HANDBOEK
IN 5 STAPPEN NAAR EEN ITEMBANK

WEGWIJS IN TOETSVRAGENBANKEN

SURF

HANDBOEK
IN 5 STAPPEN NAAR EEN ITEMBANK

WEGWIJS IN TOETSVRAGENBANKEN

INHOUDSOPGAVE

DEEL 1: INTRODUCTIE 6

1. Inleiding 7
1.1 Voor wie is deze uitgave? 7
1.2 Scope 7
1.3 Verantwoording 8

2. Definities en begrippen 9
2.1 Wat is een itembank? 9
2.2 De plek van een itembank binnen het digitale toetssysteem 9
2.3 Belangrijke begrippen 11

3. In 5 stappen naar een itembank 12

4. Succesfactoren 14

DEEL 2: STAPPENPLAN 15

5. Stap 1: Voorbereiding 16
5.1 Haalbaarheidsonderzoek 16

6. Stap 2: Plan van aanpak 18

7. Stap 3: Ontwerp 19
7.1 Ordening van items 19
7.2 Itemontwikkelproces 19
7.3 Itemkwaliteit 20
7.4 Systeem 20
7.5 Productie en planning 20
7.6 Financiën 20
7.7 Juridische aandachtspunten 21
7.8 Beheer itembank 21

8. Stap 4: Pilot 22

9. Stap 5: Realisatie 23
9.1 Nazorg 23

DEEL 3: VERDIEPING 24

10. Waarom (samenwerken aan) een itembank? 25
10.1 Waarom een itembank? 25
10.2 Waarom samenwerken aan een itembank? 25
10.3 Kansrijke vakgebieden voor een succesvolle samenwerking 26

11. Technologie: Systemen 27
11.1 Veel gebruikte toetssystemen voor onderwijs 27
11.2 Toetsitems uitwisselen tussen systemen 28
11.2.1 QTI-standaard 28
11.2.2 LTI en SCORM standaard 29
11.3 Veiligheid en schaalbaarheid van itembanken 29

12. Omvang van de itembank 30
12.1 Summatieve toetsen met een gelijktijdig toetsmoment voor alle studenten 30
12.2 Summatieve toetsen die op elk moment kunnen worden afgenomen 30
12.3 Formatieve diagnostische toetsen 31
12.4 Formatieve oefentoetsen 31

13. Kosten en baten 32
13.1 Kosten 32
13.2 Baten 33

14. Organisatie: processtappen, rollen en rechten 34
14.1 Componenten en processtappen 34
14.2 Rollen en rechten 35
14.3 Geautomatiseerde workflow 36

15. De ordening van een itembank **37**
15.1 Metadata 37
15.2 Structuur 38
15.3 Structuur en metadata gebruiken 40
15.4 Feedback en vraagtypes 42
15.4.1 Feedback 42
15.4.2 Vraagtypes 43

16. Didactiek: kwaliteit van items 44
16.1 Kwaliteit van items 44
16.2 Hoe realiseer je een kwaliteitsverbetering? 45
16.2.1 Voorlichting en training 45
16.2.2 Gecontroleerde itemontwikkeling door review 45
16.2.3 Itembankonderhoud en beheer 46
16.3 Toetsgegevens 47
16.3.1 Beschrijvende data 47
16.3.2 Psychometrische data 47

17. Juridische aandachtspunten 49
17.1 Eigendom items en itembank 49
17.2 Misbruik van itembankmateriaal 50
17.3 Het gebruik van materiaal van andere partijen in itembanken 50
17.3.1 Linken mag altijd 50
17.3.2 Licentie voor hergebruik 50
17.3.3 Het citaatrecht biedt veel mogelijkheden 51
17.3.4 Afbeelding of audiovisueel werk uit de eigen bibliotheek 51
17.3.5 Complete audio(visuele) werken 51
17.3.6 Delen van audio(visuele) werken 51
17.3.7 Bronvermelding 51
17.4 Bescherming persoonsgegevens 52
17.5 Rechtsvormen voor samenwerkingsverbanden 52

Bijlage 1: Begrippenkader 53
Bijlage 2: Drie workflows voor itemontwikkeling 59
1. Een itemontwikkelproces door twee docenten 59
2. Itemontwikkelproces itembank Toets & Leer 60
3. Itemontwikkelproces MBO kennistoetsenbank 61

Colofon 63

DEEL 1 INTRODUCTIE

1. INLEIDING

Digitale itembanken zijn een prachtige tool om de kwaliteit van toetsen te laten toenemen en de kosten te laten dalen. Wie gezamenlijk toetsvragen ontwikkelt, stemt doelen en processen af. Er is meer expertise en tijd beschikbaar per item (toetsvraag) voor ontwikkeling en kwaliteitscontrole. Dit alles komt de kwaliteit van de toetsen ten goede. Hergebruik van items kan de kosten doen dalen. Onder andere om deze redenen staat het delen van leermateriaal en dus ook toetsmateriaal hoog op verschillende agenda's van hogeronderwijsinstellingen.

Het ontwikkelen, beheren en onderhouden van itembanken vraagt om goed projectmanagement. Met deze uitgave willen we laten zien hoe je een digitaal systeem opzet om items uit te wisselen met collega's van binnen en buiten de instelling. In dit handboek bundelen we de kennis en kunde die instellingen, samenwerkingsverbanden van instellingen en SURF over dit onderwerp hebben opgedaan. Het helpt instellingen op weg die hun voordeel willen behalen met digitale itembanken. Het handboek bestaat uit een introductie, een stappenplan en een verdiepend deel.

1.1 Voor wie is deze uitgave?

Het handboek is bedoeld voor iedereen die een rol vervult in de processen rondom de ontwikkeling, het beheer en het onderhoud van een itembank en de items.

Bijvoorbeeld:

- Docenten die items beter willen ordenen of willen ontwikkelen en uitwisselen met een collega
- Onderwijsadviseur die docenten ondersteunt of die trekker is van het inrichten van een itembank
- Een vakcommunity die gezamenlijk toetsmateriaal wil uitwisselen of gezamenlijk toetsen wil uitzetten
- Een landelijk overleg dat de kwaliteit van toetsen wil verhogen
- Functioneel beheerders van een toetssysteem dat wordt gebruikt voor een itembank

1.2 Scope

Het handboek richt zich op het ontwikkelen, beheren en gebruiken van itembanken binnen een opleiding, binnen één instelling, instellingsoverstijgend, binnen een vakcommunity of zelfs landelijk. Uitgangspunt is dat de processen in de basis hetzelfde zijn. Het handboek geeft geen uitgebreide beschrijving van bestaande itembanken of toetssystemen. Hiervoor kun je terecht in de SURF thema-uitgave *Toets-en vragenbanken in het onderwijs*.[1]

Het ontwikkelen van itembanken kan onderdeel zijn van een groter implementatietraject van digitaal toetsen bij instellingen. Voor het implementeren van digitaal toetsen verwijzen wij naar enkele SURF publicaties, zoals de handreiking *Beleid voor digitaal toetsen*, het stappenplan *Digitaal toetsbeleid schrijven* en het werkboek *Veilig toetsen*.[2]

[1] https://www.surf.nl/kennisbank/2017/thema-uitgave-toets--en-vragenbanken.html

[2] https://www.surf.nl/kennisbank/2016/handreiking-beleid-voor-digitaal-toetsen.html
https://www.surf.nl/kennisbank/2016/stappenplan-digitaaltoetsbeleid-schrijven.html
https://www.surf.nl/kennisbank/2016/werkboek-veilig-toetsen.html

We gaan in dit handboek uit van de toetscyclus zoals beschreven in het *Begrippenkader voor digitaal toetsen* van SURF uit 2013.[3] Centraal staat het ontwikkelen van items (het itemontwikkelproces) en het beheer van items in een itembank. Een goede itembank vormt de basis voor goede toetsen. Over hoe je op basis van zo'n itembank goede toetsen ontwikkelt in een toetssysteem, zijn andere bronnen beschikbaar, zoals *Toetsen in het hoger onderwijs* van Berkel, Bax en Joosten-ten Brinke.[4]

1.3 Verantwoording

De inhoud van deze publicatie is tot stand gekomen in nauwe samenwerking met itembank-experts uit diverse instellingen, de special interest group (SIG) Digitaal Toetsen, Cito, 10voordeLeraar en Prove2Move. Het handboek is gebaseerd op de ervaringen van de betrokken experts met het ontwikkelen van itembanken. Verder hebben we gebruikgemaakt van verzameld materiaal van onder andere SURF en beschikbare literatuur op dit gebied.

Zie het colofon voor een overzicht van alle betrokkenen die hebben bijgedragen aan dit handboek. Neem gerust contact op met de auteurs en/of de deskundigen die een inhoudelijke bijdrage hebben geleverd aan het handboek voor meer informatie.

[3] *https://www.surf.nl/kennisbank/2013/begrippenkader-voor-digitaal-toetsen.html*

[4] *Van Berkel, H. J. M., Bax, A., & Joosten-ten Brinke, D. (2017). Toetsen in het Hoger Onderwijs. (4rd ed.). Houten/Diegem: Bohn Stafleu Van Loghum.*

2. DEFINITIES EN BEGRIPPEN

Voor het gebruik van dit handboek is een goed begrip van een aantal termen essentieel. We beschrijven kort de componenten van een digitaal toetssysteem en laten zien hoe ze zich verhouden tot de itembank.

2.1 Wat is een itembank?

Een itembank is een verzameling van items ofwel vragen (in het vervolg van dit handboek noemen we ze items). Een itembanksysteem is het digitale systeem waarin de verzameling van items is opgeslagen en waarmee ze beheerd kunnen worden. Een itembanksysteem kan meerdere itembanken bevatten. Als wij in dit werkboek spreken over itembank bedoelen we het itembanksysteem (de digitale itembank).

Definitie
Een **itembank** is een verzameling van items voor een bepaald toetsdoel. Een itembank heeft daarvoor een bepaalde ordening, meestal op basis van metadata en in de vorm van een hiërarchische structuur.
Een **itembanksysteem** is het digitale systeem waarin de verzameling van items is opgeslagen en waarmee ze beheerd kunnen worden. Een itembanksysteem kan meerdere itembanken bevatten.

In de itembank worden de vragen logisch ingedeeld. Er vindt versiebeheer plaats en er kunnen allerlei metadata worden opgeslagen over de inhoud, het vraagtype, het niveau, et cetera. Je kunt er ook psychometrische gegevens in opslaan, zoals moeilijkheidsgraad en het discriminerend vermogen van items. Soms draaien de programma's van een itembanksysteem lokaal op een computer en staat de database ook op een computer of een server. Andere itembanken bestaan uit internetomgevingen. De database en de applicatie worden vanuit de cloud benaderd. Verschillende mensen kunnen tegelijkertijd werken in zo'n cloudsysteem.

Nice to know
Waarom worden 'toetsvragen' door toetsdeskundigen eigenlijk vaak 'items' genoemd?
In de jaren 30 werd in de Verenigde Staten al gewerkt met gestandaardiseerde toetsen en rudimentaire bestanden met vragen. Al snel ontstond een beroepsgroep met bijbehorend beroepsjargon. 'Test questions' werd 'Test items'. In Nederland hebben we dit woordgebruik overgenomen en 'vernederlandst' naar 'ietums' in plaats van 'aitums'.

2.2 De plek van een itembank binnen het digitale toetssysteem

Met een digitaal toetssysteem bedoelen we software die het digitale toetsproces faciliteert. We maken onderscheid tussen een aantal kerncomponenten: auteursomgeving, itembank, afspeelomgeving en analysetool. Deze componenten bespreken we aan de hand van de toetscyclus. Niet alle stappen uit de toetscyclus worden ondersteund door digitale toetssystemen. In figuur 1 is te zien hoe de componenten van een digitaal toetssysteem aansluiten op de toetscyclus.

Figuur 1 Componenten van digitale toetssystemen afgebeeld op de toetscyclus

De *itembank* vormt de centrale component waarin de items worden opgeslagen. De bewerking en het afspelen van de items vindt plaats in andere systemen, die vervolgens de resultaten teruggeven aan de itembank. Bij de volgende onderdelen van de toetscyclus speelt de itembank een rol:

Het **ontwerp** van de toets (ook wel toetsspecificatie) gaat over aspecten als het doel van de toets, te toetsen onderdelen, de keuze voor formatief of summatief toetsen en 'wijze van toetsing'. In het ontwerp worden dus nog geen items gemaakt.

Een toets wordt samengesteld vanuit een verzameling items. Hiervoor wordt meestal een specificatietabel gebruikt, ook wel toetsmatrijs of blauwdruk van de toets genaamd. Hierin is vastgelegd hoeveel items, welke vraagtypen en welke onderwerpen van welke aard, kennis, of toepassing in de toets komen.

De **ontwikkeling** (constructie) van de items vindt plaats in de *auteursomgeving* van een toetssysteem. Ook de antwoorden (in geval van gesloten vragen) worden hierin opgenomen.

Toetssoftware biedt voor online **toetsafname** een *afspeelomgeving* of player aan. De player controleert de identiteit van de student door middel van een inlogproces, biedt de student een toets aan en slaat de antwoorden van de student op.

Het **nakijken** van de toets hangt af van de gekozen vraagtypen. Toetssoftware kan goed uit de voeten met gesloten antwoorden. De itemontwikkelaar (de docent) heeft vooraf het correcte antwoord ingegeven. Open vragen moeten worden nagekeken, waarbij systemen bijvoorbeeld 'blind' nakijken mogelijk maken en 'vraag-voor-vraag' nakijken, eventueel met behulp van ondersteunende software.

Sommige softwarepakketten bevatten een *analysetool* voor **toetsanalyse**, maar analyseren kan ook met behulp van bijvoorbeeld Excel of SPSS. Toetsanalyse is veelal gericht op het bepalen van de betrouwbaarheid van toetsen en het opsporen van items die van onvoldoende kwaliteit zijn.

2.3 Belangrijke begrippen

Formatief toetsen: Toetsing waarbij het leren van de toets voorop staat, meestal zonder cijfers. Door met toetsopgaven aan de slag te gaan en te leren van fouten en feedback wordt het leerproces gestimuleerd. Er is geen sprake van een minimaal te behalen niveau.

Psychometrische gegevens: Deze data zeggen iets over de moeilijkheidsgraad van vragen en geven aan in welke mate ze onderscheid maken of een student de stof wel of niet beheerst. Psychometrische gegevens kunnen worden opgeslagen in de itembank bij het item waarop ze betrekking hebben.

Summatief toetsen: Toetsing waarbij het zo nauwkeurig mogelijk meten van een bepaald niveau van vaardigheid voorop staat. De behaalde score wordt gebruikt voor formele toekenning van een behaald studieresultaat, zoals studiepunten of een diploma.

Toetskwaliteit: Het geheel van aspecten rondom toetsing dat de mate van transparantie, betrouwbaarheid en validiteit van toetsing behelst. Transparantie neemt bijvoorbeeld toe als er inzicht is in de wijze waarop items worden gemaakt. Betrouwbaarheid neemt bijvoorbeeld toe als items beter onderscheid maken tussen studenten met verschillende mate van beheersing van de stof of vaardigheid. Validiteit neemt toe als de items beter en vollediger de beoogde kennis of vaardigheid meten.

Samenwerkingsverbanden

In het handboek refereren we naar een aantal Nederlandse instellings-overstijgende itembanken. We hebben dankbaar gebruik gemaakt van hun kennis en ervaring. Het gaat om de volgende samenwerkingsverbanden:

deKennistoetsenbank MBO

Voor het MBO-opleidingsprofiel Zorg en Welzijn biedt deKennistoetsenbank MBO van Prove2Move formatieve kennistoetsen aan. DeKennistoetsenbank MBO is een methode-overstijgend digitaal systeem dat leren ondersteunt met goede kennisvragen. Prove2move is een coöperatie van 3 ROC's (Landstede, Deltion College en ROC van Twente).
https://www.prove2move.nl/dekennistoetsenbank/

IVTG

De interuniversitaire VoortgangsToets Geneeskunde (iVTG) is een meet-instrument om de kennisprogressie van studenten geneeskunde gedurende hun studie te meten. Het betreft een interuniversitaire samenwerkingsverband, bestaand uit 6 UMC's.
www.ivtg.nl

Toets en Leer

Toets & Leer is een samenwerkingsverband van 6 hogescholen met itembanken voor de onderwerpen Bedrijfsadministratie/boekhouden, Bedrijfseconomie, Belastingrecht, Marketing, Management en Recht.
www.toetsenleer.nl

10voordeleraar

Het programma 10voordeLeraar is onder regie van de Vereniging Hogescholen ontwikkeld en biedt landelijke kennisbases, verplichte kennistoetsen en een peer review systematiek voor alle lerarenopleidingen.
www.10voordeleraar.nl

3. IN 5 STAPPEN NAAR EEN ITEMBANK

In de praktijk ontstaan itembanken soms zonder bewust plan. Een uitprobeersel op kleine schaal groeit uit tot iets omvangrijks. Een dergelijke aanpak brengt het risico met zich mee dat de itembankontwikkeling in een later stadium stagneert, bijvoorbeeld door tijdgebrek of personele wisselingen. Het advies in deze is 'bezint eer ge begint'. Werk volgens beheerste stappen. Te snel gaan heeft mogelijk later consequenties voor het draagvlak of kan zelfs het project doen mislukken.

Wij onderscheiden 5 stappen voor het ontwikkelen en gebruiken van een itembank. Binnen iedere stap moet je voor een aantal zaken zorgen. Deze zaken zullen soms gelijktijdig om aandacht vragen, of in een andere volgorde dan hier beschreven. In elke stap wordt het resultaat, een in gebruik genomen itembank, concreter gemaakt; er wordt van grof naar fijn gewerkt.

Deel 2 van het handboek, het Stappenplan, neemt je mee door de verschillende stappen. Soms heb je daarvoor theoretische kennis nodig. Daarom verwijzen we regelmatig naar deel 3 van het handboek, de Verdieping. Dit deel bevat meer achtergrondinformatie voor het opzetten van een itembank.

Tip
Geen tijd om het hele handboek door te nemen? Hou dan in ieder geval de succesfactoren voor het ontwikkelen, beheren en gebruiken van een itembank uit hoofdstuk 4 bij de hand.

STAPPENPLAN

1 **VOORBEREIDING**

- Formuleer het doel en onderzoek de haalbaarheid van de itembank
- Oriënteer je op toetssystemen en uitwisselbaarheid
- Verdiep je in omvang en vraagproductie
- Maak een eerste kosten- en batenanalyse

2 **PLAN**

Maak een plan van aanpak, waarin je beschrijft:

- het doel
- het resultaat (wat hebben we als we klaar zijn?)
- hoe je denkt dit te realiseren

3 **ONTWERP**

- Orden de vragen
- Organiseer de processtappen en bepaal rollen en rechten
- Stel kwaliteitseisen vast en check of het systeem hieraan voldoet
- Kies een toetssysteem
- Leg afspraken over omvang en vraagproductie vast
- Regel de financiën
- Zorg voor een juridische verankering van het project
- Organiseer het beheer van de itembank

4 **PILOT**

- Ontwikkel de items
- Richt de itembank in en probeer uit

5 **REALISATIE**

- Neem de itembank in gebruik
- Zorg voor een toekomstbestendige itembank

4. SUCCESFACTOREN

Het ontwikkelen en gebruiken van een itembank kan complex zijn. De onderstaande succesfactoren leveren een belangrijke bijdrage aan de levensvatbaarheid van een itembank. Neem ze in acht, ook bij een relatief eenvoudige samenwerking. Deze factoren lopen als een rode draad door het handboek.

1. Bezint eer ge begint: neem de tijd bij de start, schat de haalbaarheid goed in en vind de balans tussen een projectmatige aanpak en laten ontstaan. Kies een werkwijze die past bij jouw situatie.

2. Stel een opdrachtgever aan en zorg dat het eigenaarschap van begin af aan goed belegd is. Zonder een opdrachtgever of afspraken over eigenaarschap is de kans op mislukking groot.

3. Bepaal voor het ontwikkelen van de itembank welke stappen je wilt doorlopen en welke resultaten je per stap wilt bereiken.

4. Bepaal het doel van je itembank en geef de itembank een duidelijke rol in het onderwijs. Houd hieraan vast. Draag deze visie regelmatig uit.

5. Wees reëel en manage verwachtingen: de ideale itembank ontstaat niet in één dag en misschien zelfs wel helemaal nooit. De ideale inrichting kun je niet vanachter je bureau bedenken. Durf dingen uit te proberen.

6. Kijk om je heen, er is al veel ontwikkeld. Vind het wiel niet opnieuw uit. Betrek experts.

7. Draagvlak is essentieel.
a. Docenten die elkaar kennen en vertrouwen en die het doel van de itembank wdelen, zijn een randvoorwaarde voor succes; steek energie in het leren kennen van elkaar.
b. Zorg voor goede begeleiding en ondersteuning van docenten.
c. Het bestaan van een netwerk van mensen die al samenwerken en kennis uitwisselen vergroot de kans op succes aanzienlijk.

8. Begin klein met de itembank en probeer verschillende inrichtingen en metadatamogelijkheden uit. Blijf aanpassen totdat de juiste inrichting en werkwijze is gevonden. Houd het eenvoudig en dicht bij de dagelijkse praktijk.

9. Onderschat de aard en omvang van de werkzaamheden en/of het project niet. Het is moeilijke materie om te doorgronden. Gedrag en werkwijzen van mensen veranderen gaat niet vanzelf.

10. Wees je bewust van technische beperkingen als je over instellingsgrenzen heen aan een itembank werkt.

DEEL 2 STAPPENPLAN

5. STAP 1 VOORBEREIDING

Begin met het formuleren van het doel van de itembank. Denk na over de rol van de itembank binnen het onderwijs, de mogelijke fasering en het eigenaarschap van de itembank. Wie heb je nodig? Hoe kun je stapsgewijs een itembank ontwikkelen die meerwaarde heeft voor het onderwijs en langere periode 'in de lucht' blijft? (Zie ook H10)

Ga na of het ontwikkelen van een itembank past binnen het beleid van de instelling en of het de onderwijsvisie ondersteunt. Past het binnen het digitale toetsbeleid? Welke kwalitatieve en kwantitatieve doelen wil jouw instelling met digitaal toetsen bereiken? Zorg dat je het doel van jouw itembank laat passen binnen dit beleid. Dit maakt de itembank kansrijker. Om een idee uit te werken zal er commitment nodig zijn in de vorm van tijd en geld. Iedere instelling regelt dit anders. Vaak gaat het om een bestuurlijk traject, waarbij diverse mensen praten/lobbyen op operationeel, tactisch en strategisch niveau. Er zijn meestal richtlijnen en procedures beschikbaar voor het opstarten van een project binnen een instelling. Denk bijvoorbeeld aan innovatiegelden die via onderwijsprijzen of stimuleringsmaatregelen beschikbaar kunnen komen. Mogelijk moet je bijvoorbeeld een startnotitie of projectopdracht formuleren om de middelen te regelen. Het kan zijn dat er al wordt gevraagd naar een initiële inschatting van de kosten ten opzichte van de baten.

Het is aan te raden om tijdens de voorbereiding minimaal de volgende zaken te bepalen en te beschrijven:

- De huidige situatie en de gewenste situatie.
- Het doel van de itembank. Waarom is het een goed idee? Wat hebben studenten eraan? Maak duidelijk hoe de itembank zich verhoudt tot andere beleids-documenten, bijvoorbeeld het digitaal toetsbeleid en het onderwijsbeleid.
- Welk resultaat ligt er als het project klaar is? Hoe ziet de exploitatie (onderhoud en beheer) er globaal uit?
- Vanuit welke theorie worden de items ontwikkeld? In de voorbereiding bedenk je op basis van welke theorie je aan de slag gaat: de klassieke testtheorie (KTT) of de item response theorie (IRT). Welke vraagtypes wil je gebruiken en hoe wordt er met feedback omgegaan? Keuzes over feedback en vraagtypes zijn van invloed op het eindresultaat. (Zie ook § 15.4 en § 16.3.2)
- Hoeveel tijd, geld en doorlooptijd is er nodig voor dit project?
- Bepaal de haalbaarheid van het project: voer eventueel een apart **haalbaarheids-onderzoek** of **vooronderzoek** uit.
- Bepaal wie de opdrachtgever en/of de uiteindelijke eigenaar van de itembank is. Stel vast wie je nodig hebt om de itembank te ontwikkelen. (Zie ook H14)

5.1 Haalbaarheidsonderzoek

Een mogelijke tussenstap in de voorbereiding van je itembankproject is een onder-zoek naar de haalbaarheid ervan. Hierin werk je één of meerdere opties voor het inrichten van een itembank uit. Kies welke optie het beste is voor jouw situatie. Kijk naar de volgende aspecten:

- *Inhoud:* welke overeenkomsten zijn er tussen de items die je in de itembank wilt opnemen? Is er een gezamenlijke ordening te ontdekken? In hoeverre denken de

docenten dat zij consensus kunnen bereiken over de inhoud van het vak en de items? Je moet dit weten om je aanpak te bepalen. Bij uiteenlopende meningen is er meer tijd nodig voor dit onderdeel. **(Zie ook H7)**

- *Kwaliteit:* wat is de kwaliteit van de toetsen eigenlijk op dit moment? Wat zijn de p-waarden en Rit-waarden? Hoe acceptabel is dit? Welke verbeterslag willen we realiseren met de itembank?
- De keuze voor het kopen of zelf ontwikkelen van de items.
  - Wat is er al op de markt aan itembanken? Check oplossingen buiten de instelling, bijvoorbeeld het inkopen van items van een externe commerciële partij.
  - Is er een mogelijkheid om aan te sluiten bij een bestaand samenwerkingsverband dat al een itembank ontwikkelt?
- De beoogde omvang van de itembank.
  - Hoeveel items zijn er al binnen de instelling die kunnen worden gebruikt?
  - Hoeveel items heb je naar verwachting nodig? **(Zie ook H12)**
- Techniek (toetssystemen en itembanksystemen). Ga na wat er al beschikbaar is binnen de instelling aan itembanksystemen en itembanken die je kunt gebruiken. Zoek de afdeling op die verantwoordelijk is voor het beheer en gebruik van ICTO-applicaties en specifiek de toetsapplicaties. Moet er iets door de instelling worden aangeschaft? **(Zie ook H11)**
- Organisatie: is er voldoende draagvlak voor de optie(s)? Ga in gesprek met organisaties die ervaring hebben met het inrichten van itembanken.
- Financiën: zet per optie de kosten en baten op een rij en neem de volgende zaken in acht.
  - Kan er voldoende tijd en geld worden vrijgemaakt voor het ontwikkelen van de itembank?
  - In hoeverre zal de itembank de prijs-kwaliteitverhouding van items kunnen verbeteren?
  - Kun je projectgeld of subsidiegeld aantrekken? **(Zie ook H13)**

5.2 Afronding voorbereidingsstap

Verwerk alle gegevens tot een opdrachtformulering en verkrijg goedkeuring van je opdrachtgever om een plan van aanpak te maken.

Wat heb je aan het einde van deze stap?

- Resultaat: inzicht in het doel, de haalbaarheid en de globale aanpak van het idee.
- Een GO/NO GO van je opdrachtgever voor de volgende stap.

6. STAP 2 PLAN VAN AANPAK

In deze stap leg je vast hoe de itembank eruitziet aan het eind van het project. Je maakt een plan van aanpak om dit resultaat te bereiken.

Stel jezelf de vraag: wat hebben we als we klaar zijn? Het antwoord is afhankelijk van het doel dat de itembank zal vervullen in het onderwijs. Formuleer met de betrokkenen eisen en wensen ten aanzien van de itembank. Eisen moeten verplicht worden gerealiseerd; wensen wil men graag realiseren, maar zijn niet noodzakelijk om tot een goed resultaat te komen. In je plan van aanpak zet je beide op een rij.

Het kan zijn dat een organisatie projectmatig werken faciliteert en dat er wordt gewerkt aan de hand van een projectmanagement-methode, zoals Prince II. Mogelijk zijn er zelfs projectleiders beschikbaar voor het trekken van grotere projecten. In dat geval zijn er binnen de instelling procedures gangbaar rondom werkwijze, fasering en besluitvorming. Check wat er beschikbaar is binnen de instelling over projectmatig werken. Grijp waar mogelijk terug op standaarddocumenten, zoals een sjabloon voor een plan van aanpak. Geef de stappen uit het stappenplan hierin een plek.

Zorg ervoor dat je plan van aanpak minimaal de volgende zaken zo gedetailleerd mogelijk beschrijft:

- Het doel en resultaat.
- Hoe je tot het resultaat wilt komen.
- De rol van opdrachtgever en wie dit op zich neemt. Idealiter is dit de toekomstige eigenaar van de itembank. Plan momenten in om met hem of haar de aanpak en plannen te bespreken. Stel vast waar je zijn of haar hulp nodig hebt.
- Bedenk wie je wanneer tijdens het project en erna nodig hebt. Beschikken zij over de juiste kennis en kunde? Denk aan de volgende expertises: opdrachtgever, projectleider, toetsexpert, psychometrist, functioneel beheerder, jurist en docenten.
- Bepaal hoe je draagvlak creëert. Hoe ga je ervoor zorgen dat de itembank daadwerkelijk wordt gebruikt? Hoe kun je docenten en studenten ertoe aanzetten om de nieuwe werkwijzen in de praktijk te brengen? Ga hierover met ze in gesprek. Bedenk hoe je het belang van de itembank ook op de tactische en strategische agenda's krijgt. Wie zijn de belanghebbenden?
- Kosten en baten: geef een inschatting van wat het ontwikkelen van de itembank kost en wat het de komende jaren zal opleveren.
- Kwaliteit: hoe kom je tot een itembank van voldoende kwaliteit?
- Planning: welke fasering heb je voor ogen? Welke doorlooptijd heb je nodig?
- Bedenk hoe je het gestelde doel behaalt en hoe je het proces om daar te komen kunt beheersen en zo nodig bijstellen.

Wat heb je aan het einde van deze stap?
Het resultaat van deze stap is een **plan van aanpak**, waarin het doel, het resultaat en hoe je dit denkt te realiseren zo concreet mogelijk zijn gedefinieerd.
Laat het plan van aanpak goedkeuren door de opdrachtgever. Door het plan van aanpak is de benodigde tijd van mensen afgestemd met leidinggevenden. Zo zorg je voor commitment en draagvlak.

7. STAP 3 ONTWERP

Aan het einde van deze stap ligt er een ontwerp dat aan de opdrachtgever duidelijk maakt wat hij krijgt. Daarbij weten de medewerkers van de volgende stap, het testen in een pilot, wat er moet worden gemaakt.

In het ontwerp worden de volgende zaken bepaald en vastgelegd:
- Ordening van items
- Itemontwikkelproces
- Itemkwaliteit
- Systeem
- Productie en planning
- Financiën
- Juridische aandachtspunten
- Beheer

Leg de afspraken die worden gemaakt per aandachtspunt duidelijk vast. De resultaten van elk onderdeel tellen op naar het totale ontwerp van de itembank en hangen dus ook met elkaar samen.

7.1 Ordening van items

Breng in kaart in welke mate de vakgebieden overeenkomen. Laat de betrokken docenten bepalen in hoeverre zij op dezelfde manier naar een vak of vakgebied kijken. Ze gaan met elkaar de diepte in om vast te stellen wat er op onderwerpniveau en vraagconstructie-niveau in de itembank wordt opgenomen. Deze fase is moeilijk, omdat er vaak verschillen zijn. Het is zaak om pragmatisch te blijven. Accepteer dat er mogelijk wat aanpassingen in het onderwijs noodzakelijk zijn, bijvoorbeeld door het kiezen van een nieuwe methode of het aanpassen van het gebruik van specifieke namen voor variabelen of concepten.

- Vragen die aan de orde komen: Hoe schrijf je de items uit? Wat vind je relevant of opleidingsspecifiek? Hoe zien de toetsmatrijzen eruit?
- Het kan zijn dat er verschillen zijn op notatie-of illustratieniveau. Hoe ga je hiermee om?
- Bepaal welke hiërarchische structuur de itembank krijgt en welke metadata wenselijk zijn. (Zie ook H15)

7.2 Itemontwikkelproces

- Bepaal wie welke stappen in het itemontwikkelproces doorloopt om de toetskwaliteit te borgen.
- Stel een indicatie op voor de verwachte ontwikkeltijd per item.
- Kijk naar de huidige processtappen. Scherp ze indien nodig aan of breid ze uit.
- Moeten al bestaande items door de ‘wasstraat’ om ze op te schonen? Hoeveel? (Zie ook H14)

7.3 Itemkwaliteit

- Bepaal en documenteer vuistregels waaraan de items moeten voldoen. Denk ook aan regels over taalgebruik en bronvermelding.
- Betrek waar mogelijk item-experts (niet iedereen maakt goede items) of een psychometricus.
- Plan review-begeleidingssessies onder leiding van een vraagexpert. Laat docenten een flink aantal bestaande voorbeelditems aanleveren. Bespreek gezamenlijk de variatie tussen de vakken, de wijze van vraagstelling en de constructiefouten in deze items. Hierdoor krijgen de betrokkenen ideeën over uniformering van vraagformulering en vraagvormen.
- Bedenk hoe je de gewenste kwaliteitsverbetering kunt meten.
- Begin met het einde in gedachten: inventariseer welke gegevens je uit de itembank wilt kunnen halen ten behoeve van het onderwijskundige proces. (Zie ook H16)

7.4 Systeem

- Bij een eenvoudig itembankproject: bekijk welk toetssysteem de instelling gebruikt. Ondersteunt het systeem het gewenste itemontwikkelproces, de gewenste rollen en rechten, de hiërarchische structuur en metadata?
- Bij een complex itembankproject: indien er meerdere systemen in gebruik zijn, maak een afweging welk systeem het meest geschikt is voor het doel.
- Past het gewenste itemontwikkelproces in het systeem? Dan kun je bepalen hoeveel werk het is om in te richten. Past het niet, bepaal dan of je het proces kunt aanpassen, of dat je liever onderzoekt of de structuur van het systeem kan worden aangepast.
- Geef aan wat er nodig is om de veiligheid van de items in de bank te waarborgen.
- Hoe schaalbaar moet het systeem zijn, nu en in de toekomst – voor zover je kunt overzien? (Zie ook H11)

7.5 Productie en planning

- In de voorbereiding heb je grofweg bepaald hoeveel items je nodig hebt in de itembank. Zet deze hoeveelheid uit in tijd (korte en lange termijn). Hoeveel vraagt dit van medewerkers?
- Hoeveel items zijn al beschikbaar bij de verschillende partijen en kunnen met weinig moeite bij elkaar gebracht worden?
- Bedenk hoe je het vullen van de itembank aanpakt. Bepaal de periode, eventuele deadlines (wanneer vinden de eerste toetsen plaats?) en streefproductie door opschaling (batches).
- Bepaal wanneer je een kwaliteitsreview doet op basis van toetsstatistieken en psychometrische gegevens. (Zie ook H12)

7.6 Financiën

- Werk de inschatting van de kosten en baten uit de voorgaande stappen verder uit. Bedenk hoe de initiële ontwikkeling zal worden gefinancierd.
- Stel vast hoe je de exploitatie van de itembank na afronding van het project gaat financieren. Stel een exploitatieplan op. Mogelijkheden om inkomsten te genereren zijn bijvoorbeeld:
  - Instellingen nemen een abonnement op de itembank, om deze vervolgens kosteloos aan studenten aan te bieden.
  - Studenten betalen voor toegang tot de itembank.

- Betrek de opdrachtgever van de itembank bij het opstellen van het budget voor exploitatie. Zorg dat hij of zij zich hiervoor verantwoordelijk voelt. (Zie ook H13)

7.7 Juridische aandachtspunten

- Stel vast wie de eigenaar is van de items in de itembank.
- Bedenk hoe je wilt omgaan met materiaal van derden dat auteursrechtelijk beschermd is, zoals afbeeldingen bij items.
- Bepaal wat je nodig hebt op het gebied van de bescherming van persoonsgegevens.
- Overleg welke afspraken je over de samenwerking met betrokken partijen wilt vastleggen. Leg minimaal vast hoeveel uur wordt geleverd voor itemontwikkeling. Betrek een juridisch expert bij instellingsoverstijgende samenwerkingen om te beoordelen wat de beste samenwerkingsmogelijkheid is en wat hierbij komt kijken. (Zie ook H17)

7.8 Beheer itembank

- Bedenk hoe je de gebruikers, zoals docenten en studenten, het beste kunt ondersteunen om het eindproduct in gebruik te nemen. Ga hiervoor met hen in gesprek en luister naar argumenten. Laat hen juist bij het ontwerp meedenken.
- Bedenk wie of wat je nodig hebt om de itembank te beheren. Wie heb je nodig op het gebied van ondersteuning en beheer na het project?
- Maak instructiemateriaal en plan instructiemomenten.
- Ga in gesprek met de eventueel beschikbare toets- en beheerorganisatie in de instelling en stem de workflow af. (Zie ook H14)

Wat heb je aan het einde van deze stap?
Een ontwerp van de itembank. Alle tussenresultaten per onderdeel tellen op tot het ontwerp van de itembank:

- *Structuur:* een landkaart/tabel met de onderwerpstructuur die is overeengekomen en afspraken omtrent bijvoorbeeld notatiewijzen.
- *Organisatie:* aanpak en gedocumenteerde organisatorische afspraken (beschrijving van de toets- en beheerorganisatie) over hoe de itembank zal worden beheerd, inclusief besturing van de itembank, inclusief benodigde kennis en kunde van betrokkenen (een opleidingsplan).
- *Systeem:* inrichtingsplan systeem; dit betreft workflow (van het itemontwikkelproces) en de metastructuur (hiërarchische structuur en metadatastructuur).
- *Kwaliteit:*
  - afspraken over de gewenste kwaliteit van de items, onder andere een beschrijving van richtlijnen en criteria waaraan de items moeten voldoen.
  - beschrijving van het werkproces voor itemontwikkeling, inclusief beheer, analyse en de rollen en rechten.
  - een indicatie van de tijd die het vergt om een item te ontwikkelen.
  - een indicatie van de toename en consistentie van de itemkwaliteit.
- *Productie en planning:* aanpak en gedocumenteerde organisatorische afspraken hoe de itembank 'te vullen' en op niveau te houden.
- *Financieel:* een businesscase en exploitatieplan.
- *Juridisch:* vastgelegde afspraken over de juridische aspecten van de items en het systeem

8. STAP 4 PILOT

Het ontwerp is klaar. Nu richt je de itembank in en maak je hem gereed voor gebruik. In deze stap wordt een itembankstructuur opgezet, metadata-categorieën ingevoerd, accounts aangemaakt, bestaande items geconverteerd, rollen en rechten verdeeld.

Voer een pilot uit waarbij je een eerste toetsperiode volledig draait. Je test in de pilot of de keten van activiteiten werkt. Maak zoveel items dat je een eerste toetsperiode kan afdekken. Check of de gemaakte afspraken haalbaar zijn. Kloppen de uitgangspunten? Weet iedereen welke rollen en rechten hij of zij heeft? Test of de systemen en de workflow technisch werken. En worden de items door studenten gebruikt zoals beoogd? Worden formatieve toetsen wel gebruikt? Evalueer en pas aan indien nodig.

De volgende activiteiten vinden plaats in de pilot:

- Richt de itembank technisch in en test hem.
- Indien er sprake is van een conversie van oude items naar de nieuwe situatie: voer de conversie uit met een kleine batch.
- Leid medewerkers op.
- Voer een eerste batch van items in op basis van workflow en afspraken.
- Stel een toets samen met de items.
- Bied de toets aan in een onderwijsonderdeel.
- Evalueer het gebruik en de resultaten van de toets.
- Zet een administratief systeem op voor personen en financiën (bijvoorbeeld als er is gekozen voor een vergoeding per geconstrueerde item of coördinatorschap).
- Evalueer het gebruik en de werkwijzen voor het aanmaken van items. Kan de kwaliteit beter onder controle worden gebracht? Kloppen de uitgangspunten, bijvoorbeeld de ingeschatte tijd voor vraagontwikkeling?
- Waar nodig: stuur bij, pas gemaakte afspraken aan, stel uitgangspunten bij.

Wat heb je aan het einde van deze stap?

- Een gedocumenteerde evaluatie van de pilot op de diverse inrichtingsaspecten en vanuit het oogpunt van de gebruikers inclusief een verbeterplan.
- Een ingerichte en geteste itembank.

9. STAP 5 REALISATIE

In deze stap neem je de itembank in gebruik. Ook onderneem je actie om het gebruik van de itembank te laten toenemen.

De pilot heeft een aantal lessen opgeleverd. Op basis hiervan kun je grootschaliger te werk gaan. Tijdens de realisatie worden grotere hoeveelheden items gemaakt en worden de toetsen vaker ingezet.

Aandachtspunten:

- Zorg dat de eigenaar van de itembank zich blijft bezig houden het volgen en bewaken van de doelstellingen en resultaten.
- Bespreek de borging regelmatig met de eigenaar van de itembank. Maken de itemontwikkelaars voldoende items? Zijn de items van goede kwaliteit? Hoe kun je zorgen voor zo min mogelijk variatie in kwaliteit? Is iedereen tevreden met het proces? Kan de docent gemakkelijk toetsen samenstellen? Vinden studenten de ontwikkelde formatieve items en feedback aantrekkelijk?
- Ga na of de businesscase haalbaar blijft. **(Zie ook H16)**

9.1 Nazorg

De itembank is geïmplementeerd. Het proces is in productie, iedereen vervult zijn taak en gebruikt de itembank op de afgesproken manier. Er is een formele eigenaar die het doel van de itembank in de gaten houdt. De itembank heeft een plek gekregen in de organisatie: er zijn mensen die ervoor zorgen dat de itembank blijft bestaan en goed wordt gebruikt. De onderwijskundig meerwaarde is evident, nu en in de toekomst.

Aandachtspunten:

- Zorg dat het onderhoud van de itembank is belegd, bijvoorbeeld bij een functioneel beheerder en de redacteur(s).
- Zorg dat de financiële administratie functioneert zoals voorzien voor de exploitatiefase. Blijf in de gaten houden of de exploitatie een passende toekomstige begroting oplevert.
- Zorg voor goede documentatie, zoals:
  - een beschrijving van de inzet van de itembank in de cursus.
  - een beschrijving van het onderhoud van de itembank voor de toets- en beheerorganisatie.
  - een beschrijving van het onderhoud van de itembank voor een eventueel coördinatieteam en voor elk auteursteam per itembank. Zie § 16.2.3.
  - een standaard instructieprogramma voor gebruik van het itembanksysteem voor nieuwe medewerkers.
  - aanpassingen aan het instructieprogramma bij updates van en aanpassingen aan het itembanksysteem.

Wat heb je aan het einde van deze stap?

- Een in gebruik genomen en beheerde itembank.

DEEL 3 VERDIEPING

INFORMATIE EN ACHTERGROND BIJ HET ONTWIKKELEN VAN EEN ITEMBANK

10. WAAROM (SAMENWERKEN AAN) EEN ITEMBANK?

Tijdens de voorbereiding leg je het doel van de itembank vast. In dit hoofdstuk beschrijven we een aantal mogelijke manieren waarop een itembank impact kan hebben op het onderwijs. We gaan ook nader in op de voordelen van samenwerken aan een itembank. Ten slotte bekijken we welke factoren binnen het vakgebied bijdragen aan een succesvol gezamenlijk itembankproject.

10.1 Waarom een itembank?

Bepaal welke rol de itembank zal gaan spelen in het leer- en toetsproces.
Vaak wordt een onderwijskundige impact nagestreefd met de inrichting van een itembank, zoals kwaliteitsverbetering of een verhoging van de efficiency.
Voorbeelden van beoogde impact zijn:

- *Psychometrisch:* het verhogen van betrouwbaarheid en validiteit van toetsen.
- Verantwoording en transparantie: door de komst van de itembank komen er meer mogelijkheden om de toetskwaliteit transparant te verantwoorden. Er is onder andere goed zicht op de totale itembank en de match met de toetsmatrijs.
- *Borging:* een beter controleerbaar en minder foutgevoelig productieproces van items, bijvoorbeeld door controleerbaar versiebeheer van items.
- *Verduurzaming:* items liggen niet meer 'in de la' van een docent, maar worden docent-onafhankelijk beheerd in de itembank. Ze verdwijnen dus ook niet als een docent vertrekt.
- *Toetsveiligheid:* het is simpel om meerdere toetsversies te produceren en zo spieken tegen te gaan.
- *Toetsveiligheid:* het wordt gemakkelijker om toetsen op een veilige manier te ontwikkelen, bijvoorbeeld doordat items niet meer via mail worden doorgestuurd.
- *Efficiency:* door hergebruik van items nemen de kosten af. Items hoeven maar eenmaal te worden ingetypt, redactie en verbetering gaan veel sneller en items kunnen gemakkelijker worden geselecteerd voor opname in een toets.
- *Onderwijskwaliteit:* door hergebruik van items kunnen docenten meer tijd aan onderwijs besteden.
- *Onderwijskwaliteit:* er ontstaan meer oefenmogelijkheden en mogelijkheden voor niveaudifferentiatie.
- *Onderwijskwaliteit:* er kunnen snellere herkansings- en betere feedback-mogelijkheden worden gerealiseerd.
- *Onderwijskwaliteit:* het studiesucces wordt hoger, doordat studenten kunnen oefenen en de stof daardoor beter beheersen.
- *Utiliteit:* investeringen in al ontwikkelde items worden effectiever terugverdiend.

10.2 Waarom samenwerken aan een itembank?

Zowel binnen de instellingen als tussen instellingen valt veel profijt te behalen uit samenwerking. Het gezamenlijk opzetten en gebruiken van itembanken kan zelfs op korte termijn leiden tot een aanzienlijke kwalitatieve en kwantitatieve winst. Hoe komt dat?

Ten eerste moeten docenten die samen een itembank inrichten het gesprek aangaan in een eenduidige 'taal'. Ze bepalen samen hoe ze over itemontwikkeling spreken en welke minimale eisen ze stellen ten aanzien van kwaliteit. Dit is op zich al kwaliteitsverhogend.

Ten tweede ligt bij samenwerking de inzet van toetsspecialisten voor de hand. Zij kunnen een bijdrage leveren aan de toetsinhoud, maar ook op het gebied van toetskundigheid, redactie, opmaak et cetera. Ook dit zal de kwaliteit van items verhogen. Zie § 16.2 over verschillende rollen bij itemontwikkeling.

Ten derde maak je bij samenwerking meer afspraken over het proces van itemontwikkeling. Dat betekent dat de itemontwikkelaars zich bewuster worden van de verschillende ontwikkelingsfasen, de verantwoordelijkheid en de borging van het proces. Als itemontwikkelaars zich aan deze afspraken houden, leidt dat tot kwaliteitsverhoging. Er is betere grip op kwaliteit doordat het proces is geborgd.

Tot slot kan samenwerking zorgen voor lagere kosten. In eerste instantie zal een samenwerking meer kosten met zich meebrengen, maar als eenmaal ontwikkelde items vaker en door meer docenten worden ingezet en dus door meer studenten worden gebruikt, wordt het gebruik van de items per saldo goedkoper. Voor een verdere toelichting zie hoofdstuk 13: Kosten en baten.

10.3 Kansrijke vakgebieden voor een succesvolle samenwerking

Een gedeelde visie op het vakgebied kan eerder leiden tot een succesvolle itembank. Hoe meer stakeholders in het vakgebied het eens zijn over de indeling van de leerstof, hoe meer het voor de hand ligt om aan itembanken te werken. Vakgebieden met een gemeenschappelijke Body of Knowledge (BoK), zoals de Lerarenopleidingen of het Medisch domein, zijn hier voorbeelden van. Ook vakgebieden waarbij de meeste opleidingen inleidende cursussen verzorgen, zijn geschikt voor het inrichten van gezamenlijke itembanken. Denk bijvoorbeeld aan Inleiding Psychologie, Inleiding in de Sociologie, Inleiding Neurobiologie, en Inleiding programmeren.

Stabiele vakgebieden lenen zich goed voor gezamenlijke itembanken die je met weinig onderhoud kan inrichten. Klassieke vakgebieden zoals logica, wiskunde, statistiek, mechanica, boekhouden en basiskennis microbiologie lenen zich goed voor een dergelijke samenwerking. Eenmaal ontwikkelde vragen verouderen niet of niet snel. Bij vakgebieden die zich gestaag ontwikkelen door nieuwe inzichten is het onderhoud van de items belangrijk. Verouderde items moeten worden opgespoord, verwijderd of aangepast. Dat proces moet systematisch worden ingericht om de kwaliteit van de vragenbank te kunnen borgen. Eenmaal per jaar gaat bijvoorbeeld de stofkam door alle items. In een samenwerkingsverband zijn hierover afspraken nodig. Voorbeelden van vakgebieden waarvoor dit geldt, zijn het medisch vakgebied, verpleegkunde, toerisme en geschiedenis. Samenwerking om kosten te delen en kwaliteit te verhogen, is hier een interessante driver.

Vakgebieden die inhoudelijk voortdurend in ontwikkeling zijn (volatiele vakgebieden), vragen nog meer nadruk op de inrichting van het itembanksysteem en het borgingsproces. Denk aan vakken waarbij recente wetgeving altijd van belang is, zoals Rechten, waar steeds aanpassing aan protocollen plaatsvindt, zoals Verpleging of waar thematisch wordt gewerkt. Er is continue aandacht nodig voor screening, aanpassing, verwijdering of aanvulling. Hiervoor moet een proces worden ingericht.

In vakgebieden waarin veel studenten worden onderwezen, zijn itembanken ook meer geschikt om gezamenlijk aan te werken, omdat de investeringskosten en onderhoudskosten per student dan veel lager zijn.

11. TECHNOLOGIE: SYSTEMEN

Tijdens de voorbereiding van het project doe je onderzoek naar het te gebruiken itembanksysteem. Bij een klein project is dit waarschijnlijk het centrale toetssysteem van de instelling. Zoek binnen de instelling de afdeling op die verantwoordelijk is voor het beheer en gebruik van ICTO-applicaties en specifiek de toetssystemen en bespreek je wensen en eisen ten aanzien van de itembank.

Primair moeten itembanksystemen het mogelijk maken om:
- items goed te kunnen invoeren, ordenen, zoeken en selecteren
- aan kwaliteitsborging van de items te werken.

Toetssystemen zijn er in vele soorten en maten en bieden een grote hoeveelheid functionaliteiten die tegemoetkomen aan deze eisen.

De kerncomponenten van een digitaal toetssysteem zijn de auteursomgeving, de itembank, de afspeelomgeving en de analysetool. In hoofdstuk 14 gaan we dieper in op de achterliggende processen van deze componenten.

11.1 Veel gebruikte toetssystemen voor onderwijs

Deze paragraaf geeft een overzicht van de meest gebruikte toetssystemen voor het hoger onderwijs. De lijst is samengesteld door experts die betrokken zijn bij de ontwikkeling van dit handboek. De markt van toetssysteem-leveranciers is dynamisch. Daarom is deze lijst niet compleet.

Veel gebruikte toetssystemen:
- TestVision Online: http://www.testvision.nl
- RemindoToets Toets: https://www.paragin.nl/remindotoets/
- MapleTA: https://www.maplesoft.com/products/mapleta/
- Cirrus Assessment: http://www.cirrusassessment.com/
- QMP: http://www.questionmark.com
- iQualify: www.infoland.nl/producten/iqualify
- Surpass: http://www.surpass.com
- Andriessen TeleToets: http://www.andriessen.nl/page/2/
- Assess.com: FastTest, http://www.assess.com/fasttest/
- Testlife: http://www.testlife.eu
- Quayn: http://www.drp.nl/Producten/WinToets-40-WEB.html
- TAO testing: http://taotesting.com

Quiz-/ Toetssystemen in Learning Management Systemen met beperkte itembankfunctionaliteit:
- Canvas: https://www.canvaslms.com/
- Blackboard: http://www.blackboard.com
- Desire2Learn BrightSpace: https://www.d2l.com/products/learning-environment/
- Moodle: https://moodle.com/
- N@tschool: http://www2.threeships.nl/nl-nl/producten/ntschool.aspx
- Drillster: https://www.drillster.com/info/

Itembanksystemen(Stand Alone programma's) met directe itemexportfunctionaliteit:
- Respondus: https://www.respondus.com/
- Wimba Diploma: http://www.wimba.com/products/diploma
- ExamView: https://www.turningtechnologies.com/examview

Toetssystemen met vragen:
- SOWISO (Wiskunde): https://sowiso.nl/nl/
- I Hate Statistics (Statistiek): https://www.ihatestatistics.com/
- Zeer Actieve Psychologie: http://zap.psy.utwente.nl/
- R/Exams (Statistiek): http://www.r-exams.org/
- Drillster: https://www.drillster.com/info/
- Quizlet: https://quizlet.com/
- ProProfs: https://www.proprofs.com/
- Itembanken van uitgeverijen, vaak behorende bij een specifiek studieboek of online leeromgeving (bijv. Pearson MyLab series (https://www.pearsonmylabandmastering.com/), McGraw Hill methoden, WebAssign e.d.)

11.2 Toetsitems uitwisselen tussen systemen

Niet iedere onderwijsinstelling in Nederland gebruikt dezelfde toetssoftware. Hoe werk je dan samen aan itembanken? Vaak is het mogelijk om eenvoudige meerkeuzevragen te converteren van het ene naar het andere systeem. Instelling A maakt bijvoorbeeld de vragen en instellingen B en C converteren ze naar hun eigen toetssysteem. In het ideale geval werk je echter in dezelfde auteursomgeving, hanteer je een gezamenlijk itemontwikkelproces en vloeien alle gemaakte vragen en data over de gemaakte items terug naar één itembank.

11.2.1 QTI-standaard

In de afgelopen jaren is onder meer in SURF-verband onderzoek gedaan naar de interoperabiliteit en uitwisselbaarheid van vragen tussen verschillende systemen. In internationaal verband wordt gewerkt aan de IMS QTI-standaard[5] (Question and Test Interoperability) voor de uitwisseling tussen toetssystemen. Deze standaard zal niet op korte termijn alle uitwisselingsvraagstukken oplossen. In de praktijk zien we dan ook dat een complexere samenwerking van meerdere instellingen voor de meeste deelnemende partijen betekent dat ze moeten werken in een ander systeem dan hun eigen instellingssysteem.

Als je bij een samenwerking niet in één itembanksysteem kan werken, beperk dan het te gebruiken vraagtype tot de multiplechoicevraag. Deze vorm is zo eenvoudig dat uitwisseling via QTI of platte tekst meestal probleemloos verloopt. Ga ervan uit dat metadata en media (afbeeldingen, geluid) niet (goed) meekomen als je items ex- en importeert.

Voer in de voorbereiding een proef uit met het uitwisselen van items. Zo kun je bepalen of het overzetten goed gaat of dat er technische ingrepen nodig zijn, of conversiegereedschappen. Probeer zo veel mogelijk in hetzelfde systeem te werken, ook vanuit verschillende instellingen.

[5] *https://www.imsglobal.org/question/index.html*

11.2.2 LTI en SCORM standaard

In de toetswereld wordt vaak gewerkt met de standaarden IMS LTI (Learning Tools Interoperability standard) en SCORM (Sharable Content Object Reference Model). Dit zijn géén uitwisselstandaarden voor items. Deze standaarden worden gebruikt om toetssystemen te koppelen aan andere systemen, zoals digitale leeromgevingen (Learning management systemen zoals Blackboard of Canvas, oftewel LMS-en). Daarbij worden met name docent- en studentgegevens tussen de systemen automatisch uitgewisseld, kunnen toetsen vanuit het LMS worden gestart en worden de scores op toetsen teruggevoerd naar de cijferlijstfuncties in het LMS. Voor de eindgebruikers ontstaat een meer eenduidige beleving van het LMS en ze hoeven niet extra in te loggen.

11.3 Veiligheid en schaalbaarheid van itembanken

Met name summatieve items moeten in veilige omstandigheden worden ontwikkeld en opgeslagen in itembanken. Hoe hoger het belang van de toetsen, hoe meer aandacht er nodig is voor veiligheid.

Ook tijdens de afname van toetsen wil je verspreiding van items tegengaan, vooral bij gekalibreerde vragen en kleine banken. Bij computerafname is het belangrijk om een afgeschermde omgeving te hebben (secure browser), maar ook goed te controleren op smartphones, brillen, horloges en andere devices die foto's kunnen maken van het scherm of van papier. Professionele surveillanten blijven noodzakelijk.

Meer informatie is te vinden in het SURF werkboek *Veilig toetsen*.[6]

Schaalbaarheid hangt nauw samen met de gewenste omvang van de itembank. Houd rekening met het toevoegen van eventuele externe auteurs, de performance van de auteursomgeving, een mogelijke groei in het aantal vragen en toe te voegen onderwerpen en metadata.

[6] *https://www.surf.nl/kennisbank/2016/werkboek-veilig-toetsen.html*

12. OMVANG VAN DE ITEMBANK

Hoe bepaal je de hoeveelheid items van een itembank? Tijdens de voorbereiding denk je al na over de benodigde omvang, omdat dit richting geeft aan de hoeveelheid tijd en middelen die nodig zijn om de bank te realiseren. Soms worden indrukwekkende hoeveelheden items genoemd, maar in feite is de benodigde omvang afhankelijk van veel factoren.

In de eerste plaats moet je weten hoeveel items je in een toets wilt opnemen. Mag en kan je de items hergebruiken? Wil je de items in een summatieve of formatieve toets gebruiken? Dit kan verschil maken voor de benodigde hoeveelheid.

In de volgende paragrafen schetsen we scenario's voor verschillende toetsen waarvoor je de items wilt gebruiken.

12.1 Summatieve toetsen met een gelijktijdig toetsmoment voor alle studenten

De meeste toetsen in het hoger onderwijs bestaan uit klassieke tentamens. Eén docent of team stelt een toets samen voor een groep kandidaten voor één cursus van één instelling. Alle studenten maken de toets op hetzelfde moment. Welke situatie treedt dan op?

Stel dat de toetsvragen **niet** worden vrijgegeven na afloop van de toets, dat je 3 verschillende varianten van de toets samenstelt en dat de toets uit 50 items bestaat. Dan is een itembank van 150 items afdoende.

Stel dat de toetsvragen **wel** worden vrijgegeven na afloop van de toets, dat je 3 verschillende varianten van de toets samenstelt en dat de toets uit 50 items bestaat. Dan is een itembank nodig waarbij de omvang bij elke toetsafname moet *groeien* met 150 items. Na 5 jaar zitten er 750 items in de itembank. Instellingen leggen vaak vast in hun toetsbeleid wat wel en niet toegestaan is ten aanzien van het hergebruik van items.

Stel dat de toetsvragen **niet** worden vrijgegeven na afloop van de toets, dat er voor elke positie van een vraag in de toets random gekozen wordt voor 1 uit 2 items en dat de toets uit 50 items bestaat. Een itembank van 100 items is afdoende.

Stel echter dat de toetsvragen **wel** worden vrijgegeven na afloop van de toets, dat er voor elke positie van een vraag in de toets random gekozen wordt voor 1 uit 2 items en dat de toets uit 50 items bestaat. Dan is een itembank nodig waarbij de omvang bij elke toetsafname moet groeien met 100 items. Na 5 jaar heb je een itembank met 500 items.[7]

12.2 Summatieve toetsen die op elk moment kunnen worden afgenomen

Er zijn ook toetsen die worden afgenomen op aanvraag van de student. Daarbij bestaat het gevaar dat studenten items gaan verzamelen en met elkaar gaan delen.

[7] *Meer informatie: Draaijer, S. & Klinkenberg, S. (2015). A practical procedure for the construction and reliability analysis of fixed-length tests with randomly drawn test items. In E. Ras & D. J. Brinke (Eds.), Computer Assisted Assessment. Research into E-Assessment (pp. 47–60). Springer International Publishing. https://doi.org/10.1007/978-3-319-27704-2_6*

Of dat studenten zich vaker voor de toets inschrijven in de hoop dezelfde items te beantwoorden waarop ze dan het correcte antwoord weten.

Over dergelijke toetsen zegt men soms dat het niet erg is dat de toetsvragen bekend zijn, als de bank maar groot genoeg is. Immers, als studenten kunnen oefenen met alle items, leren ze vanzelf de stof. Er zal echter ook een groep studenten proberen de correcte antwoorden letterlijk uit het hoofd te leren, waardoor er geen echte kennis of begrip wordt verworven. Als ze hierin slagen, maakt dit de toets minder valide. Hoeveel items er nodig zijn om dit tegen te gaan, is niet te zeggen. Itembanken van minimaal 1000-2000 items lijken noodzakelijk.

Het is in alle gevallen aan te raden studenten nieuwe vragen te laten beantwoorden. Een minimale variant is dat alternatieven uit multiplechoicevragen variëren van inhoud of van positie. Daarnaast zou gebruik kunnen worden gemaakt van geparametriseerde items. Dit zijn vragen waarbij getallen, objecten, concepten of begrippen worden getrokken uit een verzameling.[8]

Casus OU

Bij de start van het itembankproject van de Open Universiteit gold de richtlijn dat er 15 keer zoveel items in de bank moesten zitten als het aantal items in een toets. Deze eis was afgeleid van 3 tentamenmogelijkheden per jaar en de aanname dat studenten binnen 5 jaar de opleiding zouden afronden.

Voor docenten was het bijna onhaalbaar om zoveel items te maken, hoewel het voor ongeveer 50 itembanken is gelukt. De OU heeft de programmatuur laten aanpassen, zodat studenten pas na 5 keer weer eenzelfde item kregen voorgelegd. Daarmee kon de eis van 15 keer verlaagd worden naar 7 keer; een opluchting voor docenten.

Een andere eis was dat de itembank na een jaar psychometrisch nuttige analyses kon opleveren. Elk item moest dan minimaal 30 keer beantwoord zijn. Voor sommige banken was daarmee de eis van 15 keer ook veel te hoog, bijvoorbeeld wanneer er maar 10 studenten op jaarbasis het tentamen aflegden.

12.3 Formatieve diagnostische toetsen

Formatieve diagnostische toetsen die het klassieke tentamen representeren, komen overeen met het eenmalig maken van één tot drie representatieve toetsen. Het maakt niet uit of de toetsvragen uitlekken. Het is immers de verantwoordelijkheid van de student zelf om al dan niet serieus een diagnostische toets te maken.

Is de insteek dat studenten heel vaak een diagnostische toets kunnen afnemen? Dan lijkt de verstandigste keus om te bepalen hoeveel van dergelijke toetsen je wilt aanbieden. De hoeveelheid items die nodig is voor summatieve toetsen die op elk moment kunnen worden afgenomen is hier niet nodig.

12.4 Formatieve oefentoetsen

Formatieve oefentoetsen vragen minder items per onderwerp. De student heeft vooral goede instructieve items nodig over belangrijke, moeilijke onderwerpen. Per onderwerp zijn wellicht een stuk of 20 items voldoende. Als een vakgebied uit 20 onderwerpen bestaat, zijn er ongeveer 400 items nodig.

[8] *Meer lezen? Gierl, M. J., & Haladyna, T. M. (2012). Automatic item generation: Theory and practice. New York, NY: Routledge.*

13. KOSTEN EN BATEN

Tijdens de voorbereiding schat je de financiële haalbaarheid van je project in. Je maakt een globale inschatting van de te verwachten investeringskosten en exploitatiekosten van de itembank. Daarnaast kijk je naar de kwantitatieve en kwalitatieve voordelen of baten die de itembank mogelijk oplevert. Je probeert antwoord te vinden op de vragen 'waarom investeren in een itembank?' en 'wegen de kosten op tegen de baten?' Het uitwerken van een kosten-batenanalyse biedt hiervoor uitkomst. Bij grote investeringsplannen wordt steeds vaker een businesscase gevraagd. Een businesscase geeft antwoord op de vraag 'waarom willen we dit eigenlijk?'. Het is een onderbouwing van toekomstig besluit en draagt eraan bij dat iedereen begrijpt waarom het project belangrijk is.

Het ontwikkelen van itembanken kan onderdeel zijn van een groter implementatietraject van digitaal toetsen bij instellingen. Hier ligt vaak een businesscase aan ten grondslag. De kosten en baten van een te ontwikkelen itembank is dan één van de onderdelen.[9]

13.1 Kosten

Voor de ontwikkeling van een itembank is een investering nodig. Dit zijn de projectkosten. Aansluitend heb je geld nodig voor exploitatie voor het doorontwikkelen en het beheer.

Projectkosten

De initiële investering bestaat uit het creëren van een samenwerkingsverband, het ontwikkelen van een plan, het opzetten van de organisatie, het overleggen en onderhandelen over de inrichting van de bank en werkwijzen, het licenseren en opzetten van het beheer van een itembanksysteem, het ontwikkelen van een eerste hoeveelheid items voor de itembank en het afnemen van de eerste toetsen met de items uit de itembank.

Neem voor kosten de uren van de projectleden op en vermenigvuldig die met hun uurtarief. Neem kosten van externen op, net als kosten van middelen (systeem, ruimtes, reiskosten, opleidingen, materialen et cetera). Neem ook een post onvoorzien op, die je kunt onderbouwen met mogelijke risico's.

Een aantal mogelijkheden om de ontwikkeling van complexere itembanken te financieren:

- Bij de ontwikkeling van de itembanken wordt 'met gesloten beurzen' gewerkt. Er worden afspraken gemaakt wie hoeveel items maakt en reviewt en in welke termijnen. Onderling worden dus geen rekeningen verstuurd.
- Itemontwikkelaars krijgen een vergoeding per geconstrueerd en goedgekeurd item. Dan maakt het niet zoveel uit welke docent of instelling de meeste items maakt. In dit geval zal er wel out-of-pocket geld beschikbaar moeten zijn om itemontwikkelaars te bekostigen.
- De coördinerende instelling of docent krijgt een aparte vergoeding, omdat de coördinator relatief meer tijd kwijt is en een grotere verantwoordelijkheid heeft.

[9] *Voor meer informatie over de businesscase van digitaal toetsen: de whitepaper 'De businesscase van digitaal toetsen', Michiel van Geloven (PDF) en de thema-uitgave Digitale toets- en itembanken in het onderwijs, jan 2017 (hoofdstuk Kosten en baten van een itembank, bladzijde 12)*

Exploitatie

Na afloop van het project moet de itembank worden onderhouden en door-ontwikkeld. Tijdens het project maak je een plan hoe dat eruit ziet. In dit zogenoemde exploitatieplan leg je onder andere vast dat medewerkers tijd krijgen om items te blijven ontwikkelen, de itembank in de lucht te houden en te beheren. Ga ervan uit dat eens in de zoveel jaar alle items worden nagelopen, bijvoorbeeld om ze te hercoderen vanwege curriculumwijzigingen, nieuwe boeken et cetera.

Bedenk dat zonder een gezamenlijke itembank de kosten voor het ontwikkelen van items veelal niet transparant zijn. Het kost tijd en geld om elke keer op individuele basis items te maken, die mogelijk van variërende kwaliteit zijn.

13.2 Baten

Wat zijn de baten van de itembank? In § 10.1 staat een opsomming van mogelijke vormen van impact van de itembank. Probeer die kwaliteitsdoelen zo te formuleren dat ze in financiële zin kunnen worden uitgedrukt. Vaak is dat niet goed mogelijk en komt het aan op de kracht van argumenten.

Baten kunnen bestaan uit structurele exploitatie, waarbij financiering wordt geborgd door bijvoorbeeld lump-sum financiering, contributie of financiering op basis van toetsafname door studenten of docenten. Breng vooraf in kaart welke wijze van bekostiging je wilt hanteren. Zorg ervoor dat partijen zich al in de voorbereiding hieraan committeren.

Baten kunnen ook bestaan uit lagere kosten en toegenomen kwaliteit. Eenmaal ontwikkelde items worden immers gebruikt door meerdere docenten en vooral door meerdere studenten. Dat betekent dat er weliswaar kosten zijn gemoeid met het ontwikkelen van een itembank, maar dat per saldo de verhouding tussen kosten en kwaliteit per afgenomen toets per student beter is.

14. ORGANISATIE: PROCESSTAPPEN, ROLLEN EN RECHTEN

In het ontwerp van de itembank organiseer je de ontwikkeling en het beheer. Je bepaalt wie welke taken uitvoert en hoe dat gebeurt. Inzicht in de globale processtappen die bij een itembank horen, kan helpen om deze taken op een rijtje te krijgen.

14.1 Componenten en processtappen

Itembanksystemen bestaan vaak uit verschillende componenten met bijbehorende functies. Kerncomponenten zijn meestal een auteursomgeving, een itembank, een afspeelomgeving en een analysegereedschap. Binnen elk van die kerncomponenten zijn processtappen te onderscheiden. Neem de functionaliteiten van de itembank. Hierin onderscheiden we processen om items te kunnen invoeren, controleren, wijzigen, voorzien van metadata, ordenen en selecteren.

In figuur 2 staat een uitwerking van een aantal processtappen die binnen de verschillende componenten van een toetssysteem kunnen plaatsvinden.

Figuur 2. Bron: SURF Begrippenkader voor digitaal toetsen, 2013.

14.2 Rollen en rechten

Wie mag bij welke informatie in de itembank? Wie mag wat veranderen of toevoegen? Dat bepaal je door het toekennen van rollen en rechten. In de meeste itembanksystemen liggen de mogelijke rollen en rechten vast. Je kunt er zelf personen aan koppelen. Sommige systemen maken het mogelijk om zelf rollen te definiëren. Bepaal bij voorkeur eerst welke rollen en rechten je wilt kunnen onderscheiden, voordat je een keuze maakt voor een systeem.

Voorbeelden van mogelijke rollen:

Beheer

- *Super user:* mag bij alle items en kan rechten instellen voor alle gebruikers.
- *Coördinator:* heeft toegang tot een deel van de itembank en een deel van de gebruikers en kan hier wijzigingen in aan brengen.

Itemontwikkeling

- *Auteur:* mag items toevoegen en bewerken.
- *Controleur:* mag items wijzigen en van commentaar voorzien.
- *Psychometrisch specialist:* mag metadata toevoegen en bewerken en commentaar geven op items.
- *Redacteur:* zorgt voor controle op taal- en spelfouten, uniform taalgebruik en uniforme lay-out van items.
- *Illustrator/multimediaspecialist:* zorgt voor uniforme en goed bruikbare afbeeldingen en voegt deze in de items.
- *Copyright controleur:* zorgt dat gebruik van afbeeldingen en teksten juridisch gelegitimeerd is.
- *Gatekeeper:* zorgt voor een finale check van een item.
- *Lid examencommissie of lid visitatiecommissie:* kan kijken in de itembank om vast te stellen dat een eventueel itembankbeleid in een opleidingsplan wordt uitgevoerd.
- *Externe validator:* inspecteert de itembank als iemand van buiten het ontwikkelteam.

Voorbeelden van verschillende rechten in een itembanksysteem:

- Toegang tot alle items.
- Toegang tot delen van de itembank.
- Leesrechten, bewerkrechten, verwijderrechten, verplaatsrechten, toevoegrechten, commentaarrechten, metadateringsrechten.
- Rechten om toetsen samen te stellen.
- Rechten om toetsen klaar te zetten voor afname.
- Rechten om toetsanalyses uit te voeren.

Tip

Wees spaarzaam met het toekennen van verschillende rollen en rechten. Hoe complexer het aantal rollen en rechten en de workflow, des te meer tijd en energie het redactieteam kwijt is aan de inrichting en het beheer hiervan. Een gebrek aan goed beheer zal de gebruiksvriendelijkheid verminderen en het proces van itemontwikkeling stagneren.

14.3 Geautomatiseerde workflow

De huidige toetssystemen bieden in meer of mindere mate werkprocesondersteuning voor samenwerking aan items. Dit betreft het proces dat een item ondergaat van start tot gereed product en verwijdering. De inrichting van het werkproces voor itemontwikkeling ligt dus al grotendeels vast.

Let er bij de keuze voor een systeem op dat meerdere personen tegelijkertijd kunnen werken in het systeem en gelijktijdig aan dezelfde items kunnen werken. In webgebaseerde toetssystemen is dit meestal standaard mogelijk. Voor stand-alone programma's is dit meestal niet mogelijk.

Via een systeem van rollen en rechten kunnen items sequentieel aan verschillende personen ter beschikking worden gesteld voor revisie, controle of accordering. Dit is de geautomatiseerde workflow. In een workflow kan worden vastgelegd dat een item pas naar een volgende fase kan als een voorgaande fase, bijvoorbeeld grammaticale controle, is doorlopen.

In een toetssysteem kan van elk item een geschiedenis worden bijgehouden. Zo wordt inzichtelijk wie wanneer aan welk item heeft gewerkt. Als er fouten zijn ontstaan tijdens het proces, kan zo nodig een roll-back worden uitgevoerd.

Tip
Als de kwaliteit bij inzet al hoog is: richt een uitgebreider workflowproces in met diverse controlemomenten. Daarmee optimaliseer je de mogelijkheden om de kwaliteit van de items te verhogen. In § 16.2.2 lees je meer over gecontroleerde itemontwikkeling.

15. DE ORDENING VAN EEN ITEMBANK

Een goede ordening van een itembank geeft inzicht in de opbouw ervan en in de mate waarin deze geschikt is om toetsen mee samen te stellen. Ook helpt een goede ordening bij het efficiënt beheren van items. Tegelijkertijd wijst de praktijk uit dat ordening nooit een geheel sluitend systeem kan opleveren. Voor één specifieke toets is het wel mogelijk om een sluitend systeem te creëren. Gebruik je één itembank om items voor toetsen voor meerdere vakgebieden te selecteren, dan is het een vrijwel onmogelijke opgave.

Wat is de ordening van een itembank? Het vastleggen van wat iedereen bedoelt met de term is van groot belang om spraakverwarring te voorkomen. Maak goed onderscheid tussen de *structuur* van een itembank en *metadata*. In dit hoofdstuk bespreken we deze begrippen en indelingsprincipes.

15.1 Metadata

Metadata zijn toegevoegde stukjes informatie aan items. Metadata van items zijn bijvoorbeeld het hoofdonderwerp van de vraag, een subonderwerp van de vraag, de status van het item (concept, ter revisie, ter redactiecontrole, goedgekeurd, afgekeurd), de moeilijkheidsgraad of het discriminerend vermogen, de soort vraag, de taxonomische aanduiding (bijvoorbeeld kennis, toepassing, inzicht), de functie van de vraag, wie het item heeft gemaakt of bewerkt en in welke toetsen het item wordt gebruikt.

Figuur 3. Een item met enkele metadatavelden ingevuld

15.2 Structuur

De structuur van een itembank is de opdeling van de itembank in onafhankelijke, vaak hiërarchisch geordende eenheden. Dit is vergelijkbaar met een mappenstructuur zoals je die vindt in de verkenner van een computer.

Over het algemeen hebben gebruikers behoefte aan deze duidelijke hiërarchische structuur. Om inzicht te krijgen in de opbouw van een itembank, is deze structuur beter bruikbaar dan een structuur alleen gebaseerd op metadata. In de praktijk werkt het het beste om eerst de hiërarchische structuur te kiezen en dan op vraagniveau metadata toe te voegen.

De structuur en metadata zijn vaak op verschillende wijzen technisch met elkaar verweven per itembanksysteem. Gebruikers kunnen de structuur en de metadata apart gebruiken voor indeling en selectie van items, maar soms ook in combinatie. Dit maakt het doordenken van structuur en metadata extra moeilijk.

Enkele voorbeelden van bestaande itembankstructuren:

Figuur 4. Het VUmc heeft een aparte itembank binnen het itembanksysteem van Questionmark Perception. De hoofdstructuur van het VUmc volgt de afzonderlijke cursusafhankelijke toetsen (CAT) die worden afgenomen in het curriculum.

Figuur 5. De hoofdstructuur voor een itembank over statistiek van de Faculteit der Sociale Wetenschappen van de VU. Deze itembank bevindt zich in een cursus in Blackboard. De itembank volgt de onderwerpindeling binnen statistiek met verbijzondering naar Advanced/Basic en Moeilijkheidsgraad.

Figuur 6. (links) DeKennistoetsenbank bevat itembanken voor gebruikers van Pedagogisch Werk (PW), Maatschappelijke Zorg (MZ) en Verpleegkundige en Verzorgende opleidingen MBO Zorg en Welzijn. Deze itembanken zijn ingedeeld naar: vraagtypen, de structuur van de Body of Knowledge (BoK), het Kwalificatie Dossier (KD) en bijzondere kenmerken. In de afbeelding is de inrichting van de itembank van Pedagogisch Werk (PW) weergegeven.

Figuur 7. (boven) Een docent van de cursus Inleiding Microbiologie heeft een itembank in QuestionmarkLive gemaakt. De indeling volgt exact de hoofdstukindeling van het studieboek dat hij gebruikt.

Figuur 8. Indeling vragenbank LUMC. In het itembanksysteem RemindoToets heeft het Leids Universitair Medisch Centrum (LUMC) een itembank.

Figuur 8.1 (links) In eerste instantie koos het LUMC voor een themagerichte indeling op basis van disciplines uit de landelijke Voortgangstoets. Deze structuur leidde tot problemen vanwege de gebrekkige vindbaarheid van items voor docenten en het beheer van de vragen door docenten.

Figuur 8.2 (rechts) Inmiddels is voor een meer pragmatische ordening naar opleidingen en onderwijsonderdelen gekozen.

Figuur 9. Structuur van de itembank van Toets & Leer. Iedere itembank binnen het itembanksysteem RemindoToets van Toets & Leer is uit drie niveaus (categorieën) opgebouwd: *Onderwerp > Leerdoel > Toetsterm*. Hierdoor klap je slechts drie niveaus uit om een toetsvraag te zoeken, namelijk onderwerp, leerdoel en toetsterm. Het is de kunst geweest om de lijst met onderwerpen niet te algemeen, maar ook niet te gedetailleerd te maken.

15.3 Structuur en metadata gebruiken

Hoe moet je de structuur en metadata van een specifieke itembank inrichten? In de eerste plaats helpt de inrichting het *doel* van de toetsen – die worden samengesteld op basis van de items – zo efficiënt mogelijk te realiseren. Hiervoor is met name de *toetsmatrijs* leidend. Is die niet aanwezig, neem dan als uitgangspunt een sequentie van items waarmee studenten kunnen oefenen. De structuur en de metadata helpen om de items te selecteren, vindbaar te maken en 'in te laden' in een toets. Ze maken het eenvoudig om de grote hoeveelheid items zodanig te ordenen dat de gebruiker er goed mee kan werken en er bijvoorbeeld een toets met de gewenste eigenschappen mee kan samenstellen.

Tip

Voor metadata geldt het adagium: garbage in = garbage out.
Als er te weinig tijd of geld is om metadata consequent en precies bij te houden, dan wordt de bruikbaarheid van de metadata teniet gedaan. Houd liever minder metadata bij en zorg dat ze van goede kwaliteit zijn.

Een basisprincipe voor de inrichting van de structuur en de hoeveelheid metadata is het streven naar spaarzaamheid. Houd een complexiteit aan die net genoeg is om het doel te bereiken dat de toetsen dienen en de basisstappen voor kwaliteitscontrole te borgen.

Ten tweede ondersteunt de inrichting van de metadata het proces van item-ontwikkeling en kwaliteitsborging. De metadata in de vorm van de processtappen voor ontwikkeling zijn dus van belang. Zie ook § 14.1. Hetzelfde geldt voor metadata waarmee je eventueel verouderde items kunt traceren of items afkomstig van een bepaalde auteur kunt selecteren.

Op de derde plaats moet je de metadata zo inrichten dat ze de basis kunnen vormen voor zinvolle rapportages en feedback voor studenten. Is het doel van een specifieke toets dat studenten feedback krijgen over hoe zij per cel van een toetsmatrijs scoren? Schep hiervoor dan de mogelijkheid in de metadatastructuur. Is dit doel niet vooraf bepaald? Laat een dergelijke indeling achterwege.

Tip

Als in je vakgebied sprake is van een breed gedragen Body of Knowledge (BoK, kennisbasis) of (inter)nationale standaarden, zoals bijvoorbeeld CanMeds of bij de Interuniversitaire Voortgangstoets Geneeskunde (iVGT), dan ligt het voor de hand om deze indeling als hoofdstructuur te kiezen. Zeker als die structuur ook al hiërarchisch is geordend. Het gebruik ervan kan nog effectiever zijn als de structuur van het curriculum van een opleiding er ook nauw bij aansluit.

Als het curriculum (vakindeling) er niet nauw bij aansluit, wat helaas vaak het geval is, levert het kiezen van de BoK als hoofdstructuur in de praktijk problemen op. Vaak is het dan efficiënter om als hoofdstructuur de curriculumindeling (vakken) te volgen en daar BoK-gegevens als metadata aan te koppelen.

Figuur 10. Een itembank met als hoofdstructuur een indeling van hoofdstukken. Vragen kunnen worden gefilterd op basis van verschillende categorieën van metadata.

Tot slot, accepteer dat er verschillende itembanken met verschillende ordeningen en metadatastructuren binnen één itembanksysteem worden gehanteerd. Binnen hogeronderwijsinstellingen leeft vaak de wens om één uniforme ordening en metadatastuctuur te realiseren voor alle items. De praktijk laat zien dat dit niet kan. Dat is niet zo vreemd: itembanken worden in het hoger onderwijs gebruikt voor verschillende doeleinden en vakgebieden en vereisen dus een andere ordening en metadatastructuur. Probeer wel om de hoeveelheid varianten te beperken.

Tip

Het 'op papier' verzinnen van een benodigde structuur om een bepaald toetsdoel te realiseren, is heel moeilijk. Het beste kun je deze structuur bedenken door een pilot uit te voeren. Ontwikkel een matrijs en een toets met behulp van het itembanksysteem. Itereer daarna naar een optimale oplossing voor de gegeven context en het doel. Zorg ervoor dat de gebruikers van het systeem zich bewust zijn van de voor- en nadelen van de gekozen indeling, zodat er voldoende draagvlak is. Hier zit een 'politiek' aspect aan, omdat deelgebieden in een vakgebied graag hun deelgebied in de juiste 'zwaarte' vertegenwoordigd zien in een toets.

Zeker als je toetsresultaten over een langere termijn met elkaar wilt vergelijken (zoals bijvoorbeeld bij de voortgangstoetsen in het medisch domein), is het belangrijk om de toetsmatrijs niet te vaak te veranderen.

Tip
Vaak willen docenten bij een item kunnen aangeven met welke andere items in de itembank het betreffende item niet gelijktijdig in een toets mag voorkomen. Het aantal mogelijkheden dat hiervoor wordt nagelopen, neemt snel toe als de itembank groeit. Dit vergt een voortdurende handmatige handeling. Het is eenvoudiger om te onderzoeken of de toetsen zo kunnen worden samengesteld en gefinetuned dat, na de eerste selectie van items, eventueel van elkaar afhankelijke items handmatig kunnen worden verwijderd.

Tip
Het is verstandig om elk item in een itembank van een unieke code te voorzien, die zelf de hiërarchische of metadatastructuur in zich herbergt. Items worden namelijk soms per ongeluk verplaatst. Na een toetsafname waarbij alleen ruwe data overblijft, maken de codes het terugvinden van individuele items in de bank weer mogelijk. Het unieke technische ID van een item op systeemniveau zegt vaak niet veel.

Een voorbeeld van zo'n code: <leerdoel><topic><subtopic><niveau><uniek nr>

B2_AF12-3_RegcoefT is dan een item uit het Bachelor2-jaar over Hoofdstuk 12 uit het studieboek van Agresti en Finlay, met als onderwerp de regressiecoëfficiënt.

Dit vergt wel uiterst precies werken door degene die items invoeren en controleren. Het onderstreept het belang van vindbaarheid en het gebruik van te veel of te weinig tags.

15.4 Feedback en vraagtypes

Afhankelijk van het doel van je itembank maak je een keuze over het type items en de aard van de feedback van de items.

15.4.1 Feedback

Voor formatieve toetsen is het ontwikkelen van feedback per item heel belangrijk. Feedback kan de student belangrijke studeeraanwijzingen geven. Het is verstandig om vroeg in het project na te gaan welke soort feedback je wilt ontwikkelen. Zo kun je een goede inschatting maken van de benodigde ontwikkeltijd per vraag.

Vooral bij *oefentoetsen* is het belangrijk dat er bij iedere vraag inhoudelijke feedback wordt gegeven. Feedback moet voldoende specifiek en gedetailleerd zijn en zo mogelijk studeeraanwijzingen voor de student bevatten, of een verwijzing naar studiemateriaal om de stof opnieuw te bestuderen. Je kunt niet vooraf bedenken waarom studenten correcte dan wel incorrecte antwoorden kiezen. Noem die dan ook niet. Formuleer feedback zorgvuldig en bij voorkeur neutraal. Het mag studenten niet demotiveren, maar uitnodigen om verder te studeren.

15.4.2 Vraagtypes

In een itembank kun je items maken op basis van verschillende vraagtypes. De meeste toets- en itembanksystemen ondersteunen meer dan tien types.[10] Zie ook § 11.2.1.

Voor *oefentoetsen* is het gebruik van verschillende vraagtypes aan te raden. Dit maakt de toetsen aantrekkelijker. Met deze vraagtypes kun je bijvoorbeeld meerdere deelonderwerpen tegelijkertijd bevragen of gebruikmaken van de voordelen van ICT-middelen, zoals het verslepen of aanklikken van objecten. Dat elke vraag kan resulteren in een andere score, maakt weinig uit voor oefentoetsen.

Voor *summatieve* en *diagnostische* toetsen is vanuit statistisch oogpunt het gebruik van verschillende vraagtypes af te raden. Kies voor multiplechoicevragen én geef deze items allemaal een gelijk gewicht (bijv. 0 of 1 scorepunt). Deze keuze zorgt ervoor dat er geen problemen optreden met verschillende scoremogelijkheden of met de wijze van beantwoording. Dit zorgt voor meer flexibiliteit *in het random kunnen trekken* van items uit een itembank. Variëren in het aantal alternatieven (3, 4 of 5 keuzeopties in een multiplechoicevraag) levert over het algemeen geen problemen op. De scores vindt immers plaats op basis van 0 of 1 punt.

Spring voorzichtig om met items waarbij een casus wordt beschreven met daarbij deelvragen. Een dergelijk vraagtype maakt het moeilijker om flexibel items te trekken, omdat het aantal deelvragen kan verschillen per casus. Dat heeft invloed op de maximaal te behalen score.

[10] *Zie bijvoorbeeld dit overzicht van alle QTI-gedefinieerde vraagtypes: https://webapps.ph.ed.ac.uk/qtiworks/anonymous/samples#cat1*

16. DIDACTIEK: KWALITEIT VAN ITEMS

Het verhogen van de kwaliteit van toetsing is meestal één van de belangrijkste redenen om gezamenlijk aan itembanken te werken. Zie § 10.2. De kwaliteit van toetsing neemt toe als de kwaliteit van de items toeneemt. Maar hoe zorgt een tool als een gezamenlijke itembank voor een hogere itemkwaliteit? Daarover gaat dit hoofdstuk.

Er zijn veel definities voor de kwaliteit van toetsing.[11] De kwaliteit van een itembanksysteem betreft de mate waarin het systeem het mogelijk maakt om op een effectieve, efficiënte en prettige manier items te ontwikkelen, te beheren en te selecteren voor opname in een toets. De kwaliteit van de itembank neemt toe naar mate de items beter de leerdoelen uit de toetsmatrijs representeren. De kwaliteit neemt ook toe als er meer items in de itembank zitten die van hogere kwaliteit zijn.

16.1 Kwaliteit van items

Wil je de kwaliteit van items verbeteren, focus je dan op twee soorten basiskwaliteit van toetsing: validiteit en betrouwbaarheid. Met behulp van itembanken kun je in beide een verbeterslag maken. Dit geldt zowel voor formatieve toetsing als voor summatieve toetsing.

De *validiteit* van een toets gaat over de mate waarin een toets meet wat je beoogt te meten. Voor validiteit is het belangrijk dat de items relevant zijn voor een bepaald thema of leerdoel. Kwaliteitsverhoging treedt op wanneer docenten vaardiger worden in het maken van dit soort items. Training en samenwerking helpt daarbij. Itembank-systemen ondersteunen dit proces onder andere door mogelijkheden te bieden om direct commentaar te leveren, het faciliteren van samenwerking en de opzet van een verbetercyclus via een workflow. Zie § 14.1 en 14.2.

De *betrouwbaarheid* gaat over de mate waarin het toetsresultaat (de score op een toets) reproduceerbaar is en zo min mogelijk afhangt van het toeval. Voor een hoge betrouwbaarheid moet de 'ruis' of 'meetfout' zo minimaal mogelijk zijn. De *betrouw-baarheid* van toetsen kan op de eerste plaats groter worden door *meer* items op te nemen in toetsen. Door gezamenlijk aan itembanken te werken, vergroot je de voorraad aan items. Het aantal items waarmee een docent toetsen kan samenstellen, neemt toe.

De betrouwbaarheid wordt ook groter door items te construeren die beter *discrimineren* tussen de studenten die de stof in meer of mindere mate beheersen. Beter discriminerende items verhogen over het algemeen ook de validiteit: je meet beter wat je wilt meten. Formuleer items in de eerste plaats zo helder, efficiënt en objectief mogelijk om beter te discrimeren.[12]

[11] *Zie Sluijsmans, Joosten- ten Brinke en Van Schilt-Mol, 2014*

[12] *Er zijn veel handboeken en lijstjes met vuistregels om de meest bekende oorzaken voor onduidelijke items te voorkomen en daarmee de ruis. Zie bijvoorbeeld Van Berkel, H. J. M., Bax, A., & Joosten-ten Brinke, D. (2017). Toetsen in het Hoger Onderwijs. (4rd ed.). Houten/Diegem: Bohn Stafleu Van Loghum.*

16.2 Hoe realiseer je een kwaliteitsverbetering?

Je verhoogt de kwaliteit van items door fluctuatie in kwaliteit van items te verminderen en de vraagkwaliteit op een hoger niveau te krijgen. Dit inzicht komt voort uit het vakgebied van de operations management, zoals geïllustreerd in Figuur 11. In het linkerplaatje zie je dat de kwaliteit van een product of dienst gemiddeld vrij laag is, dan weer eens hoog en dan weer eens laag, maar niet stijgend. Het rechterplaatje laat zien dat de kwaliteit gestaag toeneemt en de fluctuatie kleiner wordt. Om een dergelijke kwaliteitsverbetering te realiseren, is een doelgerichte benadering nodig.

Figuur 11. Grafische weergave van kwaliteitsontwikkeling. Links: de kwaliteit varieert heel erg, de gemiddelde kwaliteit is relatief laag. Rechts: de kwaliteit wordt constanter en de kwaliteit neemt systematisch toe.

Er zijn twee manieren om de kwaliteit van nieuwe items te verhogen: ofwel door een grotere vaardigheid te ontwikkelen in het construeren van items, ofwel door review tijdens het constructieproces. Itembanksystemen bieden hiervoor workflow-ondersteuning. Zo werk je systematisch aan kwaliteitsverhoging.

Om de kwaliteit van bestaande items in de itembank te verbeteren, kun je verder analyses uitvoeren na afname van toetsen en aanpassingen doorvoeren op basis van interpretaties van die analyses. Dit is in feite een cyclisch proces, dat onderdeel uitmaakt van het onderhoud en beheer van een itembank. We werken dit in de volgende paragrafen uit.

16.2.1 Voorlichting en training

Om goede items te kunnen construeren, zijn voorlichting, training en een goede feedbackloop in itemontwikkeling onontbeerlijk. Geef bij de start van een itembank training aan de beoogde itemontwikkelaars. Dit helpt om een gezamenlijk gevoel van kwaliteit te ontwikkelen en te laten groeien. Betrek toetsexperts bij de opzet van de training. Wanneer de itemontwikkeling eenmaal op gang is, kunnen meerdere trainingen volgen, bijvoorbeeld tijdens redactievergaderingen van het auteursteam. Zie ook § 16.2.3. Het diepgaand bespreken van items op basis van toets- en itemanalyses is in feite impliciete training. Training is daarmee een continu terugkerend onderdeel van het onderhoud en beheer van de itembank.

16.2.2 Gecontroleerde itemontwikkeling door review

Door een stapsgewijze itemontwikkeling met ingebouwde reviewmomenten kun je veelvoorkomende valkuilen of onduidelijkheden voortijdig signaleren en oplossen. Het welbekende vierogenprincipe is hiervoor bedoeld. Review vindt plaats door het inschakelen van vakgenoten of experts.

Voor twee samenwerkende docenten hoef je het reviewproces en de verdeling van rollen niet uitgebreid te definiëren, zeker niet in technische zin. Tegenwoordig

wordt wel geëist dat elk item in ieder geval éénmaal door een collega-docent is bekeken. Ook moet duidelijk zijn wie een item goedkeurt. Onderschat dit niet: de consequenties van een toets die door één docent is ontwikkeld, zijn even groot als die van een landelijke toets. Een student die de toets niet heeft gehaald, kan niet afstuderen.

Itembanken voor toetsen waarmee grote belangen gemoeid zijn, vereisen meer rollen en rechten. Stel dat op basis van de itembank toetsen worden ontwikkeld voor een selectie, of toetsen zonder herkansingsmogelijkheid. De stakeholders, zoals de politiek, het publiek, bestuurders en studentenbonden, stellen hoge eisen aan transparantie en de kwaliteit van toetsing. Denk bijvoorbeeld aan landelijke toetsen, zoals de PABO-rekentoets, de toets in het kader van 10voordeleraar of de centrale toetsing in het mbo. Er is een uitgebreid, transparant reviewproces nodig. Voor een beschrijving van mogelijke rollen in dit proces, zie § 14.2. Hoe omvangrijker het team van betrokkenen, des te belangrijker is de aansturing van teams en project-ondersteuning door een coördinator.

16.2.3 Itembankonderhoud en beheer

Onderdeel van het onderhoud en beheer van de itembank is het bespreken van items die niet zo goed functioneren. Uit psychometrische analyse van deze items blijkt dat ze lage Rit-waarden of extreme p-waarden hebben. Voor een uitgebreide uitleg van deze termen, zie § 16.3.2. Probeer er gezamenlijk achter te komen wat de oorzaak is en wat je kunt doen om de items te verbeteren. Regelmatig zul je besluiten om items te verwijderen en geheel opnieuw te construeren. Deze besprekingen zijn wellicht aanleiding om de regels en richtlijnen voor items met het team aan te scherpen. Soms blijkt dat het probleem niet in de items, maar in het onderwijs zit. In dat geval zul je moeten bekijken hoe onderwijs en items beter op elkaar kunnen aansluiten.

Het komt voor dat er een herbeoordeling van items nodig is, of een aanpassing van de ordeningen. Er zijn bijvoorbeeld nieuwe inzichten in het vakgebied of wijzigingen in onderwerpen of leerdoelen, of in de metadata en de toetsmatrijs. Je kunt je dan afvragen of er nieuwe sets van items voor nieuwe onderwerpen nodig zijn. Moeten items worden gehercodeerd? Zet gerichte ontwikkelopdrachten uit voor de itemont-wikkelaars.

Redactieteam

Het team van itemontwikkelaars komt gedurende het ontwikkelproces regelmatig bij elkaar om de voortgang van de itemontwikkeling in de gaten te houden. Maak één persoon verantwoordelijk voor het itemontwikkelproces. Noem deze persoon de coördinator of hoofdredacteur. Kom met elkaar overeen dat deze persoon het mandaat heeft om mensen aan te spreken op de afspraken over het ontwikkelen van items.

De hoofdredacteur plant regelmatig een overleg in met alle betrokkenen. Werk in deze overleggen aan het onderling vertrouwen en het bespreken van problemen rondom het proces van de itemontwikkeling.

16.3 Toetsgegevens

Een krachtig middel om de itemkwaliteit te verbeteren, is het verzamelen van gegevens over items van afgenomen toetsen. In combinatie met de kwaliteitsverhoging door itembanken heb je een dubbele verbeterslag te pakken. Deze gegevens zeggen namelijk iets over de kwaliteit van de items en de mate waarin studenten presteren, of over de kwaliteit van het onderwijs. Het is belangrijk om beschrijvende en psychometrische gegevens goed te onderscheiden. Zij hebben een heel verschillend karakter, dat moet passen bij het doel dat je nastreeft met de itembank.

16.3.1 Beschrijvende data

Eenvoudige beschrijvende data van items gaan om zaken zoals het aantal keren dat een item is afgenomen of beantwoord.

Op het niveau van een toets als geheel kunnen beschrijvende data gaan over hoe vaak een student een poging heeft ondernomen om een toets te maken, wanneer de poging is geweest, hoe de student scoort, wat de slagingspercentages zijn enzovoort. Deze data zijn afkomstig uit toetsservicesystemen.

De data kun je vertalen naar een grafische overzichtspagina voor docenten of studenten. Docenten krijgen een overzicht van de progressie van een groep studenten en de onderwerpen waarmee ze moeite hebben. Studenten krijgen inzicht in hun progressie en hun positie ten opzichte van groepsgenoten. Dit levert in feite managementinformatie op. Gebruik van data op deze wijze noemt men ook wel *learning analytics*.

16.3.2 Psychometrische data

Psychometrische gegevens zeggen iets over de moeilijkheidsgraad van items en in welke mate ze onderscheid maken of een student de stof wel of niet beheerst. Daarbij maakt het uit of je gebruikmaakt van de klassieke testtheorie (KTT) of de item response theorie (IRT). In deze paragraaf leggen we beide theorieën kort uit.

De klassieke testtheorie (KTT)

De klassieke testtheorie gaat ervan uit dat de score op een toets bestaat uit *de werkelijke score en een fout in de meting (ruis).* Door statistische bewerkingen kan van één toets worden bepaald wat de betrouwbaarheid is van de toets, dat wil zeggen, in welke bandbreedte de werkelijke score van de student met een bepaalde mate van zekerheid ligt.

De belangrijkste psychometrische gegevens voor items binnen het concept van de klassieke testtheorie zijn:

- *Betrouwbaarheid:* de mate waarin de toets als geheel, dus alle items gezamenlijk, goed onderscheid maken in de mate van beheersingsgraad van de stof. Het is een maat van meetnauwkeurigheid en de mate waarin de score op een toets niet aan het toeval is te wijten. Volgens de literatuur moeten tentamens minimaal een waarde van 0,7 of hoger hebben en bijvoorbeeld selectietoetsen een waarde van 0,9 of hoger.
- *p-waarde:* het percentage correcte antwoorden van de studentenpopulatie. De p-waarde wordt vaak aangeduid als de moeilijkheidsgraad van een item, maar is in feite een graad van gemakkelijkheid. Een optimale p-waarde voor open items is 0,5 en voor multiplechoicevragen met 4 alternatieven is 0,67. Hier kan ook nog de a-waarde worden genoemd. Dat is het percentage studenten dat een bepaalde afleider heeft gekozen.
- *Rit-waarde:* de correlatie tussen de score van de studenten op de item en de score op de toets als geheel. De Rit-waarde is een maat voor het discriminerend

vermogen van een vraag. Dat is de mate waarin het item onderscheid maakt tussen de studenten die de stof beter of slechter beheersen. De Rit-waarde is een wat strengere maat voor het discriminerend vermogen, omdat de invloed van het item zelf op het discriminerend vermogen wordt uitgesloten. Volgens de literatuur ligt het streven op een Rit-waarde van minimaal 0,3. Waarden lager dan 0,1 worden als slecht beschouwd. Negatieve waarden verdienen direct aandacht. Hier komt de Rat-waarde van pas. Dat is de mate waarin de keuze voor een bepaald incorrect alternatief samenhangt met de toetsscore.

Het kan waardevol zijn om psychometrische gegevens bij de items op te slaan. Sommige toetssystemen ondersteunen zelfs een automatische update van deze gegevens na iedere beantwoording door een student. Op zich aantrekkelijk, ware het niet dat ze alleen interpreteerbaar zijn per toetsafname. Onderling kun je ze alleen vergelijken als de afnamecondities identiek zijn. Een itembank moet daarom de psychometrische gegevens per toetsafname opslaan.

De item response theorie (IRT)

Wil je weten wat de moeilijkheidsgraad en het discriminerend vermogen van items zijn, dan biedt de item response theorie (IRT) uitkomst. IRT levert zogenaamde gekalibreerde items op. Om die te verkrijgen, worden de items afgenomen bij een grote steekproef onder een groep studenten met veel spreiding in kennis en vaardigheid. In het kalibratieproces wordt heel precies gemeten hoe items onderscheid maken tussen studenten op specifieke niveaus van moeilijkheid. Itembanken met items die op deze manier zijn ontwikkeld, kun je gebruiken in zogenaamde adaptieve toetssystemen. De achterliggende technieken zijn zo complex dat ze alleen met ruime ontwikkelbudgetten kunnen worden gemaakt. Voorbeelden hiervan zijn de PABO-rekentoets (Wiscat), de Rekentuin (primair onderwijs) en de in ontwikkeling zijnde computer-adaptieve versie van de iVTG (interuniversitaire voortgangstoets geneeskunde). Ook geldt dat de items aan strengere eisen moeten voldoen dan bij klassieke toetsen, vooral voor wat betreft discriminerend vermogen.[13]

Soms kiezen docenten in de voorbereiding voor adaptief toetsen, zonder zich voldoende te realiseren dat hiervoor complexe en kostbare IRT-technieken nodig zijn. Kies hier alleen voor als de projectgelden toereikend zijn.

Voor de ordening van de itembank heeft de keuze voor IRT niet zo heel veel consequenties. Het maken van een adaptieve itembank vergt meer kennis en middelen om de kalibratie uit te kunnen voeren, maar het resultaat kan worden opgenomen in een ‘gewone’ itembank. Er zijn voldoende items nodig, verspreid over de verschillende moeilijkheidsniveaus, met meer moeilijke items dan items met een gemiddelde moeilijkheidsgraad.[14]

[13] *Zie Van der Linden, W. J., & Glas, C. A. W. (2000). Computerized adaptive testing: theory and practice. Springer.*

[14] *Voor een wat uitgebreidere handzame uitleg van de klassieke testtheorie en item respons theorie, zie De Gruijter, D. N. M. (2008). Toetsing en toetsanalyse via http://media.leidenuniv.nl/legacy/toetsing-en-toetsanalyse.pdf.*

17. JURIDISCHE AANDACHTSPUNTEN

Als je als instelling een itembank wilt ontwikkelen, hetzij alleen, hetzij samen met andere instellingen, krijg je te maken met een aantal juridische aandachtspunten. Hoe borg je het eigendom van de items en de itembank? Wat moet je regelen op het gebied van bescherming van persoonsgegevens? En met betrekking tot samenwerkingsvormen? Hoe ga je het verspreiden van items tegen? In dit hoofdstuk komen deze zaken aan de orde.

Weet je niet zeker hoe het bij jouw instelling in elkaar steekt? Vraag het na bij de afdeling juridische zaken of het juridisch bureau van je instelling.

Werk je aan een instellingsoverstijgende itembank? In de praktijk is het een behoorlijke klus dit juridisch goed doortimmerd te krijgen. Betrek een jurist met kennis van zaken. Neem de tijd om te zoeken naar iemand met deze expertise.

17.1 Eigendom items en itembank

Maak afspraken over de rechten over de items. Hiermee borg je dat deze beschikbaar blijft voor de instelling en/of het samenwerkingsverband. Items zijn beschermd door het auteursrecht. Dit recht ontstaat bij het maken van het werk. Volgens de Auteurswet heeft de auteur het exclusieve recht om te publiceren en te dupliceren.

Let op de volgende zaken:

- Voor iedere docent met een arbeidsovereenkomst geldt dat de auteursrechten toekomen aan de werkgever indien het materiaal is gemaakt bij de uitoefening van de onderwijstaken, tenzij partijen anders zijn overeengekomen.
- Indien er sprake is van een samenwerkingsverband: maak de items eigendom van het samenwerkingsverband. Specificeer dit in de samenwerkingsovereenkomst òf stel de items onder Creative Commons Licentie beschikbaar. Let op: indien de items summatief worden ingezet, is dit laatste wellicht niet wenselijk, want dan wil je juist niet dat iedereen de items kan gebruiken.
- Regel bij een samenwerking dat bij vertrek van één van de partners het eigenaarschap van de items bij het samenwerkingsverband blijft. Leg dit vast in de samenwerkingsovereenkomst.
- Als je gebruikmaakt van afbeeldingen bij items, heb je te maken met beeldbankenrecht. Het is verboden beeldmateriaal te kopiëren en te gebruiken. Dit moet de instelling of het samenwerkingsverband afkopen per beeld, of er een abonnement voor afsluiten. Tip: definieer een extra stap binnen je itemontwikkelproces, waarin de reviewer van een item checkt of het gebruikte beeld en geluid is toegestaan.
- Bepaal of er databankenrecht ontstaat. De Databankenwet beschermt tegen overnemen of herhaaldelijk opvragen van gegevens uit een databank zonder toestemming van de producent. De gegevens die in de databank staan, zijn als 'verzameling' beschermd door het databankenrecht. Als je als partij een investering doet in de databank bij het aanleggen, ontstaat er automatisch databankenrecht. Degene die investeert, mag ook exploiteren.
- Spreek met de betrokken commerciële marktpartij af hoe je content kunt terugkrijgen bij een exit-scenario. Wat voor tool levert de marktpartij? Draait de applicatie op eigen servers of is het een SaaS (software-as-a-service)-dienst? Indien de applicatie is geïnstalleerd op eigen servers is het terugkrijgen geen

probleem, maar bij een SAAS-oplossing wordt het lastiger. Maak hierover afspraken bij de start van de dienstverlening.
- Laat je goed adviseren als je een commerciële partij betrekt, zodat jouw belangen zijn afgedekt. Voorkom dat een andere partij ervandoor gaat met de waardevolle inhoud van de itembank.

17.2 Misbruik van itembankmateriaal

Instellingen willen meestal voorkomen dat de items 'op straat' terecht komen, vooral als het summatieve items betreft. Vaak komen toetsen en items terecht op www.studeersnel.nl, www.stuvia.nl en www.knoowy.nl. Spreek de website aan op het auteursrecht. Zij zijn geen eigenaar van de items en mogen ze dus niet publiceren. Als je ze vraagt om de items te verwijderen, moeten zij hieraan gehoor geven.

17.3 Het gebruik van materiaal van andere partijen in itembanken

Je kan niet zomaar allerlei video's en afbeeldingen van anderen overnemen in items. Hieraan zijn wettelijke regels verbonden. Wel zijn er veel mogelijkheden om rechtmatig bestaand materiaal te gebruiken en anders is er de mogelijkheid om sommige afbeeldingen of video's speciaal voor je itembank opnieuw of na te maken. Neem hiervoor budget op in je project.

Ga na of je instelling, of de instelling waarmee je wilt samenwerken, een regeling heeft getroffen met Stichting Pro (Publicatie en Reproductie Organisatie). Dit orgaan monitort auteursrechten binnen het onderwijs en beschermt de bij haar aangesloten uitgeverijen. Pro controleert instellingen op het gebruik en hergebruik van publicaties van commerciële partijen. Instellingen kunnen ervoor kiezen om Stichting Pro jaarlijks een afkoopbedrag te betalen. Daardoor hoeven deze instellingen geen administratie bij te houden voor korte stukken tekst die worden overgenomen.

Als je instelling géén afkoopbedrag betaalt, moet je expliciet rekening houden met auteursrecht. Soms moet je voor een overgenomen bron betalen, maar vaak ook niet. Hieronder sommen we de belangrijkste vuistregels op voor het gebruik van beeld en geluid in items. Deze informatie is ontleend aan www.auteursrechten.nl.

17.3.1 Linken mag altijd

Je mag linken naar afbeeldingen, beeld- en geluidsfragmenten op het internet. Openbaarmaking op het internet moet echter wel rechtmatig zijn. Een beeld- of geluidsfragment stream je rechtstreeks. Downloaden en vertonen is verboden, tenzij de licentie dit toestaat.

Voor summatief toetsen kan streamen problematisch zijn. Bij toetsing wordt vaak gebruikgemaakt van materiaal dat juist niet openbaar toegankelijk mag zijn. Bovendien wil je niet dat het materiaal nèt van het internet is verdwenen op het moment van de toets.

17.3.2 Licentie voor hergebruik

Afbeeldingen en audio(visuele) werken met een licentie die hergebruik toestaat, bijvoorbeeld een Creative Commons-licentie, kun je zonder problemen downloaden en vertonen. Afhankelijk van de gebruikte CC-licentie zijn de mogelijkheden ruimer of beperkter. Check daarom altijd even de betekenis van de licentie op https://creativecommons.nl/uitleg/.

17.3.3 Het citaatrecht biedt veel mogelijkheden

Het citaatrecht maakt het mogelijk om afbeeldingen of audio(visuele) werken te gebruiken. Je moet aan de volgende voorwaarden voldoen:

- De citaten dienen ter ondersteuning van de inhoud van het onderwijs. Je mag ze niet gebruiken ter versiering. Ook mag je geen wijzigingen aanbrengen in het te citeren fragment.
- De omvang van het citaat is gerelateerd aan het doel dat je ermee nastreeft. In de praktijk komt dit neer op: gebruik altijd korte fragmenten.
- Afbeeldingen mag je wel in hun geheel 'citeren'.
- Je doet aan bronvermelding.

17.3.4 Afbeelding of audiovisueel werk uit de eigen bibliotheek

Mogelijk heeft je instelling gebruikslicenties afgesloten, bijvoorbeeld voor Academia[15], die gebruik van afbeeldingen en audio(visuele) werken binnen het onderwijs van je instelling toestaan, zonder dat je voor ieder gebruik aparte toestemming moet vragen en een vergoeding betalen. Ga na binnen je instelling of er dergelijke licenties zijn.

17.3.5 Complete audio(visuele) werken

Je mag een compleet audio(visueel) werk gratis afspelen of vertonen zonder toestemming, mits:

- het afspelen een educatief doel tijdens een les of werkgroep dient en deel uitmaakt van het onderwijsprogramma.
- het afspelen fysiek in de onderwijsinstelling plaatsvindt.
- het afspelen plaatsvindt in het kader van onderwijs zonder winstoogmerk.
- Een kopie van het werk in een toetsapplicatie zetten, zodat studenten het thuis kunnen afspelen, mag niet. Je moet hiervoor toestemming vragen aan de maker.

17.3.6 Delen van audio(visuele) werken

Je mag delen van audio(visuele) werken plaatsen in een toets, mits:

- de vertoning/het afspelen uitsluitend dient ter toelichting bij niet-commercieel onderwijs. Het is dus aanvullend en niet onderwijsvervangend.
- de plaatsing gebeurt in een besloten omgeving, waartoe alleen studenten toegang hebben. Loggen studenten in voor de toets? Dan voldoe je aan deze voorwaarde.
- er een billijke vergoeding wordt betaald aan de rechthebbenden. Neem hiervoor contact met ze op.

De vertoning hoeft niet fysiek in het leslokaal plaats te vinden. Studenten mogen de werken thuis zien of afspelen.

17.3.7 Bronvermelding

Als je gebruikmaakt van materiaal van derden waarvoor je moet betalen, neem je contact op met de rechthebbenden (of het werk laten namaken). Het contact verloopt via de volgende organisaties:

- VIDEMA: https://www.videma.nl
- BUMA/STEMRA: http://www.bumastemra.nl/
- PICTORIGHT: http://www.pictoright.nl/auteursrecht/

Meer informatie:

- https://www.auteursrechten.nl/
- http://www.onderwijsenauteursrecht.nl/
- https://IUSmentis.com
- https://creativecommons.nl/

[15] *https://www.academia.nl/*

17.4 Bescherming persoonsgegevens

De Wet bescherming persoonsgegevens (WBP) beschermt de privacy van gebruikers van de itembank (docenten, medewerkers en studenten van de instelling). In mei 2018 treedt de Algemene Verordening Gegevensbescherming (AVG) in werking. Vanaf die datum heb je bewerkersovereenkomsten nodig om hun privacy te borgen.

Zodra medewerkers of studenten een account aanmaken voor bijvoorbeeld een toetsapplicatie en de instelling of leverancier toegang heeft tot deze persoonsgegevens, moet er een bewerkersovereenkomst worden getekend tussen de instelling en de leverancier.

Let op de volgende zaken:

- De instelling en of het samenwerkingsverband moet ervoor zorgen dat deze vereisten vanaf mei 2018 zijn gerealiseerd. Je hoeft je hier als individuele docent, of medewerker geen zorgen over te maken.
- Het gebruik van een SAAS (software-as-a-service)-dienst vraagt om extra handelingen. De instelling of het samenwerkingsverband moet zorgen voor een bewerkersovereenkomst met de leverancier. Bij het inloggen in de applicatie staat er dan vaak: 'er zijn gebruiksvoorwaarden van toepassing'. De gebruikers kunnen dit wel of niet accepteren, waarmee de privacy is geborgd. De instelling of het samenwerkingsverband neemt zelf voldoende passende maatregelen. Er is afstemming met de leverancier nodig in een contract en medewerkers moeten weten waar ze aan toe zijn. Als het gaat om een applicatie op de eigen servers, dan is de instelling of het samenwerkingsverband zelf verantwoordelijk voor het systeem en de gegevens. Zorg ook in dat geval voor duidelijke communicatie met medewerkers.
- Lees meer over de Algemene Verordening Gegevensbescherming en de overgang ernaartoe op de SURF thema-pagina *Algemene Verordening Gegevensbescherming (AVG)*. [16]

17.5 Rechtsvormen voor samenwerkingsverbanden

Ben je van plan om met meerdere instellingen een itembank te ontwikkelen? Leg de rechtsvorm en de onderlinge verhoudingen vast. Stem ook af wie wat op welk moment aanlevert. Dit leg je bijvoorbeeld vast in statuten en of een samenwerkingsovereenkomst.

Raadpleeg voor meer informatie de SURF whitepaper *Samenwerkingsvormen en licentiemodellen voor itembanken* uit 2013.[17]

[16] *https://www.surf.nl/themas/beveiliging/beleidsondersteuning-privacy/algemene-verondening-gegevensbescherming-avg/index.html*

[17] *https://www.surf.nl/kennisbank/2013/whitepaper-samenwerkingsvormen-en-licentiemodellen-voor-itembanken.html*

BIJLAGE 1: BEGRIPPENKADER

Onderstaande begrippenkader is gebaseerd op het Begrippenkader voor digitaal toetsen van SURF uit 2013[18], aangevuld met begrippen die relevant zijn bij het ontwikkelen van een itembank.

Adaptieve toets	Een toets waarbij de kandidaten gemakkelijker of moeilijker opgaven moeten beantwoorden afhankelijk van de vaardigheid van de kandidaat, die wordt ingeschat op basis van voorgaand items. Bij een op IRT gebaseerde adaptieve toets is sprake van gekalibeerde items waarvan die op een gegeven vaardigheidsniveau van de kandidaten het meest nauwkeurig kunnen meten. Bij een computer adaptieve toets gebaseerde toets worden geen items meer aan de kandidaat aangeboden als het niveau van vaardigheid met de gewenste nauwkeurigheid is vastgesteld. Zie ook: item response theorie.
Afspeelomgeving	De component van een toetssysteem waarmee toetsen en items aan de kandidaten worden aangeboden en de antwoorden geregistreerd worden. • zorgt dat studenten na inloggen de toets kunnen maken • wordt vaak gecombineerd met een zogenaamde secure browser, die ervoor zorgt dat de student niet ongeoorloofd software kan gebruiken om op internet te zoeken, met derden te communiceren, etc. • registreert de antwoorden en stuurt deze terug naar de itembank
Auteursomgeving	De component van een toetssysteem waarmee items kunnen worden ingevoerd. • is een interface waarin items kunnen worden ontwikkeld en voorzien van metadata • biedt keuze uit allerlei vraagtypen (meerkeuze, aanwijsvraag, open vraag, etc.) • geeft de mogelijkheid multimediale componenten op te nemen • ondersteunt (meestal) de workflow en legt het reviewproces vast • ondersteunt het reviseren van items (bijvoorbeeld na de eerste toetsafname)
Betrouwbaarheid	De mate waarin de score op een toets niet van het toeval afhankelijk is. De mate waarin een toets bij gelijke condities dezelfde uitslag oplevert. In statische termen wordt betrouwbaarheid beschreven in de mate waarin een meting vrij is van meetfouten (ook wel ruis of error genoemd). Zie ook: Misclassificaties
Body of Knowledge (BoK)	Een Body of Knowledge (BoK) is een gemeenschappelijke kennisbasis voor een vakgebied of sector waaraan de professional zijn theoretische en praktische kennis, inzichten en methoden ontleent. Het gaat niet alleen om theorie, maar ook om beproefde inzichten en methoden van de beroepsgroep in kwestie.

[18] *https://www.surf.nl/kennisbank/2013/begrippenkader-voor-digitaal-toetsen.html*

Businesscase	Een businesscase of een haalbaarheidsstudie beschrijft de afweging om een project of taak te beginnen of te stoppen. In de business-case worden de kosten tegen de baten afgewogen, rekening houdend met de risico's. Een businesscase is niet bedoeld om uitsluitend een financiële afweging te maken. Juist ook kwalitatieve overwegingen worden in een goede businesscase nadrukkelijk in de beoordeling betrokken. Een aspect als kwaliteitsverbetering kan voor een instelling heel belangrijk zijn, terwijl het zuiver financiële gewin daarvan nauwelijks te bepalen is. Bekijk ook de whitepaper De businesscase van digitaal toetsen: https://www.surf.nl/kennisbank/2014/whitepaper-de-businesscase-van-digitaal-toetsen.html
Cronbach's alfa	Een psychometrische maat die bij een niet herhaalde toets (dus éénmalige afname zoals een tentamen) een waarde geeft waarin de meting onderhavig is aan meetfouten. Cronbach's alfa is geschikt voor polytoom gescoorde items. Vragen waar naast correct-incorrect (0 of 1) ook met bijv. 0, 1, 2, 3 etc. punten kan worden gerekend. *Zie ook:* KR20
Diagnostische toetsen (onderdeel van formatieve toetsen)	Een formatieve toets waarbij een kandidaat inzicht krijgt in zijn progressie door de leerstof. Een diagnostische toets is representa-tief voor wat betreft inhoud en niveau met een eventueel volgende summatieve eindtoets. *Zie ook:* Oefentoets.
Digitaal toetssysteem	Het geheel van alle technische componenten en functies om items en toetsen te ontwikkelen, af te nemen, analyseren en op te slaan.
Digitale toetscyclus	Dit is het cyclische proces om van leerdoelen, via items, toetsen, afname en analyse toetsing te realiseren en voortdurend te verbeteren. Vaak wordt in het centrum van deze cyclus de ontwikkeling en beheer van de itembank gesitueerd.

Formatieve feedback	Feedback waarbij de opbrengst voor het leerproces voorop staat. De feedback is bedoeld om studenten te stimuleren om zich verder te verdiepen in de stof. Bij incorrecte antwoorden worden vaak aanwijzingen gegeven om het correct antwoord te kunnen achterhalen of simpelweg het correcte antwoord. Bij correcte antwoorden wordt soms verdiepende stof of een ander voorbeeld besproken.
Formatieve toetsing	Toetsing waarbij het leren van de toets voorop staat. Er wordt vaak afgezien van het geven van cijfers (dit wordt ook aanbevolen in de literatuur). Door met toetsopgaven aan de slag te gaan en te leren van fouten en feedback wordt het leerproces gestimuleerd. Er wordt niet gestreefd naar een bepaald minimaal te behalen niveau. Soms ook aangeduid met Assessment for Learning of Assessment as Learning.
Geparametriseerde vraag	Een item bestaand uit een vaste hoofdstructuur met variabele elementen (getallen, objecten, concepten, principes), zodat ze elke keer een nieuw item vormen. Sommige itembanksystemen genereren elke keer een unieke nieuw item voor de docent, of een unieke set van waarden voor elke student.
IMS	Een IMS standaard is vastgelegd door het IMS Global Learning Consortium, een community van hogeronderwijsinstellingen, leveranciers en overheidsinstellingen die samen uitwisselingsstandaarden ontwikkelen. Toen IMS in 1997 begon, was de officiële naam Instructional Management Systems (IMS) project.
Item response theorie (IRT)	Item response theorie (IRT) is een theorie waarbij ervanuit gegaan wordt dat de kans dat een kandidaat een item correct beantwoord bepaald wordt door de vaardigheid van de kandidaat als functie van het moeilijkheidsniveau van de item en het onderscheidend vermogen van dit item. IRT is een theorie die uitgaat van de karakteristiek van items, terwijl de klassieke testtheorie uitgaat van de karakteristiek van een toets. Om IRT zinvol te kunnen toepassen moet op basis van een groot aantal toetsafnames bij een populatie waarin alle vaardigheidsniveaus aanwezig zijn, deze karakteristieken voorafgaand aan een toets worden bepaald (zgn. kalibreren). *Zie ook:* Adaptieve toets
Item	Een opdracht/taak waarop een kandidaat het correcte antwoord moet geven. Wordt ook wel aangeduid als vraag, toetsvraag of toetsitem.
Itembank	Een verzameling van items voor een bepaald toetsdoel. Een itembank heeft daarvoor een bepaalde ordening, meestal op basis van metadata en in de vorm van een hiërarchische structuur.
Itembanksysteem	Een technische voorziening (software) om itembanken in op te kunnen slaan en bewerken. Itembanksystemen kunnen losse systemen zijn of onderdeel van een toetssysteem. In een toetssysteem zijn ook functies opgenomen om toetsen samen te stellen, toetsen af te nemen, toetsen te scoren en toetsanalyses uit te voeren.

Klassieke testtheorie (KTT)	De klassieke testtheorie (KTT) heeft als basisaanname dat de geobserveerde score op een toets bestaat uit de werkelijke score plus een meetfout. Uitgaande van deze basisveronderstellingen kunnen grootheden zoals Cronbach's alfa worden berekend, die een schatting is voor de grootte van de meetfout.
KR20	Een psychometrische maat die bij een niet herhaalde toets (dus éénmalige afname zoals een tentamen) een waarde geeft waarin de meting onderhavig is aan meetfouten. Het enige verschil tussen KR20 en Cronbach's alfa is dat er bij KR20 uitsluitend met dichotoom gescoorde items kan worden gerekend, zoals multiplechoicevragen. Dus items die correct of incorrect zijn en 0 of 1 punt kunnen opleveren. *Zie ook:* Cronbach's alfa
Learning analytics	Learning analytics is het verzamelen en analyseren van studiedata die een student tijdens het online leren genereert. De studiedata worden omgezet naar waardevolle informatie en kan bijdragen aan de verbetering van de onderwijskwaliteit. *Zie ook:* https://www.surf.nl/themas/onderwijsinnovatie-met-ict/learning-analytics
LTI	LTI (Learning Tools Interoperability standard) is een IMS standaard om toetssystemen te koppelen aan andere systemen zoals digitale leeromgevingen (LMSen zoals Blackboard of Canvas). Daarbij worden docent- en studentgegevens tussen de systemen automatisch uitgewisseld, kunnen toetsen vanuit de LMS worden gestart en worden de scores op toetsen teruggevoerd naar de cijferlijstfuncties in de LMS. *Zie ook:* SCORM
Meetfout	De meetfout bij een toets is de afwijking van de score op een toets die niet verklaard kan worden. De werkelijke score is de geobserveerde score minus de meetfout. *Zie ook:* ruis en error.
Metadata	Toegevoegde gegevens aan items die aanvullende karakeristieken beschrijven. Die gegevens kunnen gebruikt worden om de items te structureren (zie structuur), zoeken, filteren en selecteren.
Misclassificaties	Het aandeel kandidaten dat onterecht gezakt danwel geslaagd is op een toets. Toetsen met lage betrouwbaarheid hebben een hoge mate van misclassificaties. Zie ook: Betrouwbaarheid
Oefentoetsen (onderdeel van formatieve toetsen)	Een toets waarbij een kandidaat kan oefenen met de leerstof. De kandidaat kan zelf onderwerpen en niveau kiezen. Vaak worden deze toetsen gebruikt om vooral met moeilijke begrippen of concepten te oefenen. Soms wordt enige adaptiviteit ingebouwd zodat de afspeelomgeving items op een best passend niveau wordt aangeboden aan de kandidaat. *Zie ook:* diagnostische toets.

Psychometrische data	Psychometrische data betreft te berekenen gegevens van afgenomen toetsen en items die iets zeggen over de kwaliteit van de toetsing. Enerzijds zeggen zij iets over de betrouwbaarheid en moeilijkheidsgraad van toetsen als geheel. Anderzijds zeggen ze iets over het discriminerend vermogen en de moeilijkheidsgraad van afzonderlijke items uit die toetsen. *Zie ook:* Klassieke testtheorie, item response theorie, Cronbach alpha, p-waarde, Rit-waarde
P-waarde	Het percentage kandidaten dat een item correct heeft beantwoord. Wordt ook de moeilijkheidsgraad van een item genoemd. Hoe hoger de p-waarde, des te gemakkelijker is de item. P-waarde wordt alleen gebruik in relatie tot de klassieke testtheorie.
QTI	QTI (Question and Test Interoperability) is een IMS standaard voor de uitwisseling van items en toetsen tussen toetssystemen.
Rir-waarde	De correlatie tussen de score van de kandidaten op de item en de score op de toets als geheel minus de betreffende item zelf. De Rir-waarde is een wat strengere maat voor het discriminerend vermogen dan de Rit-waarde omdat bij de Rir-waarde de invloed van de item zelf op het discriminerend vermogen wordt uitgesloten. Rir-waarde wordt alleen gebruik in relatie tot de klassieke testtheorie. *Zie ook:* Rit-waarde en Klassieke testtheorie
Rit-waarde	De correlatie tussen de score van de kandidaten op de item en de score op de toets als geheel. Het is maat voor het discriminerend vermogen van een vraag ofwel de mate waarin de vraag onderscheid maakt tussen de studenten die de stof beter of slechter beheersen. Rit-waarde wordt alleen gebruik in relatie tot de klassieke testtheorie. *Zie ook:* Rir-waarde en Klassieke testtheorie
SCORM	SCORM (Sharable Content Object Reference Model) is een IMS standaard om toetssystemen te koppelen aan andere systemen zoals digitale leeromgevingen (LMSen zoals Blackboard of Canvas). Daarbij worden met name docent- en studentgegevens tussen de systemen automatisch uitgewisseld, kunnen toetsen vanuit de LMS worden gestart en worden de scores op toetsen teruggevoerd naar de cijferlijstfuncties in de LMS. *Zie ook:* LTI
Stabiel domein	Een vakgebied dat weinig veranderingen meer ondergaat. Zie ook: Volatiel domein
Structuur	De manier waarop een verzameling van items logisch geordend zijn, bijvoorbeeld in een boomstructuur.
Summatieve toetsing	Toetsing waarbij het zo nauwkeurig mogelijk meten van een bepaald niveau van vaardigheid voorop staat. De behaalde score op een dergelijke meting wordt gebruikt voor formele toekenning van een behaald studieresultaat zoals studiepunten of een diploma. Denk bijvoorbeeld aan examens of tentamens. Soms ook aangeduid met Assessment of Learning.
Toetsitem	*Zie:* Item

Toetskwaliteit	Het geheel van aspecten rondom toetsing dat de mate van transparantie, betrouwbaarheid en validiteit van toetsing behelst. Transparantie kan bijvoorbeeld toenemen als inzicht is in de wijze waarop items worden gemaakt, betrouwbaarheid kan toenemen als items beter onderscheid maken tussen studenten met verschillende mate van beheersing van de stof of vaardigheid, validiteit kan toenemen als de items beter en vollediger de beoogde kennis of vaardigheid meten.
Toetsmatrijs	Een toetsmatrijs is een overzicht hoe toetsopgaven in een toets verdeeld zijn over de stof. Vaak wordt per onderdeel van de stof tevens de gewenste type vaardigheid (kennis, inzicht of toepassing) aangegeven. Soms ook aangeduid als toetsmatrix, toetplan, specificatietabel of blueprint.
Toetsveiligheid	De mate waarin correcte antwoorden items en toetsen niet op een ongeoorloofde manier door kandidaten gegeven worden (fraude). Bijvoorbeeld doordat items zijn gestolen of in het publieke domein terecht zijn gekomen, danwel dat kandidaten tijdens de toetsafname de beschikking hebben over ongeoorloofde bronnen, danwel dat kandidaten toetsuitslagen kunnen wijzigen. Zie o.a. de SURF uitgave Handboekveilig toetsen: https://www.surf.nl/kennisbank/2016/werkboek-veilig-toetsen.html
Toetsvraag	*Zie:* item
Vakgebied	Een vakgebied is een terrein van beroepsmatig toegepaste kennis en vaardigheden waar mensen zich in kunnen specialiseren. Voor het uitwisselen van ervaring en vakkennis is meestal een jargon nodig.
Vragenbank	Een systeem waarin alle items zijn verzameld.
Volatiel domein	Een vakgebied dat aan veel veranderingen onderhevig is door nieuw ontwikkelde kennis, methoden of technieken. *Zie ook:* Stabiel domein
Workflow	Een logische, vastgelegde volgorde van activiteiten die moeten worden uitgevoerd om een van te voren gedefinieerde uitkomst te verkrijgen. Deze stappen kunnen sequentieel zijn, maar kunnen ook parallel worden uitgevoerd afhankelijk van doel van het proces.

BIJLAGE 2: DRIE WORKFLOWS VOOR ITEMONTWIKKELING

In deze bijlage beschrijven we drie praktijkvoorbeelden van workflows voor itemontwikkeling.

1. Een itemontwikkelproces door twee docenten

Een coördinator/examinator van de faculteit Aard- en Levenswetenschappen van de Vrije Universiteit ontwikkelt en onderhoudt een itembank over het onderwerp Microbiologie. Elk jaar worden ongeveer 60 nieuwe items gemaakt.

- De coördinator is de itembank gestart op basis van de itembank van de uitgeverij McGrawHill: Prescott's Microbiology. De structuur van de itembank volgt de hoofdstukken uit dat boek. Die itembank wordt door de uitgeverij geleverd in Word formaat. De docent kopieert de inhoud één voor één naar het online itembanksysteem QMLive.
- Tijdens het invoeren voert de coördinator een kwaliteitscheck uit: ongeveer 80-90 procent van de items zijn relevant en van goede kwaliteit. De rest van de items verwijdert de coördinator.
- Een nieuw item maakt de docent eerst in Word.
- Een week voor uitvoering van het cursusonderdeel laat de coördinator nieuwe items reviewen door de uitvoerend docent. Reviewen gaat via Word.
- Op basis van de feedback van de uitvoerend docent past de coördinator het item aan. Hij voert het in in het itembanksysteem.
- Na afloop van de toets en op basis van de psychometrische analyse past de docent een item direct aan in de itembank voor de volgende jaren.

Het kost de docent ongeveer 0,5 uur om een goed item exclusief feedback te ontwikkelen. De coördinator maakt vooral andere vraagtypes dan multiple-choicevragen, vaak met afbeeldingen. Dat kost relatief veel tijd per item.

2. Itemontwikkelproces itembank Toets & Leer

Toets & Leer is een samenwerkingsverband van zes hogescholen met itembanken voor de onderwerpen Bedrijfsadministratie/boekhouden, Bedrijfseconomie, Belastingrecht, Marketing, Management en Recht.

De hogescholen leveren docenten voor de itemontwikkeling. Ze werken met gesloten beurzen. Het streven is om 300 items op jaarbasis per hogeschool te ontwikkelen.

- De docent ontwikkelt een item en voert dit in in het itembanksysteem. De docent informeert een docent-collega van één van de andere hogescholen, meestal per groepje van items.
- De collega-docent controleert de items en zorgt voor eventuele wijzigingen en aanvullingen. Aanvullingen en opmerkingen worden in de toetsapplicatie opgeslagen. De collega-docent bericht de oorspronkelijke auteur en deze controleert de wijzigingen.
- Het item wordt toetstechnisch gecontroleerd door een extern bureau. Eventueel ontvangt de oorspronkelijke auteur een opmerking of een verzoek tot wijziging. Dan volgt hernieuwde controle.
- Het item in de toetsapplicatie wordt definitief goedgekeurd.
- In het itembankproject vindt alleen op incidentele basis terugkoppeling plaats op basis van een psychometrische analyse.

Er bestaat de wens om een redacteur in te voegen in het proces. Het kost gemiddeld ongeveer 0,5 uur om een goed item inclusief feedback te ontwikkelen. De meeste items zijn multiplechoicevragen (ongeveer 80 procent) en invulvragen (korte tekst of getal, ongeveer 20 procent).[19]

[19] *Een handleiding voor de ontwikkeling van items staat hier: https://docs.google.com/document/d/1UeWvNhpo1_CFZJjnDhuJzsplyAtzKf1pFVlh8gYOc3Y/edit?usp=sharing*

3. Itemontwikkelproces MBO kennistoetsenbank

De betrokken mbo's leveren docenten voor de itemontwikkeling. Itemontwikkelaars worden betaald voor 4 uur itemontwikkeling per week. Per opleiding zijn ongeveer 6-10 itemontwikkelaars actief.

- De itemontwikkelaars construeren in een team pre-concept-items met feedback, geordend naar zowel het kwalificatie-dossier als de Body of Knowledge (BoK). Deze pre-concept-items gaan meerdere keren langs meerdere docenten (interne validatie). Labels aan items maken duidelijk welke items nog moeten worden bekeken (metatags). Aanvullingen en commentaar worden bijgehouden in het itembanksysteem.
- Er vinden maandelijkse face-to-face werksessies met toetsdeskundigen en docenten van de deelnemende mbo's plaats. Tijdens die sessies bespreken de teams nieuwe items per kwalificatiedossier. Hieruit ontstaan concept-items.
- De itemontwikkelaar past de concept-items zo nodig aan. Hieruit ontstaan gepubliceerde items. Vóór publicatie vindt taalredactie plaats.
- Vertegenwoordigers uit het werkveld en onderwijs controleren de gepubliceerde items (externe validatie). Ze zien in het itembanksysteem welk conceptitems wachten op een beoordeling. De vertegenwoordigers voegen aanvullingen en commentaar toe in de toetsapplicatie.
- De feedback van de externe validator gaat terug naar de docententeams en wordt zo nodig besproken tijdens de werksessies. Aansluitend worden items definitief als goedgekeurde items aan de itembank toegevoegd.
- Het aanpassen is een continu proces. Gebruikers kunnen opmerkingen plaatsen bij items die tot aanpassingen kunnen leiden. Bij grote aanpassingen is het mogelijk om opnieuw interne en externe validatie toe te passen.
- In het itembankproject vindt alleen op incidentele basis terugkoppeling plaats op basis van een psychometrische analyse.

Het kost gemiddeld ongeveer 1 uur om een goed item inclusief feedback te ontwikkelen.

COLOFON

Author
Silvester Draaijer *Vrije Universiteit*
Jenny de Werk *SURFnet*

Advies
Wil de Groot-Bolluijt *SIG digitaal toetsen*

Redactie
Marjolein van Trigt

Met de inhoudelijke inbreng van	*Meer informatie*
Sander Schenk Hogeschool Rotterdam	www.toetsenleer.nl
Jane Groenendijk Prove2Move	www.dekennistoetsenbank.nl
Jeroen Donkers Maastrichtuniversity	www.ivtg.nl
Marjolein Wijnker DAS	
Peter Roos 10voordeleraar	www.10voordeleraar.nl
Angela Verschoor Cito	

Met dank aan
Desiree Joosten-Ten Brinke *Fontys Hogescholen*
Ludo van Meeuwen *Technische Universiteit Eindhoven* en *SIG digitaal toetsen*
Rob Reichardt *NHL Stenden Hogeschool* en *SIG digitaal toetsen*
Kirsten Veelo *SURFnet*

Vormgeving
Vrije Stijl, Utrecht

Photography
Flickr.com

februari 2018

Copyright

Beschikbaar onder de licentie Creative Commons Naamsvermelding 4.0 Nederland.
https://creativecommons.org/licenses/by/4.0/deed.nl

SURF

+31 (0)88 787 30 00
www.surf.nl