

# De financiële dienstverlener aan banden

## Deel 2: De eerste bevindingen

□ Henk Kreeft

In het vorige nummer van *EXAMENS* schreef Henk van Berkel over de Wet op het financieel toezicht (Wft). Financiële dienstverleners moeten aan bepaalde eisen voldoen en daartoe examens afleggen. Er is een College Deskundigheid Financiële Dienstverlening (CDFD) opgericht dat onder andere als taak heeft exameninstellingen aan te wijzen die gerechtigd zijn tot het afnemen van de examens. Het erkennen van exameninstellingen is één ding, het handhaven van erkende kwaliteit is wat anders. Daartoe ontwikkelde het CDFD een auditsysteem. In het afgelopen jaar zijn vijf exameninstellingen ge-audit. Hier beschrijft een lid van de visitatiecommissie zijn ervaringen tijdens dit auditproces.

### Inleiding

Bij de start van de eerste visitatieronde gaf het College Deskundigheid Financiële Dienstverlening (CDFD) twaalf exameninstellingen een voorlopige erkenning om Wft-examens af te nemen. Exameninstellingen verkrijgen een voorlopige erkenning van twee jaar op basis van door de exameninstellingen aangeleverde documenten. Over het algemeen zijn het instellingen met een staat van dienst op het gebied van examinering in de financiële wereld. Zes van de exameninstellingen staan open voor een ieder die zich heeft voorbereid op de Wft-examens, de overige

zes zijn hogescholen die ervoor willen zorgen dat hun studenten bij afronding van de studie ook over de Wft-diploma's beschikken.

Het CDFD heeft vervolgens de organisatie *Quality Assurance Netherlands Universities* (Qanu) aangewezen om het visiteren van exameninstellingen en een kwaliteitsbeoordeling te organiseren. Op basis daarvan wordt over het al of niet verlengen van de erkenning geadviseerd. De Qanu heeft, onder verantwoordelijkheid van het CDFD, een visitatiecommissie van vijf personen ingesteld. De voorzitter en drie leden zijn deskundigen op diverse werkgebieden van de financiële dienstverlening. Examendeskundigheid wordt geleverd door de auteur van dit artikel. In de hier besproken periode heeft deze commissie vijf exameninstellingen gevisiteerd.

De situatie die de visitatiecommissie aantrof bij de vijf exameninstellingen was heel verschillend. Drie instellingen zijn vanaf begin 2007 druk bezig geweest de bestaande examens om te bouwen naar examens voor de Wft-modulen. Een tweetal instellingen ging ervan uit dat dit niet nodig was en dat de bestaande examens voldoende garantie boden om Wft-diploma's te kunnen verstrekken.

Zoals in het vorige nummer van *EXAMENS* is beschreven, zijn in het kader van dit visitatieproces zes criteria gedefinieerd die in samenhang een goed beeld moeten vormen van de kwaliteit van examens en exameninstellingen. Dit artikel zal vooral ingaan op de techniek van examineren.

### Examenvorm

Wat betreft de examentechniek zijn vrijwel alle exameninstellingen doorgegaan op de weg die zij reeds volgden. Drie van de vijf instellingen werken en werkten met digitale afname waarbij de examens *at random* worden samengesteld uit een itembank. Deze instellingen beschikken niet zelf over afnamefaciliteiten zoals testcentra en een server voor het aanbieden van de examens. Ze schakelen hiervoor gespecialiseerde examenproviders in. Analyses van

## Vier exameninstellingen hebben een positief advies verkregen



### Het auditproces

*Het auditproces omvat de volgende stappen:*

1. Het opstellen van een notitie door de exameninstellingen waaruit zou moeten blijken dat wordt voldaan aan de eisen.
2. Het instellen van een visitatiecommissie door het CDFD. De experts in de commissie beschikken over vakdeskundigheid en deskundigheid op het gebied van de examens. Zij bezoeken alle instellingen.
3. Een bezoek van de commissie aan het exameninstituut en het vormen een oordeel over het wel of niet voldoen aan de gestelde criteria.
4. Een rapportage van de commissie aan het college met een advies over het al dan niet verlengen van de erkenning en aanbevelingen voor verbeteringen.
5. Besluitvorming door het college over (voorwaardelijke) verlenging of intrekking van de erkenning.

deze examenproviders uitgevoerd. Twee exameninstellingen nemen de examens op papier af. Zij laten de analyses van examens eveneens extern uitvoeren.

### Dekking van de toetstermen

In de Wft staan eindtermen. Een van de taken van het College is deze eindtermen om te zetten in concrete(re) toetstermen. Bij de toetstermen is door het CDFD een volledige toetsmatrijs meegeleverd. Exameninstellingen zijn echter niet verplicht deze toetsmatrijs te hanteren. De toetsmatrijs kent inhoudscategorieën en gedragscategorieën. De gedragscategorieën zijn: 'kennis', 'begrip', 'toepassen procedure' en 'toepassen inzicht'. In de toetstermen is de combinatie van inhoud en gedrag al vastgelegd. Elke toetsterm komt slechts voor met één type gedrag. Drie exameninstellingen bieden examens per module aan. Zij hebben toetsmatrijzen van het CDFD één op één overgenomen en vertaald naar hun algoritmen voor de toetssamenstelling. De twee andere instellingen die een geïntegreerd examen aanbieden, hanteren ook adequate toetsmatrijzen, maar deze blijken moeilijk of niet terug te voeren op de Wft-toetstermen.

### Itembank

Vier van de vijf exameninstellingen werken met een itembank waaruit zij examens samenstellen. Bij de instellingen die werken met flexibele afnamen worden examens *at random* per kandidaat samengesteld uit de itembank. De itembank moet een bepaalde omvang hebben om de geheimhouding van de vragen te waarborgen. Ook moeten vragen over alle

toetstermen in de itembank aanwezig zijn om via *at random* trekking een representatief (ten opzichte van de toetsmatrijs) examen te realiseren. Er is discussie mogelijk over de omvang van de itembank. De meeste exameninstellingen gaan uit van vijf keer het aantal vragen van de toets. De visitatiecommissie heeft als norm aangehouden minimaal drie vragen per toetsterm. Gezien het grote aantal toetstermen vergt dit bij de meeste modules een grotere omvang van de itembank dan het criterium van vijf keer het aantal vragen. De itembanken zoals die nu beschikbaar zijn, voldoen nog niet aan het criterium van de visitatiecommissie hoewel er de afgelopen tijd bijzonder hard gewerkt is. Zij voldoen echter ook niet aan het criterium dat exameninstellingen zelf hebben geformuleerd. De vulling van de itembank is een van de belangrijkste aandachtspunten voor de verdere examenontwikkeling.

### Toetslengte

Een examen kan worden opgevat als een steekproef uit het aantal mogelijke vragen. Met behulp van de toetsmatrijs verkrijgt het examen een bepaalde samenstelling. Met een langer examen kan doorgaans een betere steekproef worden gerealiseerd. Bezien moet worden of elke categorie van de toetsmatrijs in een examen is opgenomen en of de lengte van het examen in overeenstemming is met het aantal te examineren categorieën.

Het aantal vragen van examens die telkens één module beslaan, blijkt proportioneel te zijn aan het aantal categorieën in de toetsmatrijs. Er zijn ook examens die méér dan één module beslaan. Dan volstaan 30 of



## *De vulling van de itembank is een van de belangrijkste aandachtspunten*

40 vragen niet, zo bleek de commissie. In die gevallen vond de commissie teveel niet-getoetste eindtermen. Bovendien bestaat er een relatie tussen de lengte van examens en de toetsbetrouwbaarheid. Veertig meerkeuzevragen om één module te toetsen lijkt daarom wel het minimum.

### **Variatie in vraagvormen**

De variatie in vraagvormen is bij de meeste examens niet groot. Een examen bestaat vaak uit alleen driekeuze-items. Dit hangt niet samen met de digitale afname. Eén van de exameninstellingen die de afname heeft gedigitaliseerd, hanteert juist vele verschillende vraagvormen. Een zeer open examenvorm wordt ook gehanteerd. Het gaat dan om een casuïstiekexamen waarvan de kandidaten aan de hand van een aantal richtvragen zelf bepalen hoe zij hun kennis etaleren. Voor de visitatiecommissie is de variatie in vraagvorm op zichzelf geen aandachtspunt, maar wel de vraag of met de gehanteerde vraagvormen de toetstermen kunnen worden gerealiseerd, dus ook de toetstermen in de categorie toepassing. Over het algemeen blijkt dit het geval.

### **Moeilijkheidsgraad en gelijkwaardigheid**

Vragen verschillen in moeilijkheidsgraad. Examens kunnen daarom ook in moeilijkheidsgraad verschillen. Voor exameninstellingen is het een uitdaging om te garanderen dat examens op verschillende momenten afgenomen, gelijkwaardig zijn. De visitatiecommissie moet constateren dat de meeste exameninstellingen ervan uitgaan dat de *at random*-samenstelling voldoende garantie is voor de gelijkwaardigheid van de examens. Een controle op de feitelijke juistheid van deze aanname wordt nergens uitgevoerd. Wel worden vragen verwijderd die na ongeveer honderd keer te zijn voorgelegd te moeilijk (minder dan 30 procent correct) of te gemakkelijk (meer dan 90 procent correct) blijken. Bij examens die eenmalig worden gebruikt, moet het oordeel van de vakinhoudelijke deskundigen gelijkwaardigheid garanderen. Een methode die een van de instituten daarbij gebruikt, is het vooraf indelen van vragen in drie moeilijkheidsgroepen.

Tijdens de visitatie bleek dat geen van de instituten

de gelijkwaardigheid van examens op verschillende momenten werkelijk kan garanderen. Geen van de exameninstellingen is vertrouwd met de mogelijkheden om itembanken te kalibreren<sup>1</sup> om op deze wijze de gelijkwaardigheid van examens en de kwaliteit van vragen in de bank te kunnen garanderen. De itembanken zoals deze voor de Wft-examens worden ingericht, lenen zich evenwel bij uitstek voor deze techniek.

Er kan ook een verschil in moeilijkheidsgraad zijn tussen de examens van de verschillende exameninstellingen over dezelfde module. De inhoudsdeskundigen van de visitatiecommissie hebben dit soort verschillen inderdaad waargenomen. Zij hebben zich echter vooral gericht op een ondergrens voor zover zij die op basis van toetstermen en hun ervaring met het beroepenveld konden aangeven. De globale indruk van de commissie was dat de examens daaraan voldeden.

### **Beoordeling**

Alle meerkeuze-examens worden met de computer nagekeken. Voor het zeer open examen casuïstiek van een van de exameninstellingen, is een beoordelings-

### **Een intern kwaliteitszorgsysteem**

De visitatiecommissie neemt kennis van de beschikbare documentatie van de exameninstellingen en stelt daarover vragen, mede aan de hand van een beoordelingskader, aan de deskundigen van het instituut. Dit geeft een behoorlijk compleet beeld van hoe er wordt gewerkt en welke kennis wel en niet in huis is. In de dagelijkse praktijk die de commissie niet heeft gezien, kan het een en ander natuurlijk wel eens minder fraai lopen dan beoogd. Maar dat is iets waar elke organisatie mee te maken heeft. Voor een visitatiecommissie is het van belang of de organisatie kennelijk middelen heeft om eventuele feilen te constateren. Met andere woorden, een goed functionerend intern kwaliteitszorgsysteem is belangrijk.

## *De exameninstellingen kunnen de gelijkwaardigheid van examens niet garanderen*



procedure ontwikkeld die een redelijke objectiviteit moet kunnen garanderen. De instelling beschikt echter niet over gegevens of dit ook daadwerkelijk zo is.

### **Normering**

De normering van vrijwel alle examens is absoluut. Dat wil zeggen, de zak/slaaggrens is van tevoren vastgesteld (op ongeveer 70 procent van het aantal vragen) zonder dat bijstelling achteraf in de procedure opgenomen is. De 70-procent-norm heeft onder andere te maken met de raadkans bij de meerkeuzevragen. De onduidelijkheid over de gelijkwaardigheid van examens maakt echter absoluut normeren discutabel. Voor wie een moeilijker examen heeft gehad, zou immers de norm lager moeten worden gesteld en voor wie een makkelijker examen kreeg, moet de norm wat hoger liggen. Overigens is het bijstellen van de norm achteraf niet de meest ideale oplossing voor het normeren van *at random* samengestelde digitale examens. Het vooraf kalibreren van de itembank en op basis daarvan de examens samenstellen is de beste oplossing.

Alleen bij het genoemde examen casuïstiek is sprake van een relatieve norm die op basis van de beoordeling van het examenwerk door de beoordelaars wordt vastgesteld.

### **Evaluatie van examens**

Examens behoren achteraf te worden geanalyseerd en geëvalueerd. Alle instellingen beschikken waar het meerkeuzevragen betrof over toetstechnische analyses. De productie van de analyses is uitbesteed. Dit hoeft geen bezwaar te zijn als men de gegevens maar weet te interpreteren. Dit was lang niet altijd het geval. De percentages goed beantwoorde vragen spelen bij elke instelling een rol. Maar andere analysegegevens worden vaak niet gebruikt omdat de kennis om de gegevens te interpreteren ontbrak.

De exameninstellingen met examens op papier hebben als enige gegevens over de toetsbetrouwbaarheid. Exameninstituten die met digitale afnamen werken, beschikken over minder gegevens. De conclusie is gerechtvaardigd dat de digitale afnamen van examens een grote winst betekent voor de flexibiliteit van de examinering. Er is de mogelijkheid om dagelijks examen te doen in plaats van een



paar keer per jaar. Maar dit is ten koste gegaan van de toetstechnische analyses die de kwaliteit van de itembank moeten bewaken. Van geen enkel digitaal examen zijn gegevens over de toetsbetrouwbaarheid beschikbaar, en bij de berekeningswijze van een discriminatie-index<sup>2</sup> per vraag kan men grote vraagtekens plaatsen. De klassieke toets- en itemanalyses zijn geschikt voor traditionele pen-en-papier-afnamen. Bij examentechnieken zoals digitale afnamen, behoren andere analysetechnieken. Kalibreren van de itembanken en samenstellen van examens op basis van de kalibratiegegevens zouden voor het aanbieden van de Wft-examens een aanzienlijke kwaliteitsverbetering opleveren.

### **Bezwaren**

Elke exameninstelling heeft een bezwarenprocedure. Kandidaten kunnen hun examen inzien en bezwaar





## *Samenstellen van examens op basis van de kalibratiegegevens zal kwaliteitsverbetering opleveren*

aantekenen tegen een bepaalde vraag. Het bezwaar wordt door de examencommissie beoordeeld op inhoudelijke gronden. Als het bezwaar wordt toegekend, vindt een correctie van de uitslag plaats voor die kandidaat. Bij de digitale examens is een correctie achteraf voor alle kandidaten die de betreffende vraag ook voorgelegd gekregen hebben, niet uitvoerbaar. Deze ongelijkheid in behandeling van degene die bezwaar aantekenen en degenen die dat niet doen, pleit ook weer voor een goede kwaliteitscontrole vooraf, waaraan kalibratie van itembanken veel kan bijdragen.

### **Toetsbeleid**

Bij slechts een enkele instelling trof het visitatieteam een uitgeschreven toetsbeleid aan. Natuurlijk beschikken alle instituten over instructies voor examencommissies en zijn belangrijke procedures vastgelegd. Bijeenvoegen van alle keuzes die op het gebied van examinering zijn gemaakt, gecombineerd met prestatie-indicatoren in een samenhangend toetsbeleid, levert een instrument op dat voor kwaliteitsbeheersing een belangrijke rol kan vervullen. De commissie heeft zoiets zelden aangetroffen.

### **Andere aspecten**

In het voorgaande zijn vooral de examenteknische aspecten van de Wft-examens aan de orde gekomen, maar er zijn ook andere aspecten in de visitatie beoordeeld. De exameninstellingen zijn sterk in het organiseren van examens. Informatie aan kandidaten, de toegankelijkheid voor kandidaten, examenreglementen, frauderegelingen en bezwaarprocedures waren alle in orde. De gescheiden bedrijfsvoering tussen opleidingen en exameninstituten heeft veel aandacht gekregen en tot enkele opmerkingen geleid.

### **Adviezen**

De beoordeling van de visitatiecommissie is omgezet in adviezen aan het CDFD over het al of niet verlengen van de voorlopige erkenning. De herkenbaarheid van de Wft-eind- en toetstermen in de examens en de mate waarin de itembank is gevuld, speelden daarbij de belangrijkste rol. Deze punten zijn van belang omdat die inzicht geven in de mate waarin de examens van verschillende exameninstel-

lingen gelijkwaardig zijn. De andere punten waarover de visitatiecommissie met de deskundigen van de exameninstellingen heeft gesproken, zijn bedoeld als punten die kunnen leiden tot kwaliteitsverbetering. Uiteindelijk hebben vier exameninstellingen een positief advies verkregen. Het CDFD heeft die adviezen onverkort overgenomen.

### **Conclusie**

De inhoudelijke beoordeling van de examenvragen en de analyses van de relatie tussen de toetsmatrijs en het betreffende examen, vergt waarschijnlijk meer tijd dan het gedeelte van de dag dat er aan kon worden besteed. Doordat de visitatiecommissie voor het grootste deel uit inhoudsdeskundigen bestond, was het wel mogelijk een onderbouwde uitspraak over de inhoud te doen. Belangrijk was dat een gesprek tussen de deskundigen van de exameninstellingen en de leden van de visitatiecommissie leidde tot begrip en acceptatie van de visie op de gehanteerde werkwijzen. De ervaring was dat de gesprekken in dit opzicht goed zijn verlopen.

De winst van een visitatieprocedure, waarin alles in detail wordt bekeken en besproken, is dat argumenten over en weer worden toegelicht en gewogen en dat er meer begrip ontstaat voor de beslissingen die het CDFD uiteindelijk moet nemen. De gesprekken zullen naar verwachting leiden tot verbeteringen in de kwaliteit van de Wft-examens.

### **Literatuur**

Berkel, H.J.M. van (2008). De financiële dienstverlener aan banden. Deel 1: Herkomst, criteria en proces. *EXAMENS*, 4(2), 14-17.

### **Noten**

- 1 Kalibreren is het schatten (berekenen) van de psychometrische eigenschappen (parameters) van een verzameling items met het doel om deze items op dezelfde schaal te brengen.
- 2 De discriminatie-index geeft aan in hoeverre een item correct is gemaakt door kandidaten die op de gehele toets hoog scoren.

☐ Drs. H.P. Kreeft is zelfstandig gevestigd adviseur. In het CDFD is hij de examendeskundige. Hij schreef deze bijdrage op persoonlijke titel. E-mail: [hpkreeft@xs4all.nl](mailto:hpkreeft@xs4all.nl).