

# 5

## HET AFNEMEN VAN TOETSEN MET DE COMPUTER

Silvester Draaijer, Janneke van der Hulst en Rick de Graaff

Wat zijn de voordelen van toetsen afnemen via een computer?

In welke situaties is het voor een docent haalbaar om de computer te gebruiken?

Welke mogelijkheden bieden beschikbare toetsprogramma's?

Hoe kan een docent kiezen uit de beschikbare programma's?

- 5.1 Twee voorbeelden van toetsafname met de computer
- 5.2 De computer als toetsmachine
  - 5.2.1 Afsluitende tentamens afnemen met de computer
  - 5.2.2 Formatieve toetsen afnemen met behulp van de computer
  - 5.2.3 Feedback
- 5.3 Itembanken
  - 5.3.1 Goed gebruik van een itembank
  - 5.3.2 Itembanksystemen in soorten en maten
- 5.4 Toetsprogramma's
  - 5.4.1 Stand van zaken en ontwikkelingen
  - 5.4.2 Mogelijkheden van de toetsprogramma's
  - 5.4.3 De keuze voor een toetsprogramma
  - 5.4.4 Nakijken van antwoorden in tekstvorm
- Samenvatting



## 5

Dit hoofdstuk behandelt het gebruik van de computer bij zowel tentamens als bij oefentoetsen en formatieve toetsen. Het doel van het hoofdstuk is inzicht te bieden in de mogelijkheden en beperkingen van de computer bij toetsing. De nadruk ligt op relatief eenvoudig te gebruiken computerondersteunde toetssystemen en bevat praktische adviezen voor docenten die computerondersteund toetsmateriaal willen ontwikkelen of selecteren. Veel aandacht in dit hoofdstuk gaat uit naar de mogelijkheden om via de computer directe feedback te geven.

### 5.1 Twee voorbeelden van toetsafname met de computer

#### PAK-toetsing bij het VUmc

Bij de opleiding tot basisarts aan het VUmc (Vrije Universiteit medisch centrum) is met ingang van academisch jaar 2005/2006 een nieuwe toetsvorm geïntroduceerd. Zo'n driehonderd studenten maken bij elk blok, in dertig minuten, een zogenaamde Parate Kennistoets (PAK). Die toets bestaat uit veertig vragen waar studenten direct het antwoord op moeten geven. Een student moet minimaal negentig procent van de vragen correct beantwoorden om voor de toets te slagen. Belangrijk is het gebruik van afbeeldingen en het meteen terugkoppelen van het resultaat naar de studenten. De toets is cumulatief, dat wil zeggen dat studenten bij elk volgende blok ook de PAK-vragen van de voorgaande blokken moeten kunnen beantwoorden. Met andere woorden, aan het eind van bijvoorbeeld blok 12 krijgen studenten at random veertig vragen te beantwoorden uit de blokken 1 tot en met 12. Gedurende het blok kan de student onbeperkt oefenen op de hele vragenbank, via het internet, dus tijd- en plaatsonafhankelijk. Alle vragen zijn in één systeem ondergebracht.

Vanwege de behoefte om met afbeeldingen te werken, de directe terugkoppeling en het cumulatieve karakter van de PAK, koos het VUmc voor het gebruik van een toetsvragenbank en afname van de toets via het beeldscherm. De randvoorwaarden daarvoor waren gunstig: er was een centraal toetssysteem op de VU beschikbaar (Questionmark Perception) en een pc-zaal met ruim 140 werkstations op de faculteit zelf. Binnen een onderwijsvernieuwingproject zijn de mogelijkheden en consequenties uitgezocht en is het gelukt alle betrokkenen (bestuur, docenten, examencommissies, QMP-beheerder, systeembeheerders, surveillanten, zaalbeheerders) op één lijn te krijgen. Hierdoor werden alle financiële, technische en organisatorische randvoorwaarden gerealiseerd. In september 2005 is de eerste toets met succes afgenomen. Sinds die tijd is de PAK als beeldschermtoets een regulier onderdeel in het toetsprogramma.

### Toepassing Questionmark Perception bij de International Business School (IBS) van de Hanzehogeschool<sup>1</sup>

De docente Engels stond in 1998 voor een dilemma. 'Ik geef les op onder andere de Duits-talige IBS-opleiding. De beheersing van de grammatica is bedroevend. Gemiddeld veertien onvoldoendes van een totaal van twintig op een toets. Echt zware onvoldoendes. En de cijfers van de voldoende lagen tussen de 7,5 en 9,5. Het lesgeven aan zo'n groep verloopt erg onprettig omdat de zwakke studenten het niet meer zien zitten, terwijl de betere studenten zich gaan vervelen door gebrek aan tempo.'

Zij koos daarom voor een andere strategie. Zij bracht het aantal contacturen voor grammatICA terug van twee naar één en in de vrijgekomen tijd ging zij computerondersteund aan de slag met Questionmark Perception. Via dat programma legde de docente de studenten een groot aantal verschillende (oefen)toetsvragen voor over de stof. Per vraag kreeg de student specifieke feedback op allerlei mogelijke foute (en ook correcte) antwoorden. Ze paste de vragen steeds aan met behulp van de 'rapportage'functie die zowel individuele antwoorden als een verzameling van antwoorden van een groep studenten kan laten zien. Een ander voordeel vond ze dat ze gelijktijdig op verschillende niveaus met de studenten kon werken. 'Als de studenten niet voldoende scores, gaan ze een stapje terug, als ze heel hoog scoren slaan ze een deel over.' Momenteel haalt twee derde van haar studenten een voldoende voor grammatica. Dat is een verdubbeling ten opzichte van vijf jaar geleden. 'Maar de tijd die je overhoudt door minder contacturen weegt niet op tegen de tijdsinvestering die nodig is om een goed draaiend testprogramma te maken', waarschuwt zij. 'Ik heb het in mijn eigen tijd ontwikkeld, hetgeen nu eigenlijk ondenkbaar is geworden. De hogeschool zou docenten daarom de ruimte moeten geven iets te ontwikkelen. En het is belangrijk dat je iets ontwikkelt dat je kunt delen met collega's.'

Zoals uit deze voorbeelden blijkt, staat het *ontwikkelen en afnemen* van toetsen op de computer hier centraal. Uiteraard is de computer ook in andere fasen van het toetsproces een belangrijk hulpmiddel, zoals bij het verwerken en analyseren van tentamenresultaten (zie hoofdstuk 3).

## 5.2 De computer als toetsmachine

Bij het afnemen van toetsen met de computer vindt directe interactie plaats tussen de student en de computer. De student werkt (meestal individueel) aan de computer, die reageert door bijvoorbeeld een bepaalde vervolgvraag aan te bieden of door

<sup>1</sup> <http://www.hanze.nl/home/Over+de+Hanzehogeschool/Nieuws/hanzekrant/2003-02-06duitsers.htm>.

feedback te geven op het antwoord van de student. Bij dit interactief afnemen van toetsen is het zinvol om toetsen met verschillende doelen te onderscheiden. Meestal wordt onderscheid gemaakt tussen toetsen met een summatief doel (tentamineren) en toetsen met een formatief doel (zie hoofdstuk 1 voor dit onderscheid). Bij zowel afsluitende tentamens als bij formatieve (oefen)toetsen biedt het gebruik van de computer voordelen, maar er worden verschillende eisen gesteld.

Als het gaat om het afnemen van afsluitende *tentamens* via de computer worden er onder andere zwaardere eisen gesteld aan de beveiliging en aan de bedrijfszekerheid van de programmatuur. Bij afsluitende tentamens staat er immers voor de student én voor de opleiding veel op het spel. Het mag niet voorkomen dat computers niet goed werken, dat uitslagen zoekraken of dat de integriteit van de data wordt aangetast. Vaak zijn de voordelen van de inzet van de computer voor tentamens alleen te realiseren bij een bepaalde schaalgrootte. Het vereist in de praktijk een bredere inspanning dan die van individuele docenten alleen (zie het eerste voorbeeld in de inleiding).

Het afnemen van *formatieve toetsen* stelt minder hoge eisen. Bij deze toetsen is de oefenfunctie, de bijsturing van het leerproces en de zelfbeoordeling belangrijker dan de beoordeling en selectie van studenten. Bij deze toetsen is vooral de mogelijkheid van het geven van directe feedback van belang. Op dit moment zijn er al veel goede mogelijkheden om met de computer oefentoetsen en formatieve toetsen af te nemen. Deze mogelijkheden liggen binnen het bereik van steeds meer docenten, bijvoorbeeld wanneer de toetsmogelijkheden van een digitale leeromgeving worden gebruikt.

### 5.2.1 AFSLUITENDE TENTAMENS AFNEMEN MET DE COMPUTER

Er zijn voordelen aan het afnemen van tentamens met behulp van de computer. Studenten krijgen via de computer direct een beoordeling voor hun toets, zonder tussenkomst van een docent. Dat is prettig voor de student en scheelt de docent nakijktijd. In tegenstelling tot pen-en-papier tentamens kan ook multimediaal materiaal (foto's, grafieken, geluid, video, animaties, simulaties) in een toets worden opgenomen. Als er eenmaal een set van voldoende omvang met degelijke vragen beschikbaar is, wordt het mogelijk om hieruit steeds opnieuw tentamens samen te stellen (zie paragraaf 5.3). Het *at random* kiezen uit de vragenbank en daarbij het variëren binnen een vraag in de volgorde van de antwoorden maken afkijken tijdens de toetsafname vrijwel onmogelijk.

Hoewel het afnemen van tentamens met behulp van de computer voordelen heeft, blijkt het een behoorlijke inspanning te kosten om deze voordelen echt te realiseren.

*Afnamecapaciteit*

In principe kunnen studenten via het internet vanaf iedere willekeurige plaats het tentamen maken, maar in de praktijk zijn er nog (te) veel problemen met de beveiliging (zoals de controle of studenten wel degenen zijn voor wie ze zich uitgeven en het voorkómen dat studenten ongeoorloofde bronnen raadplegen). Het afnemen van tentamens via het internet gebeurt dan ook nog nauwelijks. Als er tentamens worden afgenomen op de computer vindt dat plaats binnen de instelling of eventueel op een vaste toetslocatie elders. Dit betekent dat de computerfaciliteiten binnen de instelling geschikt moeten zijn voor het afnemen van tentamens (denk aan de inrichting en capaciteit van computerzalen, het beveiligen van het netwerk, het installeren van software). Uiteraard vereist dit beslissingen die het bereik van de individuele docent te boven gaan, het management zal er initiatief voor moeten nemen.

*Vragenbank*

Willen de voordelen van vragenbestanden goed tot hun recht komen, dan vereist dat een ruime hoeveelheid vragen, voorzien van relevante meta-informatie (bron, auteur, onderwerp, eerder gebleken moeilijkheidsgraad en kwaliteit). Het ontwikkelen van een dergelijk vragenbestand is secuur werk en is bij voorkeur geen zaak van een individuele docent, maar moet afdelingsbreed of zelfs opleidingsbreed worden aangepakt.

*Vraagtype*

Als de vragen automatisch moeten worden nagekeken, kunnen er eigenlijk alleen gesloten vraagtypen gebruikt worden. Een computerondersteund tentamen voor het toetsen van bijvoorbeeld probleemoplossen of reflectieve vaardigheden vraagt daarvoor speciaal ontworpen programma's en uitgekende protocollen.

*Beleid*

Het invoeren van het toetsen achter de computer vraagt een management dat dit tot speerpunt maakt van het beleid.

In het kader staan aandachtspunten voor docenten die overwegen te tentamineren met behulp van ICT. Hoe meer de toetssituatie voldoet aan deze punten, des te groter is de kans dat toetsen met ICT een succes wordt.

**Aandachtspunten bij het overwegen van tentamineren met ICT***Commitment*

- Vindt het management of de onderwijsdirectie de inzet van ICT voor toetsing belangrijk en is het bereid extra tijd te geven voor pilots en het ontwikkelen van expertise?

- Vindt het systeembeheer de inzet van ICT voor toetsing belangrijk, wil het meewerken en heeft het gelegenheid om daar extra tijd en energie aan te besteden?
- Vindt het roosterbureau de inzet van ICT voor toetsing belangrijk en heeft het zeggenschap over de inroostering van pc-zalen?
- Zijn er voldoende collega's die de inzet van ICT voor toetsing belangrijk vinden? Vallen er afspraken te maken over werkverdeling en beheer van vragenbestanden? Hebben de docenten hier tijd voor en zijn zij gemotiveerd?

#### *Infrastructuur*

- Is er voor iedere student die moet worden getoetst een pc beschikbaar? Indien dat niet zo is, is het dan mogelijk om studenten bijvoorbeeld in twee sessies direct achter elkaar te toetsen?
- Indien gebruik wordt gemaakt van meer pc-zalen, is er dan bereidheid meer surveillanten in te zetten?
- Zijn de pc's geschikt om de software waarmee het tentamen wordt afgenomen vlot en zonder problemen te draaien?

#### *Cursus*

- Studenten moeten tijdens de leerfase kunnen oefenen met de manier van tentamineren; biedt het cursusprogramma ruimte om studenten te laten oefenen?
- Is er een locatie beschikbaar of een website waar studenten hun resultaten kunnen inzien?

#### *Kosten en baten*

- Is er sprake van een vak/blok/cursus dat minimaal een aantal jaren wordt gegeven?
- Is de inhoud van het onderwijsonderdeel zo stabiel dat de vragen minimaal een aantal jaren bruikbaar blijven?
- Gaat het om grote groepen studenten?

(naar: Magné, 1991)

### 5.2.2 FORMATIEVE TOETSEN AFNEMEN MET BEHULP VAN DE COMPUTER

Formatieve toetsen zijn vooral bedoeld als leermiddel en kunnen in twee hoofdcategorieën worden ingedeeld: diagnostische toetsen en oefentoetsen.

*Diagnostische toetsen* zijn bedoeld om de student en docent informatie te geven over het verloop van het leerproces ('diagnostische functie'). Aan studenten geeft een dergelijke formatieve toets inzicht in de mate waarin ze (onderdelen van) de leerstof beheersen. Een groot voordeel van toetsafname met de computer is dat studen-

ten *snel* feedback krijgen op hun prestaties. Hierdoor is het leereffect groter dan wanneer de feedback pas na verloop van tijd wordt gegeven. Bovendien motiveert het studenten wanneer ze feedback krijgen op hun antwoorden. De formatieve toets helpt de docent na te gaan hoe de studenten (individueel en als groep) presteren, zodat de docent het onderwijs kan bijsturen en hij weet welke studenten problemen hebben met de leerstof. Dit vooronderstelt registratie van de antwoorden van studenten. Feedback is vaak gericht op bijspijkeren van voorkennis of advies over tentamenvoorbereiding.

*Oefentoetsen* zijn bedoeld om studenten te helpen zich de leerstof eigen te maken ('oefenen met de stof'). Deze toepassing biedt studenten de gelegenheid om gericht te oefenen met de leerstof, waarbij ze leren van de feedback. Studenten die moeite hebben met de stof kunnen individueel aan de computer werken zodat ze extra oefening krijgen; de oefentoets krijgt dan een remediërend karakter. Feedback bevat vaak aanwijzingen voor verdere studie.

Afhankelijk van het karakter van de toets (diagnostische toets of oefentoets) worden andere eisen gesteld aan de toets. Zie hiervoor het kader.

#### **Aanwijzingen voor het maken van interactieve diagnostische toetsen en oefentoetsen**

Als de toets een diagnostische functie heeft:

- moet de inhoud van de formatieve toets representatief zijn voor het niveau van de eindtoets en de leerdoelen goed dekken;
- moet de feedback zowel een totaalscore bevatten als informatie over de beheersing van de verschillende onderdelen van de leerstof om antwoord te kunnen geven op de vraag aan welke onderdelen de student extra aandacht moet besteden;
- wordt de feedback gegeven nadat de formatieve toets geheel is gemaakt;
- moet de registratie zowel gegevens per student als over de hele groep bevatten; eventueel moeten resultaten kunnen worden weggeschreven naar het studentvolgsysteem van de opleiding.

Als de toets bedoeld is om met de leerstof te oefenen:

- ontwerp dan oefeningen die eventueel eenvoudiger zijn dan de eindtoets (bijv. leren toepassen van sleutelconcepten uit de leerstof) en maak de studenten duidelijk dat het niveau niet representatief is voor de eindtoets;
- maak dan vooral vragen over lastige onderdelen van de leerstof;
- maak dan een keuzemenu zodat studenten kunnen kiezen voor vragen over bepaalde onderwerpen;
- maak dan meer vragen per onderwerp en geef bij voorkeur feedback na iedere vraag zodat studenten snel van hun fouten kunnen leren;



- plaats dan de vragen in oplopende volgorde van moeilijkheid zodat studenten niet direct stuklopen;
- maak dan gebruik van een variatie aan vraagtypen en van multimediaal materiaal om de student extra te motiveren om met het materiaal aan de slag te gaan;
- voorzie dan de eerste opgaven van aanwijzingen of suggesties in welke richting de oplossing moet (kan) worden gezocht of verdeel de opgaven in kleine stappen die voorafgaan aan de uiteindelijke hoofdooplossing.

(naar: Keller, 1983; Van Berkel, 1999; Van Vliet, 2001)

Het gebruik van de computer bij diagnostische toetsen en oefentoetsen biedt duidelijk voordelen: snelle feedback, registratie en de gelegenheid voor individueel oefenen via het internet, dus tijd- en plaatsafhankelijk. Het ontwikkelen van degelijke toetsen (met of zonder computer) vergt echter een behoorlijke inspanning van de docent. Het is daarbij niet de ‘techniek’ die de meeste tijd vraagt (met de software is meestal in ongeveer een dag te leren werken), maar het formuleren van degelijke vragen, antwoorden en feedback. Het maken van diagnostische toetsen en oefentoetsen komt dan ook vooral in aanmerking wanneer:

- er sprake is van een relatief stabiel vakgebied zodat de vragen een aantal jaren meekunnen;
- er sprake is van een flink aantal studenten dat het onderwijs volgt en er weinig gelegenheid is voor het geven van individuele begeleiding of feedback;
- studenten moeite hebben met de leerstof of moeilijk kunnen bepalen of ze de leerstof adequaat beheersen;
- de leerstof en leerdoelen zich lenen voor gesloten vraagtypen; als het bijvoorbeeld gaat om het formuleren en onderbouwen van meningen zijn open vraagvormen geschikter.

Omdat het maken van diagnostische toetsen en oefentoetsen een flinke tijdsinvestering vraagt, is het van belang om studenten te stimuleren er ook daadwerkelijk gebruik van te maken. Het blijkt dat alleen het beschikbaar stellen van toetsen niet voldoende is om te zorgen dat studenten ze ook maken (en misschien maken juist de studenten die de oefening het meeste nodig hebben er het minste gebruik van!). Het gebruik wordt gestimuleerd door bijvoorbeeld een bijeenkomst in te roosteren waarop studenten met deze toetsen werken, of studenten de diagnostische toets of oefentoets als huiswerkopdracht te geven. Uiteraard is het mogelijk om op een of andere manier een beloning te geven voor het maken van de toets, zoals een bonus of afrondingsvoordeel bij de eindtoets, of deelname aan de toets als ingangseis voor het tentamen. Daarmee worden wel de eisen aan de beveiliging van de toetsen weer strenger.

**Hoeveel tijd kost het de docent om een oefentoets via de pc aan te bieden?**

Hieronder staat een eenvoudig rekenvoorbeeld voor het opzetten en uitvoeren van een oefentoets met vierkeuzevragen. Uitgangspunt is dat alle technische voorzieningen aanwezig zijn, dat de docent niet intensief hoeft te overleggen met roosterbureaus en systeembeheerders.

De docent wil de studenten tweemaal een toets van een halfuur aanbieden op een ingeroosterd tijdstip in een computerzaal. Elke oefentoets bevat twintig, voornamelijk tekstuele vragen. Alle studenten krijgen dezelfde toets (de docent maakt géén gebruik van uitgebreide itembankmogelijkheden). In het voorbeeld wordt ervan uitgegaan dat het bedenken van een eerste versie van één vraag met feedback ongeveer een halfuur in beslag neemt.

- Leren omgaan met software (over alle stappen verspreid): 8 uur.
- Bepalen van de leerstof waarover de vragen gaan, bijvoorbeeld met een toetsmatrijs: 4 uur.
- Formuleren van 40 vragen met feedback:  $40 \times 0,5$  uur = 20 uur.
- Toets bespreken met collega (eventueel laten maken): 4 uur.
- Aanpassen vragen, herformuleren, aanpassen feedback:  $40 \times 0,2$  uur = 8 uur.
- Tweemaal afnemen van de toets door studenten inclusief voorbereidingstijd en nazorgtijd:  $2 \times 2$  uur = 4 uur.
- Verwerken van resultaten, direct vragen aanpassen en feedback: 4 uur.
- Alvast aanmaken van 10 nieuwe vragen voor volgende periode:  $10 \times 0,5$  uur = 5 uur.

Totaal: 57 uur.

### 5.2.3 FEEDBACK

Feedback kan verschillende vormen hebben:

- uitsluitend aangeven of een vraag correct of fout beantwoord is;
- aangeven of de student al voldoet aan de tentameneisen;
- inhoudelijk uitleggen waarom een vraag correct of fout beantwoord is;
- studeeraanwijzingen of hints geven waarmee de student de vraag opnieuw kan proberen te beantwoorden.

De feedback blijft bij formatieve toetsen en oefentoetsen bij voorkeur niet beperkt tot alleen een cijfer of percentage correct beantwoorde vragen. Vooral bij oefentoetsen is het belangrijk dat er bij iedere vraag inhoudelijke feedback wordt gegeven. Feedback moet inhoudelijk voldoende specifiek en gedetailleerd zijn, zo mogelijk studeeraanwijzingen voor de student bevatten of een verwijzing naar (digitaal) studiemateriaal waarin studenten de stof nog eens kunnen bestuderen.

Anders dan bij toetsen die schriftelijk worden afgenomen, moet bij computerondersteund toetsen de feedback vooraf worden geformuleerd. Er moet dus worden geanticipeerd op mogelijke correcte en foute antwoorden. In theorie is het mogelijk om voor ieder mogelijk antwoordalternatief specifieke feedback te formuleren, maar in de praktijk blijkt dit lastig. Bij vragen op kennisniveau is er vaak niet zoveel verschil tussen feedback op verschillende foutieve antwoorden, terwijl het bij vragen op inzichtniveau moeilijk is te voorspellen welke (redeneer)fout de student heeft gemaakt. Meestal kan men zich daarom beperken tot twee soorten feedback (bij correct respectievelijk fout antwoord) of tot een modelantwoord (bij open vragen).

Als het wél belangrijk is om de feedback verder te differentiëren, is het aan te raden om de feedback niet automatisch te koppelen aan een specifiek antwoordalternatief, maar om de student zelf uit een menu te laten kiezen voor passende feedback, zoals uitleg, een voorbeeld, een verwijzing of een hulpvraag. De student wordt dan gedwongen nog na te denken over de leerstof. Hier kan een positief leereffect van uitgaan. Studenten zijn meestal goed in staat om te beoordelen waarom ze het correcte antwoord niet konden geven, om daar vervolgens de meest passende feedback bij te kiezen. Overigens bestaat deze mogelijkheid alleen in meer geavanceerde toetsprogramma's. In de meeste eenvoudige programma's is de mogelijkheid voor feedback beperkt tot twee feedbackteksten: feedback bij een correct antwoord op de vraag en feedback die getoond wordt bij een incorrect antwoord.

Feedback op een beeldscherm kan confronterend zijn. Feedback moet dan ook zorgvuldig geformuleerd worden. De tekst in het kader bevat enkele aanwijzingen voor het formuleren van inhoudelijke feedback.

#### **Aanwijzingen bij het geven van feedback**

- Geef pas feedback wanneer zeker is dat de student een oprechte poging heeft gedaan om een antwoord te formuleren. Werp bijvoorbeeld een drempel op door feedback te geven:
  - in de vorm van een nieuwe vraag;
  - in de vorm van een studeeraanwijzing als de vraag niet correct is beantwoord;
  - pas als minimaal 80% van de vragen correct is beantwoord;
  - pas als het antwoord correct is.
- Formuleer feedback beknopt.
- Houd de toon van feedback positief/neutraal; feedback bij verkeerd beantwoorde vragen mag nooit neerbuigend zijn.
- Geef bij correct beantwoorde vragen aan waarom het gekozen antwoord correct is en waarom de andere antwoorden fout zijn; een student kan door gokken het correcte antwoord hebben gekozen.

- Geef bij gedeeltelijk correct of fout beantwoorde vragen aan wat het juiste antwoord is, waarom dat het juiste antwoord is, waarom het door de student gekozen antwoord fout is en waarom de andere antwoorden fout zijn.
- Lange antwoorden op open vragen kunnen niet automatisch worden nagekeken, feedback kan hier worden gegeven in de vorm van een antwoordmodel waarmee de student het eigen antwoord kan vergelijken. Geef hierbij ook een beoordeling aan ('Als je ten minste drie van deze punten hebt genoemd, is de vraag voldoende beantwoord.').
- Zorg dat de vraag nog te zien is als de feedback verschijnt, herhaal zo nodig de vraag in de feedback.
- Feedback is meestal tekstueel maar soms is een andere vorm duidelijker, zoals een afbeelding.
- Gebruik geen speciale effecten (zoals geluid of humoristische afbeeldingen) in feedback, dat kan studenten afleiden of irriteren.
- Geef studenten zo mogelijk de gelegenheid te kiezen uit verschillende soorten feedback, aansluitend bij verschillende fout- of redeneertypen.
- Behalve inhoudelijke informatie kan feedback ook informatie geven voor strategieën of suggesties om verder te gaan. Soms is een hint waarmee een student zelf de oplossing kan zoeken, zeer leerzaam. Uiterlijk na de derde poging moet het correcte antwoord worden gegeven en/of moet de student verder kunnen met de volgende vraag.
- Alle informatie die nodig is om de vraag te beantwoorden, moet in de vraag staan (het is frustrerend voor studenten als noodzakelijke informatie pas achteraf wordt gegeven in de feedback).

(naar: Fenrich, 1997; Rayne, 2002)

### 5.3 Itembanken

Sommige toetsprogramma's bieden de mogelijkheid om een gestructureerd bestand van vragen aan te leggen, waaruit eenvoudig toetsen kunnen worden samengesteld. Zulke toetsprogramma's worden ook vaak aangeduid met de term *itembanksystemen*. De itembankfunctie is een hulpmiddel bij het beheer van de vragen (eenmaal aangemaakte vragen kunnen eenvoudig opnieuw worden gebruikt en aangepast) en bij het samenstellen van toetsen op basis van bijvoorbeeld onderwerp of moeilijkheidsgraad.

Een itembank bestaat uit items (vragen en de bijbehorende antwoordmogelijkheden of een antwoordmodel) en informatie over die items, zoals het onderwerp waarbij de vraag hoort, het cognitieve niveau van de vraag, wat voor type vraag het is, wanneer de vraag voor het laatst in een toets is opgenomen. Welke informatie over de vragen in de itembank kan worden opgenomen, ligt vast in de structuur van de itembank. Voordat een itembank wordt gevuld, moet dus eerst worden besloten

welke informatie over de items moet worden vastgelegd. Het is belangrijk goed na te denken over de structuur van de itembank, omdat aan de hand van deze gegevens later de toetsen kunnen worden samengesteld. Er zijn vele indelingen van vraagkenmerken mogelijk (zie het kader).

### Mogelijke indelingen van vraagkenmerken

#### *Hiërarchische structuur*

- op basis van toetsmatrijs: toets ▶ hoofdonderwerpen ▶ subonderwerpen
- boek ▶ hoofdstuk ▶ paragraaf ▶ bladzijde
- vakken ▶ hoofdthema's ▶ onderwerp ▶ deelonderwerp
- curriculum ▶ studiefase ▶ semester ▶ blok
- rollen ▶ competenties ▶ doelen ▶ trefwoorden.

#### *Categorisering*

- kennisvraag, toepassingsvraag, inzichtsvraag
- meerkeuzevraag, invulvraag, numerieke vraag, aanwijsvraag, juist/onjuistvraag, parallelvraag, casusvraag
- tentamenvraag, oefenvraag, voortgangsvraag
- overige gegevens: datum afname, RIT- of RIR-waarde bij laatste afname (dit is een indicatie voor de bijdrage aan de toetsbetrouwbaarheid), trefwoord, auteur, bron, moeilijkheidsgraad, geschatte beantwoordingsduur, afhankelijkheid ten opzichte van andere vragen.

(naar: Van Dijk, 1991; Dopper, 2000; Teelen, 2000)

In de praktijk blijkt het ontwerpen van een itembank moeilijk te zijn door de heterogeniteit van curricula, cursussen, toetsvormen en vraagvormen die moet worden ondersteund. De meest werkbare indeling lijkt die op basis van toetsmatrijzen. Immers, als duidelijk is op basis van welke criteria toetsen worden samengesteld (binnen een curriculum, fase of cursus), dan kan van daaruit een duidelijk verband worden gelegd met de indeling van de itembank.

Wanneer een goed gevuld vragenbestand beschikbaar is, kan een toets worden samengesteld. Er kan bijvoorbeeld een toets worden gemaakt bestaande uit een bepaald aantal vragen, met een bepaalde verdeling over onderwerpen en cognitieve niveaus. Het hanteren van een toetsmatrijs is hierbij aan te bevelen (zie hoofdstuk 2). Ook is het mogelijk om vragen en antwoordalternatieven in willekeurige volgorde aan te bieden tijdens een toets, waardoor studenten verschillende versies van dezelfde toets voor zich krijgen. Hierdoor neemt de kans op fraude af of kunnen studenten op verschillende tijdstippen toch dezelfde toets maken.

### 5.3.1 GOED GEBRUIK VAN EEN ITEMBANK

Goed gebruik van een itebank vraagt regelmatig tijd en aandacht. Het aanleggen van een vragenbestand is geen eenmalige gebeurtenis. Dikwijls kan een vliegende start worden gemaakt door al beschikbare vragen in te voeren. Vervolgens groeit de itebank in de loop van de tijd. Met enige regelmaat zullen vragen moeten worden aangepast of verbeterd. Dat kan bijvoorbeeld blijken uit de opmerkingen van studenten na afloop van een toets of uit opmerkingen van collega's die meewerken aan het opzetten en onderhouden van het vragenbestand. Het kan ook blijken uit statistische informatie die verkregen is door middel van een toets- en itemanalyse. In de meer geavanceerde itebanksystemen zijn deze analysefuncties meestal ingebouwd.

Als de itebank goed wordt bijgehouden, komen de voordelen ervan pas echt goed tot hun recht: de itebank is een hulpmiddel om eenvoudiger samen met collega's te werken aan een vragenbestand, en om inzicht te ontwikkelen op de totale hoeveelheid vragen, de kwaliteit ervan en de verdeling over de leerstof.

Overigens leidt een bestand met goede vragen niet automatisch tot goede toetsen. Het is van belang om de selectie van vragen voor een toets zorgvuldig uit te voeren (zie hoofdstuk 2). Daarnaast moet de onderlinge afhankelijkheid van de vragen, als ze in één toets staan, worden gecontroleerd. Zeker als het bestand parallelvragen bevat (verschillende vragen over hetzelfde onderwerp) of casusvragen (een vraag bestaande uit een stam met daarbij deelvragen over verschillende onderwerpen), bestaat de kans dat er in een automatisch gegenereerde toets van elkaar afhankelijke vragen of antwoorden zitten (een vraag geeft bijvoorbeeld het antwoord weg op een andere vraag).

### 5.3.2 ITEMBANKSYSTEMEN IN SOORTEN EN MATEN

Itbanksystemen bestaan in soorten en maten. De meer geavanceerde systemen, zoals *Examiner*, *TestVision* of *Questionmark Perception*, bestaan uit verschillende modules, zoals speciale modules voor vraag- en toetsconstructie, gebruikersbeheersystemen en item- en toetsanalyses. De voordelen van dit soort ingewikkelde itebanksystemen komen beter tot hun recht als een hele opleiding, faculteit of instelling er gebruik van maakt en niet alleen de individuele docent. De centrale aanschaf van één itebanksysteem voor een hele instelling lijkt dus voor de hand te liggen. In de praktijk blijkt dit alleen te werken als de faculteit of instelling een duidelijke gemeenschappelijke visie en aanpak heeft wat toetsen betreft (bijv. in de vorm van een gestandaardiseerde toetsvorm, een gestandaardiseerd toetsverwerkingssysteem of een gestandaardiseerde studievoortgangstoets). Bovendien moeten de te toetsen leerdoelen vergelijkbaar zijn; de programmatuur is vaak niet voor alle vakgebieden even geschikt.

Daarnaast bestaan er toetsprogramma's met eenvoudige itembankmogelijkheden die zijn te gebruiken door individuele of groepen docenten, zoals de toetsmodules in de digitale leeromgevingen *WebCT* of *Blackboard*. Vergelijkbare 'losse' programma's om zelf interactieve toetsen voor het internet te maken zijn bijvoorbeeld *Winttoets*, *Brownstone Diploma*, *MicroTest*, *Respondus* of *Hot Potatoes*. Deze programma's bieden niet alle in deze paragraaf beschreven mogelijkheden van itembanken, maar wel de mogelijkheid om toetsen te genereren uit een vragenbestand. In de volgende paragraaf ligt de nadruk op deze eenvoudiger software.

## 5.4 Toetsprogramma's

Er bestaan veel toetsprogramma's en er bestaan veel internetsites waar docenten zelf toetsen kunnen aanmaken. Vooral voor het maken van diagnostische toetsen of oefentoetsen zijn ze heel geschikt. De gebruiksvriendelijkheid van dit soort programmatuur is de laatste jaren sterk verbeterd.

### 5.4.1 STAND VAN ZAKEN EN ONTWIKKELINGEN

Een interessante ontwikkeling is dat steeds meer uitgevers en gespecialiseerde instituten kant-en-klare computeroefeningen en toetsen bij hun studieboeken aanbieden waarmee studenten hun begrip van de stof kunnen toetsen. Vaak kunnen die toetsen worden ingelezen in standaardtoetsomgevingen of in een digitale leeromgeving.

Het is vrij eenvoudig om toetsvragen uit te wisselen tussen systemen, mits alleen gebruik wordt gemaakt van eenvoudige typen vragen zoals meerkeuze of multiple respons. Deze uitwisseling vindt meestal plaats op basis van een internationale standaard voor vraaguitwisseling: Question and Test Interoperability (QTI). Het is handig om er bij de keuze voor een programma op te letten of het programma deze standaard ondersteunt. Als er echter meer geavanceerde vraagtypen worden gebruikt, blijkt de uitwisselbaarheid tussen programma's maar in beperkte mate mogelijk.

### 5.4.2 MOGELIJKHEDEN VAN DE TOETSPROGRAMMA'S

Verschillende toetsprogramma's bieden verschillende mogelijkheden. Deze paragraaf geeft aan de hand van een aantal hoofdkenmerken een globaal overzicht van de mogelijkheden van de programmatuur.

#### *Typen vragen*

Toetsprogramma's bieden meestal een selectie uit onderstaande typen aan:

- Open vragen

- aanvulvragen (ook wel gatenvragen of *fill in the blank*); deze kunnen automatisch worden nagekeken door ze te vergelijken met vooraf opgegeven correcte antwoorden
- essay-vragen; deze kunnen niet automatisch worden nagekeken (zie ook in paragraaf 5.4.4, Semi-automatisch nakijken van tekstdocumenten).
- Gesloten vragen (kunnen allemaal automatisch nagekeken worden)
  - één correct uit verschillende alternatieven (*multiple choice*)
  - een of meer correct uit alternatieven (*multiple response*)
  - rangordenen
  - paren zoeken (*matching*)
  - aanwijzen (*hot spot*)
  - numeriek antwoord
  - variabel numeriek antwoord
  - algebraïsch antwoord.
- Samengestelde vragen/itembank (met combinaties van bovenstaande typen)
  - casusvragen (een gegeven stamvraag waarover meerdere deelvragen worden gesteld).

Het is bij vrijwel alle programma's mogelijk om afbeeldingen, geluid, video of simulaties toe te voegen in zowel de vraag- als de antwoordalternatieven.

#### *Itembanking*

Met de meeste programma's kunnen vragenbestanden worden gemaakt waaruit toetsen worden gegenereerd. De mate waarin deze functie is uitgewerkt in verschillende programma's loopt echter zeer uiteen. De toetsprogramma's verschillen sterk in structurering en categorisering. Wanneer het toetsen met de computer op een grotere schaal wordt ingezet, moeten de itembankmogelijkheden van het programma uitgebreider zijn.

#### *Gedifferentieerde feedback, scoring, hints en adaptiviteit*

De meeste programma's kunnen feedback of scores bepalen voor de hele toets en per vraag. Sommige kunnen feedback per antwoordalternatief verwerken. De feedback bestaat vaak alleen uit tekst; andere programma's kunnen ook plaatjes of animaties aanbieden. In sommige programma's is het mogelijk de student alternatieve vragen aan te bieden afhankelijk van de gegeven antwoorden (*branching*) of studenten de keuze geven om een of meer hints op te roepen bij het beantwoorden van een vraag.

#### *Registratie van resultaten en toetsanalyse*

Bij *stand-alone* programma's en bij sommige internet-toetsomgevingen vindt er geen centrale registratie van resultaten plaats. Andere programma's registreren de resultaten automatisch (zoals de meeste toetsmodules in digitale leeromgevingen) of sturen de resultaten per e-mail naar de docent. Registratie is voorwaarde om een



toetsanalyse te kunnen uitvoeren en om cijfers te kunnen toekennen. Sommige programma's gaan zo ver dat per student de resultaten over verschillende toetsen en de eventuele progressie kunnen worden bijgehouden. Bij de meeste toetsprogramma's blijven de toetsanalyses beperkt tot het berekenen van gemiddelden en standaardafwijkingen van vragen; voor het bepalen van de betrouwbaarheid of de kwaliteit van individuele vragen wordt bij meer geavanceerde toetsprogramma's vaak aparte analysessoftware meegeleverd.

#### *Losse programma's of één geïntegreerd programma*

Sommige toetsprogramma's hebben afzonderlijke onderdelen (deelprogramma's) voor de vraagconstructie, *itembanking*, het samenstellen van toetsen, de toetsafname, de resultaatregistratie en de toetsanalyse. Programma's die geschikt zijn voor tentamendoelinden zijn meestal zo opgebouwd. Andere programma's (met name de op het internet gebaseerde programma's) maken meestal géén gebruik van losse programma's. Alles wordt daarbij via één interface (de webbrowser) afgehandeld. Dit komt over het algemeen het gebruiksgemak ten goede. Deze programma's zijn meestal geschikt voor formatieve toetsing. Nieuwe ontwikkelingen, zoals de eerder beschreven PAK-toets, laten zien dat ook via het web nog veel te verwachten valt qua toetsing.

#### *Schriftelijke toetsafname*

Een aantal toetsprogramma's biedt de mogelijkheid om de toets op papier af te drukken. Zo kan een toets schriftelijk worden afgenomen; daarmee kunnen de organisatorische en beveiligingsproblemen worden vermeden die een rol kunnen spelen bij tentamineren op de computer (zie paragraaf 5.3.1). Als hierbij gebruik wordt gemaakt van gestandaardiseerde antwoordformulieren voor een optische lezer, kunnen de resultaten toch automatisch worden verwerkt. Daarmee vervalt echter de mogelijkheid om studenten directe feedback te geven, eenvoudig verschillende versies van dezelfde toets aan te bieden en verschillende typen vragen door elkaar te gebruiken.

### 5.4.3 DE KEUZE VOOR EEN TOETSPROGRAMMA

Welke criteria moeten worden gebruikt om een goede keuze te maken voor een toetsprogramma? Wie een keuze wil maken voor een bepaald toetsprogramma, kan beginnen met het opstellen van een tabel waarin wordt aangevinkt welke functies in de verschillende programma's aanwezig zijn. Hoewel dit voor oriëntatie op de software nuttig is, is het een weinig praktische benadering om een keuze te bepalen. Hoogstwaarschijnlijk zullen de verschillen tussen de geschikte pakketten zo klein of juist zo divers blijken te zijn, dat een afweging niet valt te maken aan de hand van de tabel. De doorslag wordt dan vaak gegeven door overwegingen zoals verspreidingsgraad van de software (de 'installed base'), licentiekosten, kennis bij

systeembeheer, door collega's en collega-instellingen reeds opgebouwde ervaring of vertrouwen in de leverancier. Toch is wel een aantal globale aandachtspunten en aanbevelingen te geven.

*Aandachtspunt 1.* In de eerste plaats is het van belang of het programma vooral bestemd is voor afsluitend tentamineren of voor diagnostisch toetsen en oefentoetsen. Tentamineren stelt hoge eisen aan veiligheid, resultaatregistratie en -analyse en mogelijkheden voor gedifferentieerde scoring. Bij gebrek aan voldoende beveiligbare infrastructuur is de mogelijkheid om toetsen op papier af te drukken van belang. Het programma moet de belangrijkste gewenste vraagtypen goed ondersteunen. Bij diagnostische toetsing en oefentoetsen zullen de eisen hoog zijn voor de mogelijkheid om feedback te kunnen verzorgen en eventueel voor het kunnen kiezen uit een groter aantal vraagtypen.

*Aandachtspunt 2.* Bij steeds meer handboeken en lesmethoden worden kant-en-klaar toetsen geleverd, ofwel op cd-rom, ofwel via een website. Veel methoden leveren de vragen in een kant-en-klaar itembankprogramma dat het mogelijk maakt om de vragen bijvoorbeeld rechtstreeks via Blackboard aan studenten aan te bieden. Vanzelfsprekend is het zinvol om de inzet van deze software serieus te overwegen alvorens te besluiten zelf interactieve toetsen te ontwikkelen.

*Aandachtspunt 3.* Indien er in de opleiding is gekozen voor een standaardtoetsprogramma (of een digitale leeromgeving met ingebouwde mogelijkheden voor toetsen), is het belangrijk dat de mogelijkheden van dat programma tegemoetkomen aan de wensen van de gebruiker. Het voordeel van centrale ondersteuning en het gebruik van een standaard is groot. Centrale bekostiging en het delen van kennis en ervaring met collega's weegt heel vaak op tegen het kopen en implementeren van een nieuw programma.

Mocht het standaardprogramma op essentiële punten toch tekortschieten, dan is oriëntatie op programma's die door collega's in andere opleidingen gehanteerd worden een optie.

#### 5.4.4 NAKIJKEN VAN ANTWOORDEN IN TEKSTVORM

De meeste toetsprogramma's zijn niet in staat om automatisch vragen van het 'essay'-type (open vragen) na te kijken, omdat de formulering van het correcte antwoord niet eenduidig is en er vaak verschillende correcte antwoorden mogelijk zijn. Toch kan de computer ook bij het nakijken van dit type vragen een hulpmiddel zijn.

##### *Semi-automatisch nakijken van tekstdocumenten*

De toetsprogramma's die hiervoor zijn besproken, kunnen open vragen niet automatisch nakijken. Een interessante toepassing die dit deels kan, is het programma

CODAS (Computerondersteund documentanalysestelsel), een systeem dat in staat is om semi-automatisch open vragen na te kijken. Het programma, dat is ontwikkeld aan de Erasmus Universiteit Rotterdam, vergelijkt de antwoorden van studenten onderling (De Mulder & Van Noortwijk, 2000). Voor een goed verloop van het beoordelingsproces kijkt de docent eerst een aantal studentantwoorden met de hand na en merkt deze aan als goed, voldoende, onvoldoende of slecht. Het programma rangordent daarna de antwoorden van de overige studenten naar de mate waarin deze op de correcte of slechte antwoorden lijken. Vervolgens kijkt de docent weer een aantal antwoorden na die bovenaan en onderaan op de lijst zijn gezet, past zo nodig de beoordeling aan en laat het programma opnieuw de antwoorden rangordenen. Het principe is hier dat alle antwoorden die als ‘correct’ of ‘onvoldoende’ worden beoordeeld door de docent worden nagekeken. Het systeem bespaart de docent tijd, omdat het gros van de antwoorden zich meestal in de middenmoot bevindt (deze scoren ‘voldoende’) en die worden niet allemaal door de docent nagekeken.

De antwoorden moeten worden ingeleverd in de vorm van een tekstverwerkingsbestand. Ze moeten voldoende lang zijn (per student minimaal een A4'tje tekst) en, in het geval er meer vragen zijn, over één onderwerp gaan. Het programma blijkt de slechtste opdrachten goed te onderscheiden (deze eindigen gewoonlijk onder in de rangorde), maar de beste uitwerkingen staan vaak niet helemaal bovenaan (Combrink-Kuiters, Elffers, De Mulder & Van Noortwijk, 2000). Het programma is geschikt voor het nakijken van opdrachten waarin studenten beschrijvende antwoorden moeten geven (bijv. een theorie of een principe in eigen woorden weergeven). De methode is minder geschikt voor opdrachten waarin het gaat om het ontwikkelen van taalvaardigheid of waarin eigen mening of creativiteit een grote rol speelt. Het programma is op enkele plaatsen in Nederland naar tevredenheid van docenten en studenten toegepast (Draaijer, Van Gastel, Peeters, Frinking & Reumer, 2004).

#### *Handmatig commentaar invoegen in tekstdocumenten*

De behoefte om tekstdocumenten eenvoudig van commentaar te voorzien neemt toe doordat studenten vragen en opdrachten steeds vaker inleveren via een digitale leeromgeving of via e-mail. Er bestaat een aantal hulpmiddelen om tekstverwerkingsdocumenten eenvoudig digitaal te becommentariëren:

- De meest directe methode is de ‘redigeerfunctie’ van Microsoft Word, WordPerfect of Adobe Acrobat. Deze functie registreert opmerkingen, wijzigingen, commentaar, verwijdering en dergelijke in het na te kijken document. Aangebrachte wijzigingen zijn herkenbaar door afwijkende kleur of lettertype. Deze mogelijkheid van becommentariëren is uit te breiden door in deze programma's de tekenmodule te activeren. Zo kunnen eenvoudig pijlen, cirkels en opmerkingen worden ingevoegd.
- Vrij nieuw zijn internetprogramma's die het mogelijk maken online commentaar bij teksten toe te voegen. Via een knoppenmenu is het eenvoudig veelge-

bruikte beoordelingstermen in de tekst te plaatsen. Het programma houdt het aantal opmerkingen bij. Aan het eind van de tekst genereert het meer algemene feedback of een beoordeling in punten. Deze functie is bijvoorbeeld geïntegreerd in het plagiaatdetectieprogramma Turnitin en SafeAssignment/MyDrop Box (zie ook hoofdstuk 21).

- Verschillende opleidingen laten ook studenten onderling commentaar toevoegen aan tekstdocumenten. Deze vorm van *peer assessment* is zeer geschikt bij formatieve toetsing: studenten leren niet alleen van het krijgen, maar ook van het geven en formuleren van feedback. Dit kan zowel met de redigeerfunctie van de tekstverwerker als met speciale internetprogramma's. Wel is het belangrijk dat studenten daarbij aanvullende sturing krijgen, bijvoorbeeld via een beoordelingschecklist of feedbackcategorieën. In de digitale schrijfomgevingen Concourse of Turnitin is dergelijke ondersteuning geïntegreerd.

### Samenvatting

- De computer biedt interessante toetsmogelijkheden, zeker voor *diagnostische toetsen* en *oefentoetsen*. Het belangrijkste voordeel is de directe feedback aan studenten en de mogelijkheid om multimediale informatie toe te voegen (foto's, afbeeldingen, diagrammen, video's en dergelijke).
  - Interactieve diagnostische toetsing biedt de docent de mogelijkheid om de studievoortgang van studenten te bewaken en zo nodig bij te sturen.
  - Interactieve oefentoetsen bieden studenten de mogelijkheid om op eigen gelegenheid te oefenen.
  - Voor diagnostische toetsen en oefentoetsen is van belang dat het toetsprogramma goede feedbackmogelijkheden heeft en laagdrempelig is in het gebruik.
- Interactief afsluitend *tentamineren* met de computer is zeker mogelijk, maar de randvoorwaarden op het terrein van beveiliging beperken tot nu toe de inzet op grote schaal.
  - De software is meestal complex, minder gebruiksvriendelijk en kostbaar, zodat samenwerking op het niveau van de opleiding, faculteit of instelling is vereist.
  - Inzet op grote schaal is alleen zinvol als sprake is van gestandaardiseerde toetsvormen en procedures. Dit levert een spanningsveld op, omdat er binnen een instelling juist veel verschillende systemen, onderwijsvormen en vakgebieden bestaan die elk hun eigen eisen stellen.
- Er is veel software waarmee docenten zelf toetsen kunnen maken. Daarnaast worden bij steeds meer studieboeken kant-en-klare interactieve toetsen ontwikkeld.
- Houd er rekening mee dat het zelf maken van interactieve toetsvragen een behoorlijke tijdsinvestering vergt. Dit is de moeite waard als er sprake is van een

stabiel vakgebied, een groot aantal studenten en een duidelijke noodzaak voor kwaliteitsverbetering van het onderwijs of het toetsen.

- Itembanken bieden de mogelijkheid snel toetsen samen te stellen, die op de computer of op papier worden afgenomen. Het opzetten en onderhouden van een itembank vergt regelmatig tijd en aandacht.
- Het kiezen van een toetsprogramma is niet eenvoudig, want toetsprogramma's bestaan in vele verschillende soorten en maten. Ga bij een keuze uit van een beperkt aantal echt belangrijke eisen en vermijd de valkuil te zoeken naar het schaap met de vijf poten.
- De computer biedt een aantal hulpmiddelen om open antwoorden in tekstvorm semi-automatisch na te kijken.

### Geannoteerde literatuur

Combrink-Kuiters, C.J.M., Elffers, H., Mulder, R.V. de, & Noortwijk, C. van (2000). Computer-ondersteund nakijken van open vragen: een onderzoek naar de betrouwbaarheid en gebruikersvriendelijkheid van de CODAS software. Meppel: Edu' Actief.

In dit boek rapporteren en evalueren de ontwikkelaars van CODAS hun (eigen) ervaringen met de nakijkmodule en fraudecheckmodule van de software.

Draaijer, S., Gastel, L. van, Peeters, V., Frinking, P., & Reumer, C. van e.a. (2004). Handboek flexibilisering van toetsing. Utrecht: Digitale Universiteit.

Dit handboek biedt een uitgebreide en zeer leesbare inleiding in (digitaal) toetsen, en de functies en mogelijkheden van digitaal toetsen. Op de DU-website Flexibilisering van toetsing worden alle resultaten van het project 'Flexibilisering van Toetsing' gepresenteerd. De lezer kan casussen inzien en er staan links naar andere nuttige sites. Dit handboek kan als de rode draad van de website worden beschouwd: in het boek wordt regelmatig naar andere delen van de website verwezen. <http://www.du.nl/flextoets>

Parshall, C.G., Spray, J.A., Kalohn, J.C., & Davey, T. (2002). Practical considerations in computer-based testing. New York: Springer-Verlag.

Dit boek gaat uitgebreid in op de praktische kanten van computersystemen voor toetsen. Het boek behandelt zowel de technische kanten (zoals beschikbare software), verschillende stappen in het toetsproces (zoals het maken van vragen, het samenstellen van toetsen, psychometrische aspecten, adaptief toetsen, onderhoud van vragenbestanden) als kwesties zoals de reacties van studenten op afname van toetsen per computer. Het boek is interessant voor toetsspecialisten.

Via de website van het ICT Onderwijscentrum VU ([www.icto.vu.nl](http://www.icto.vu.nl)) zijn alle internetadressen aan te klikken die in dit hoofdstuk genoemd worden, en zijn actuele ontwikkelingen, achtergrondinformatie en voorbeeldprogramma's te vinden.

