1-6-4-2-3-5.

Per fout of verwisseling 1p aftrek.
(2 punten) |
| 3 | successie (1 punten) |
| 4 | Van de schelpen van dode schelpdieren uit zee.
(1 punten) |
| 5 | D (2 punten) |
| 6 | D (2 punten) |
| 7 | A (1 punten) |
| 8 | A (2 punten) |
| 9 | C (1 punten) |
| 10 | B (2 punten) |
| 11 | A (1 punten) |
| 12 | C (1 punten) |
| 13 | plat (1 punten) |
| 14 | A (2 punten) |
| 15 | C (2 punten) |
| 16 | F (2 punten) |
| 17 | Dit deel heeft als functie het in zich opnemen van een bevruchte eicel. 1p
Als er geen eicel is bevrucht, heeft dit deel geen functie meer. 1p
(2 punten) |
| 18 | D (2 punten) |
| 19 | B (2 punten) |
| 20 | Kans dat Julia heterozygoot is: $\frac{2}{3}$. Voor Romeo geldt hetzelfde. 1p
Kans dat twee heterozygoten een kind met de ziekte krijgen is : $\frac{1}{4}$. 1p |

- Dus $2/3 \times 2/3 \times 1/4 = 4/36 (= 1/9)$ 1p
(3 punten)
- 21** A B C D (4 punten) (1 strafpunt per fout)
- 22** D (2 punten)
- 23** De maag geeft het voedsel niet goed door aan de darmen. 1p
Daardoor kunnen de darmen het voedsel niet goed verteren. 1p
(2 punten)
- 24** A (2 punten)
- 25** C (2 punten)
- 26** C (2 punten)
- 27** C (2 punten)
- 28** C (2 punten)
- 29** A (1 punten)
- 30** $60/12 = 5$ m/s
(1 punten)
- 31** C (1 punten)
- 32** A (1 punten)
- 33** D (2 punten)
- 34** B (2 punten)
- 35** A (2 punten)
- 36** C (2 punten)
- 37** Boven: klein, hoge, hoog 1p
Onder: groot, lage, laag 1p
(2 punten)
- 38** --- (infoscherm)
- 39** A C (2 punten) (1 strafpunt per fout)
- 40** B (1 punten)
- 41** A (2 punten)
- 42** B (2 punten)

