

De kracht van digitaal toetsen

Stap voor stap naar
meer kwaliteit en gemak

Brigitte Hellings

Colofon

De kracht van digitaal toetsen

Stap voor stap naar meer kwaliteit en gemak

Copyright © 2023 - Optimum Assessment

Geactualiseerde uitgave

De kracht van Digitaal Toetsen

Stap voor stap naar meer kwaliteit en gemak

Copyright © 2019 - Optimum Assessment

Auteur: Brigitte Hellings

Initiële uitgave

Digitaal Toetsen? Help!

Copyright © 2014 - Andriessen en Partners

Auteur: Manon Bonefaas

Redactie

Natasja Kroon (Optimum Assessment)

Jantina Dorreman (Optimum Assessment)

Vormgeving

OOGT - Merkontwerpbureau

Drukwerk

Q-Promotions

De kracht van digitaal toetsen is een uitgave van:

Optimum Assessment

Almystraat 10A

5061 PA Oisterwijk

013-5286363

info@optimumassessment.com

www.optimumassessment.com

Voorwoord

De aandacht voor persoonlijke groei en professionele ontwikkeling neemt toe. Toetsing wordt dan ook steeds belangrijker om de opgedane kennis en vaardigheden te duiden, erkennen en waarderen. Wij geloven dat intuïtieve toetssoftware een krachtig instrument is om groei te stimuleren, op een mooie, flexibele en efficiënte wijze.

In de afgelopen 20 jaar hebben wij al veel organisaties mogen ondersteunen en adviseren op het gebied van (digitaal-) toetsen en examineren. Als partner in digitaal toetsen is Optimum Assessment continu bezig met het vernieuwen en doorontwikkelen van haar dienstverlening en producten. Dat doen we in nauwe samenspraak met onze klanten die elk hun eigen kleuring willen geven aan de toetsing. De ontwikkeling van ons toetssoftware platform heeft inmiddels een vlucht genomen en ook in de toekomst willen we ruim aandacht geven aan nieuwe technische (maatwerk) oplossingen om nog beter aan de diversiteit aan wensen van onze klanten tegemoet te komen.

Met deze publicatie delen we onze visie op en enthousiasme voor digitale toetsing. 'De kracht van digitaal toetsen' is een actualisering van de oorspronkelijke uitgave met de titel 'digitale toetsing. Help!' dat in 2014 is uitgebracht. Deze vernieuwde versie is praktisch ingestoken en bevat onder andere handreikingen voor de toetslogistiek bij digitale toetsing en een checklist om de stap van schriftelijk naar digitaal te zetten. Tevens is aanvullend ingezoomd op kwaliteitscriteria rondom toetsing, de criteria die gebruikt kunnen worden in de zoektocht naar passende toetssoftware en de laatste innovatieve ontwikkelingen rondom digitale toetsing.

Ik ben blij dat ik dit boekje kan aanbieden zodat wij onze kennis kunnen delen. Veel leesplezier!

Martin van den Berge
Oisterwijk, januari 2023

Inhoud

1 Inleiding	7
2 Waarom digitaal toetsen?	8
2.1 Wat is digitale toetsing?	10
2.2 Aanleiding digitaal toetsen	10
2.3 Vooroordelen	11
2.4 Voordelen van digitale toetsing	17
2.5 Formatief versus summatief toetsen	19
3 Van schriftelijk naar digitaal	24
3.1 Toetslogistiek en digitale toetsing	26
3.2 Stappen naar digitale toetsing	36
3.3 Digitaal toetsen gefaseerd invoeren	41
4 Kwaliteit van examenvragen	46
4.1 Kwaliteitscriteria examenvragen	48
4.2 Analyse van examenvragen	51
4.3 Vraagsoorten	54
4.4 Scoring van verschillende vragen	57
5 De keuze van toetssoftware	60
6 Innovatie in digitale toetsing	68
7 Literatuur	74

Bijlage: Checklist - stappen naar digitaal toetsen

1 Inleiding

Dit boekje is bedoeld voor iedereen die overweegt om met digitaal toetsen te starten of om de digitale toetsing een 'schwung' te geven. Het tweede hoofdstuk gaat in op het begrip digitale toetsing en belicht diverse redenen waarom je digitaal zou toetsen. Het derde hoofdstuk beschrijft allerlei keuzes die je kunt maken, als je van schriftelijke naar digitale toetsing gaat. Laat het je inspireren om digitale toetsing vorm te geven, op een manier die bij jou en je organisatie past. Hoofdstuk 4 geeft aan hoe je, in het licht van digitale toetsing de kwaliteit aandacht kunt geven. In hoofdstuk 5 vind je verschillende zaken die van belang zijn bij de keuze van een digitaal toetsstelsel. Wellicht helpt het je om zicht te krijgen op allerlei mogelijke vragen die je aan leveranciers van toetssoftware systemen kunt stellen. Tot slot besteden we aandacht aan innovatie in toetsing.

Na het lezen van dit boekje heb je een duidelijk beeld van wat digitale toetsing behelst, hoe je het praktisch toepast en wat het in de praktijk kan opleveren. Je zult merken dat digitaal toetsen de kwaliteit van toetsing in hoge mate bevordert en dat het tegelijkertijd het proces noemenswaardig versimpelt. Zeker op termijn bemerk je de voordelen qua efficiëntie. Laat dit boekje je inspireren en motiveren om aan de slag te gaan met digitale toetsing.

Waarom digitaal toetsen?

- In dit hoofdstuk zoomen we in op wat digitaal toetsen precies inhoudt. We weerleggen een aantal vooroordelen, geven concreet inzicht in de praktische voordelen en besluiten het hoofdstuk met de meerwaarde van de inzet van digitaal toetsen bij verschillende toetsvormen.

toetsen

2.1 Wat is digitale toetsing?

Digitale toetsing is het afnemen van een examen via een beeldscherm (computer of tablet), met behulp van toetssoftware. De examenvragen staan in een database opgeslagen op één locatie, die via een beveiligde verbinding toegankelijk is. Digitaal toetsen is een iets bredere term dan online toetsen, omdat je ook digitaal kunt examineren zonder internetverbinding. Digitale toetsing is ook net wat breder dan digitaal examineren, omdat examinering verwijst naar een formele vorm van toetsing, waaraan een certificaat of diploma vast zit. Bij toetsing hoeft dat niet het geval te zijn, toetsing kan natuurlijk ook verwijzen naar tussentijdse toetsing. Voor meer informatie over het onderscheid tussen toetsing en examinering verwijzen we naar paragraaf 2.5.

2.2 Aanleiding digitaal toetsen

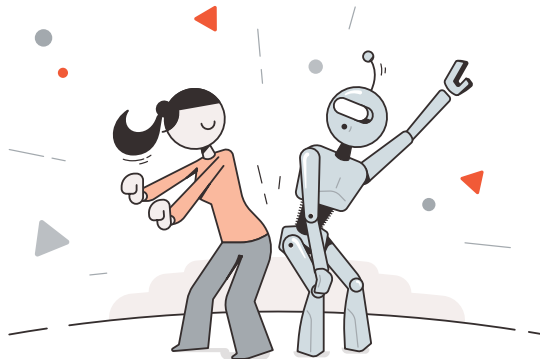
Wat maakt dat (examen)organisaties en opleidingsinstituten willen overstappen van examens 'op papier' naar digitaal toetsen? De aanleiding is vaak de wens tot vernieuwing van het examen, qua inhoud, wijze van afname en flexibiliteit voor wat betreft tijdstip en plaats. Waarom dan niet 'digitaal gaan'? Een andere vaak gehoorde aanleiding is de wens tot vereenvoudiging van de correctie. Met digitale toetsing is verzending ten behoeve van de correctie van examenwerken bijvoorbeeld niet meer nodig en je kunt met gesloten vragen de kosten voor de correctie zelfs volledig wegstrepen. Tot slot vormt e-learning vaak een aanleiding tot digitaal toetsen: leren verloopt digitaal en tussentijdse toetsen vormen daar onderdeel van.

De aanleiding tot digitale toetsing kan divers zijn, maar zo ook de twijfel. Bij de zoektocht naar wel of niet digitaal toetsen en

met welk systeem, komen er gaandeweg meestal twijfels in de vorm van vooroordelen naar voren. Laten we deze aannames benoemen en ontkrachten.

2.3 Vooroordelen

Er bestaan verschillende 'vooroordelen' over digitaal toetsen. Deze aannames kunnen zorgen voor (onterechte) twijfel over het positieve effect of de toegevoegde waarde van digitale toetsing. Misschien herken je ze?



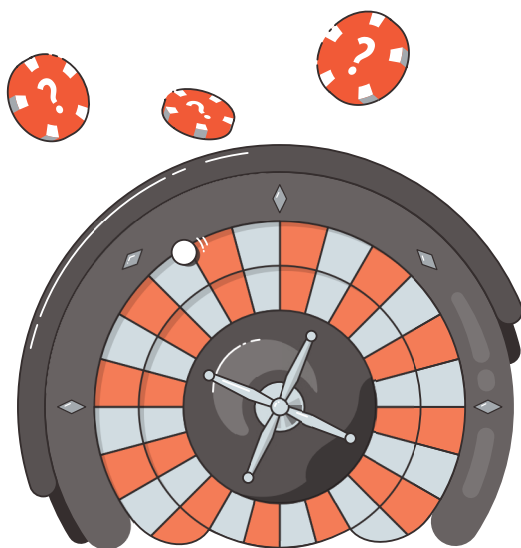
- **Alles wordt door de computer nagekeken, waardoor de menselijke maat ontbreekt;**

Deze aanname komt voort uit de gedachte dat een digitale toets bestaat uit alleen meerkeuzevragen. Er bestaan hoe dan ook veel soorten gesloten vragen, die we in hoofdstuk 3 beschrijven en toelichten. Naast gesloten vragen zijn er echter ook systemen met de mogelijkheid om open vragen in het examen op te nemen.

Open vragen zijn vragen waarbij de kandidaat zelf een antwoord moet formuleren. Deze vragen kunnen 'gewoon' nagekeken worden door een beoordelaar of docent, indien gewenst zelfs door meerdere.

- **Met meerkeuzevragen gaat iedereen gokken. Het examen wordt dan veel te makkelijk;**

Het is zelfs zo dat open vragen vaak gemakkelijker zijn dan gesloten vragen, uiteraard uitgaande van kwalitatief goede vragen. Dat je kunt 'gokken', helpt dus lang niet altijd voor het behalen van een goede score. Hoe komt dat? Bij gesloten vragen is er een gokkans, dat wil zeggen de kans dat een kandidaat een goed antwoord geeft, terwijl hij het antwoord niet weet. Deze gokkans wordt echter vaak verdisconteerd in de cesuur,



de grens voldoende - onvoldoende. Daarnaast vergen gesloten vragen het vermogen om zowel de vraag als de antwoordalternatieven goed te interpreteren. Tot slot vereisen gesloten vragen, mits goed gesteld, een degelijk en soms wat meer abstract begrip van de leerstof.

De investering van het opstarten heb je er al snel uit.

Vergelijk de gokkans ook eens met de 'blufkans' van open vragen: ook al weet je het antwoord niet precies, toch kun je een bijna juist antwoord formuleren. Een beoordelaar kent dan punten toe aan het geformuleerde antwoord, ook al is dat niet precies gelijk aan het beoogde antwoord.

Een voorbeeld van de moeilijkheidsgraad van gesloten vragen is een exameninstituut dat overstapt van digitale toetsing middels meerkeuzevragen en open vragen naar enkel meerkeuzevragen. Wat bleek? De kandidaten hadden dusdanig veel moeite met de meerkeuzevragen, dat het slagingspercentage halveerde. Vervolgens is besloten weer op de oude wijze te toetsen, middels een combinatie van meerkeuze en open vragen, waardoor slagingspercentages weer stegen. Onderzoek (Jordan & Mitchell, 2009) beschreven op het weblog van Sally Jordan (2009) bevestigt dit eveneens).

- o **Bij digitale toetsing moet je allerlei vraagvormen gebruiken;**

Als je dan eenmaal voor digitale toetsing gaat, dan zou je kunnen denken dat kandidaten een 'leuk' en innovatief examen willen, inclusief illustraties en filmpjes. Belangrijk is echter een goed examen, waarbij de vraagsoort past bij de inhoud van de examenvraag.

- o **Digitale toetsing is duur;**

Meestal doelen potentiële gebruikers van een online toetsstelsel op de investering die nodig is bij de start van digitale toetsing. Het is inderdaad niet goedkoop om een database op te zetten met toetsen en examenvragen. Maar diezelfde investering die je maakt voor een database, ben je ook kwijt na een aantal keer schriftelijk toetsen. (Zie ook de aanname hieronder over de omvang van de database.) Alleen bij digitale examens blijft de inhoud bewaard (en geheim) en is deze geschikt voor hergebruik. De investering van het opstarten heb je er dan al snel uit. Bovendien kun je ook gefaseerd een database met toetsen en examenvragen opbouwen. Hierover vind je meer informatie in hoofdstuk 2.

- o **Digitale toetsing stelt hoge eisen aan beschikbare faciliteiten;**

Het is een misvatting dat je een groot computerlokaal nodig hebt om digitaal te kunnen toetsen. Het ideale van digitale toetsing is dat dit niet gebonden is aan tijd of plaats. Een plek met een internetverbinding en (bij summatieve toetsing) toezicht, in wat voor vorm dan ook

(één op één, op een toetslocatie of met een supervisor op afstand, met een cameraverbinding en telefoon), is feitelijk voldoende. Dat maakt het examen enorm flexibel voor kandidaten en heb je in de basis weinig faciliteiten nodig. Zelfs een internetverbinding is niet persé noodzakelijk, wanneer je kiest voor offline afname of een hybride vorm van papieren afname.

- **Je hebt training nodig voordat je aan digitale toetsing kunt beginnen;**

Natuurlijk helpt het om een training te volgen, dan haal je immers het maximale uit de toetssoftware. Maar... niet iedereen hoeft alles te weten. Onderscheid verschillende rollen voordat je gaat trainen, bijvoorbeeld de rol van supergebruiker, invoerder van vragen of toezichthouder bij de afname. Dat scheelt enorm in de benodigde trainingstijd bij de implementatie van digitale toetsing.

- **Hoe groter de vragenbank, hoe beter;**

Je vragenbank moet natuurlijk wel representatief zijn voor de stof die getoetst wordt en er moeten genoeg 'versies' in zitten bij de start van digitale toetsing. Maar houd de vragenbank, binnen deze richtlijnen, zo klein mogelijk. Voor het gebruik van statistische gegevens, bijvoorbeeld hoeveel procent van de kandidaten de vraag goed heeft, moet een vraag namelijk een groot aantal keer zijn gesteld. Als er veel vragen in de vragenbank zitten, kan dat lang duren. Daarnaast is onderhoud van een grote vragenbank bewerkelijk en dus ook duur.



- **In geval van storing ligt de examenafname plat;**

Het houdt de gemoederen wel eens bezig: wat als er storingen in de internetverbinding zijn, kunnen we dan geen examens meer afnemen? De juiste toetssoftware slaat echter tussentijds alle antwoorden op én maakt tevens offline afname mogelijk. Het is overigens ook altijd mogelijk om een schriftelijk examen te genereren. In geval van een internetstoring natuurlijk, maar ook op locaties waar überhaupt geen internetverbinding is, bijvoorbeeld in gevangenissen. Een digitaal examen kan dus ook schriftelijk worden afgenomen.

- **Bij digitale toetsing is geen inzage mogelijk;**

Na afname van het examen kan het examenwerk van de kandidaat die zijn examen wil inzien uit de database worden gehaald, inclusief de gegeven antwoorden, de

juiste antwoorden en eventuele feedback naar aanleiding van de correctie. Het gemaakte examen is hoe dan ook altijd reproduceerbaar.

- **Dyslectisch of slechtziend? Digitale toetsen zijn toch geen maatwerk.**

Net zoals bij schriftelijke toetsing is bij digitale toetsing instelbaar per kandidaat hoe lang een examen mag duren. Voor bijvoorbeeld dyslectische kandidaten kun je extra tijd instellen. Daarnaast is het mogelijk om voor slechtzienden een groter lettertype in te stellen en voor blinden en slechtzienden kan de software het examen verklanken.

Wellicht hebben we eventuele twijfels ten aanzien van digitaal toetsen kunnen wegnemen. Wat we nog niet hebben benoemd, maar wat wel een groot (zo niet het grootste) voordeel is van digitale toetsing is dat de kwaliteit van de examens toeneemt, dan wel gemakkelijker te realiseren is. In hoofdstuk 3 komt de kwaliteit van de examens nader aan de orde. De volgende paragraaf zet dit en alle andere voordelen van digitaal toetsen op een rij.

| 2.4 Voordelen van digitale toetsing

Hieronder geven we inzicht in een aantal voordelen van digitale toetsing, kijkend vanuit het perspectief van verschillende gebruikersgroepen: kandidaten enerzijds en opleidingsinstituten, bedrijfsleven en exameninstellingen anderzijds.

Vanuit kandidaten:

- Flexibele afname, qua tijd en locatie;
- Geen zorgen over handschrift;
- Uitslag sneller bekend - de correctie kan vaak in minder tijd en bij gesloten vragen kan de uitslag al direct na afloop van het examen bekend zijn;
- Afname kan onder bepaalde voorwaarden vanuit huis plaatsvinden, met toezicht op afstand.

Vanuit opleidingsinstituten, bedrijfsleven en exameninstellingen:

- Overzichtelijke opslag van examenvragen, op één plaats, binnen de structuur van een database;
- Hergebruik examenvragen;
- Zicht op de kwaliteit van vragen, door snelle en gebruiksvriendelijke analyse van de kwaliteit van vragen;
- Variatie in vraagvormen mogelijk (niet verplicht);
- Mogelijkheid open vragen;
- Vereenvoudigde correctie;
- Minder kosten voor het ontwikkelen van examenvragen;
- Continue verbetermogelijkheid van vragen, waardoor betere kwaliteit van vragen;
- Examens worden gelijkwaardiger, doordat de toetsmatrijs leidend is;
- Toegankelijkheid en overzicht van uitslagen;
- Borging van wijze en inhoud van toetsing;
- Zicht op kwaliteit van opleidingen;
- Zicht op kwaliteit van vakmensen, zoals docenten, correctoren en toetsconstructeurs;
- Afname van het examen kan maatwerk zijn;
- Toezicht op afname kan onder bepaalde voorwaarden op afstand plaatsvinden;

- Mogelijkheid om bij toetsing tijdens het leerproces direct en eenvoudig feedback te geven.

2.5 Formatief versus summatief toetsen

Digitale toetsing kent zijn meerwaarde bij tussentijdse, niet-officiële toetsing op vorderingen die de student maakt. In dit geval spreken we van formatieve toetsing. Formatieve of diagnostische toetsing is gericht op toetsen en leren en heeft als doel beter leren. Deze vorm van toetsing geeft de opleider inzicht in datgene waar in de opleiding nog aandacht aan besteed moet worden.

Formatieve toetsen worden vaak gebruikt om te (be)oordelen, maar het oordeel heeft geen consequentie. De volgende toetsomschrijvingen worden vaak gebruikt bij formatieve toetsing:

- instaptoetsen;
- diagnostische toetsen;
- voorkennistoetsen;
- voorbeeldexamens;
- tussentijdse toetsen;
- zelftoetsen.

Formatieve toetsing heeft als doel:

- monitoren van de voortgang van studenten;
- zelftoetsing door studenten;
- de student helpen leren;

- informatie verzamelen als input voor de inrichting van het leerproces.

Summatieve toetsen zijn bedoeld als (eind)oordeel. Summatieve toetsen hebben:

- ... consequenties: je mag iets wel of niet. Een summatieve toets is vaak een selectieve toets. Denk aan het halen van een rijbewijs. Nadat je bent geslaagd voor zowel het theorie- als het praktijkexamen mag je een auto besturen;
- ... toezicht. Op deze manier kun je aantonen dat de examenafname betrouwbaar verloopt en kun je fraude zoveel mogelijk voorkomen;
- ... een formele status. Je kunt er een diploma of certificaat mee halen of niet (als je zakt);
- ... (vaak) landelijke bekendheid. Het kan zijn doordat de opleiding gesubsidieerd is, maar ook omdat opleidingsinstituten examens afnemen volgens gevestigde (beroeps- en/of branche) standaarden en examenprogramma's. Fijn voor de student, dat geeft meerwaarde aan het diploma!;
- ... strikte geheimhouding. Een examen waarvan je de inhoud van tevoren kent, heeft geen nut. Wat meet je hiermee?

Digitaal toetsen in de praktijk bij Stichting Vakbekwaamheid Horeca (SVH)

Al sinds 2002 maakt Stichting Vakbekwaamheid Horeca (SVH) gebruik van digitaal toetsen, naast schriftelijke toetsing. Sinds 2014 wordt de techniek van digitaal toetsen ook op de schriftelijke toetsen toegepast. De inzet van de techniek van digitaal toetsen bij schriftelijke examens houdt onder andere in dat voor het samenstellen van schriftelijke toetsen uit dezelfde database wordt geput als bij digitale toetsing. Daarnaast worden de scores en resultaten van de schriftelijke examens na de afname digitaal verwerkt en geanalyseerd.

Voordelen van de inzet van digitale toetsing bij schriftelijke examens

Er zijn drie grote voordelen die de kwaliteit van toetsing en de service richting kandidaten bevorderen:

1. (Kwaliteit.) Op deze manier heb je steeds unieke toetsen, zonder dat je de toetsen handmatig hoeft samen te stellen. Het systeem stelt namelijk automatisch de toetsen samen, volgens de toetsmatrijs. Doordat iedere toets uniek is, neemt de kans op fraude af én wordt de kans dat vragen bekend raken veel kleiner. Geen student heeft namelijk dezelfde toets.

2. (Kwaliteit.) SVH maakt in de digitale examens en het schriftelijke equivalent daarvan gebruik van meerkeuzevragen. De antwoorden op de schriftelijke examens worden na afname ingelezen. Op deze manier kan er nog van een ander voordeel van digitale toetsing gebruik worden gemaakt: Middels de toetssoftware vindt analyse van de vragen plaats waardoor er een continue kwaliteitsbewaking is.
3. (Service.) Het is niet nodig om vooraf te weten wie er komt, wat flexibiliteit geeft aan kandidaten. Zelfs als niet ingeschreven kandidaten toch examen willen doen, kunnen ze deelnemen, omdat er altijd extra schriftelijke exemplaren aanwezig zijn. Pas na de examenafname wordt de naam van de kandidaat in het systeem gekoppeld aan het resultaat van de kandidaat.

Combinatie van zowel schriftelijke als digitale afname

Omdat examens bij SVH zowel schriftelijk als digitaal worden afgenomen, is er grote flexibiliteit qua toetsmoment en toetslocatie. Voor digitale toetsing zijn faciliteiten nodig, zoals laptops en een stabiele online verbinding, faciliteiten die niet nodig zijn voor schriftelijke afname. De inzet van schriftelijke naast digitale afname maakt een nog grotere spreiding van locaties mogelijk, waarmee een nog betere service kan worden gegeven aan kandidaten en opleiders.

Gelijkwaardigheid examens en kwaliteit van toetslocaties

Doordat de examens worden gegenereerd volgens een toetsmatrijs, is de inhoud gewogen en zijn er feitelijk geen verschillen tussen de unieke examens (zowel schriftelijk als digitaal). Ook het gebruik van verschillende toetslocaties ten behoeve van digitale toetsing verloopt volgens kwaliteitsnormen. Voor de afname van schriftelijke examens is er eveneens controle op de kwaliteit, doordat eigen freelancers, zoals examenleiders en toezichhouders, worden ingezet. Voor beide examenvormen wordt er daarbij nog toezicht gehouden door een externe partij, die beoordeelt of de afgesproken procedures worden gevolgd.

Voordelen en één aandachtspunt van digitale afname

Voor kandidaten is examen doen per computer prettig: het scheelt schrijfkrimp bij open vragen, bevordert de leesbaarheid voor de corrector en een kandidaat kan het antwoord ook gemakkelijk weer wijzigen tijdens de examenafname. Tot slot worden de antwoorden centraal opgeslagen, zodat je altijd bij de stukken kunt. Bij de verwerking van antwoorden worden ook menselijke fouten automatisch voorkomen. Bij digitale toetsing is er wel een stevig aandachtspunt en groot belang voor wat betreft de beveiliging van informatie, deze moet niet onderschat worden en heeft bij SVH de hoogste prioriteit.

Van schriftelijk naar digitaal

- In dit hoofdstuk laten we zien wat het voor de processen betekent als de toetsing digitaal verloopt. Verder zetten we hier uiteen welke stappen je kunt zetten richting digitaal toetsen en welke besluiten er genomen moeten worden. Een belangrijk stap, en direct ook de stap waarin digitale toetsing het meest verschilt van schriftelijke, is de inrichting van de vragenbanken. Tot slot gaan we dieper in op de mogelijkheid om digitaal toetsen gefaseerd in te voeren.

digitaal

3.1 Toetslogistiek en digitale toetsing

Onder toetslogistiek verstaan we het hele proces van het voorbereiden en 'klaarzetten' van een toets tot en met het verstrekken van de uitslag en analyse van de examenvragen. We laten in deze paragraaf zien wat de toetslogistiek inhoudt in geval van digitale toetsing. Daarbij maken we onderscheid in het proces vóór het examen, tijdens de examenafname en het proces na het examen.

Vóór het examen

Zowel bij schriftelijk als bij digitaal toetsen zijn er een aantal aspecten die je voorafgaand aan het examen bepaalt. Het grote verschil is dat je bij digitaal toetsen deze aspecten in principe maar één keer instelt. Daarbij gaat het op hoofdlijnen om de volgende keuzes:

- ontwerp van de toets;
- samenstellen van de toets;
- aantal vragen;
- cesuur bepalen;
- opmaak toets;
- manier van inschrijven;
- keuze vast moment of flexibel.

Ontwerp van de toets

Voor een toets gebruik je een toetsmatrijs, een overzicht van hoe de leerstof over de examenvragen is verdeeld. De toetsmatrijs is meestal gekoppeld aan eindtermen, leerdoelen of exameneisen en geeft het aantal vragen per leerdoel aan of het gewicht van de vragen. De toetsmatrijs zorgt ervoor dat ieder examen dezelfde opbouw heeft. Voor een digitale toets is dat niet anders. Aan de hand van de toetsmatrijs ontwerp je de

complete toets en richt je de structuur van de vragenbank (de database) in. Ook hier geldt, een goed begin is het halve werk.

Vragen kunnen hergebruikt worden en het onderhoud is gemakkelijker.

Samenstellen toets: versies of at random

Het mooie van digitaal toetsen is dat je de toetssoftware zo kunt instellen dat er automatisch een examen wordt klaargezet, waarbij het systeem de vragen kiest. Dit heet at random toetsen. Je moet het systeem nog wel vertellen hoe het de toets 'klaar' moet zetten. Bij at random toetsen worden de vragen apart in de database gezet volgens een bepaalde structuur, bijvoorbeeld een opbouw die overeenkomt met de exameneisen. Het is dus niet zo dat uit een willekeurige bak met 200 vragen er zomaar 40 worden getrokken. Je wilt dat de verdeling van onderwerpen in het examen verloopt volgens het vooraf bepaalde ontwerp van de toets. Het grootste voordeel is dat vragen automatisch worden hergebruikt en dat het onderhoud van de vragen gemakkelijker is. Na de initiële opening van de vragenbanken zijn kostenbesparing en efficiëntie dan ook aan de orde bij digitale, at random toetsen.

Mocht je deze stap te lastig vinden, kun je kiezen voor het klaarzetten van vaste (versie)toetsen. De database bevat dan hele toetsen, bestaande uit een vaste opbouw van vragen, precies zoals bij schriftelijke toetsen. Het systeem kiest dan

steeds een compleet examen voor een bepaalde student. Ook deze variant kent een duidelijke kostenbesparing boven schriftelijke toetsen, omdat de inhoud geheim blijft en dus ook kan worden gebruikt om nieuwe toetsen samen te stellen. Voor meer informatie over at random toetsen verwijzen we naar de blog 'voordelen van at random toetsen':

Een extra optie bij at random toetsing is om toetsvragen die bij verschillende examens horen, in één vragenbank (database) op te nemen. Dat doe je wanneer er overlap is tussen examenvragen in verschillende examens. Voorbeeld van twee examens met overlap: een examen over Nederlandse vogelsoorten en een examen over Nederlandse broedvogels. Een andere toepassingsmogelijkheid is in geval van verschillende doelgroepen die licht afwijkende exameninhouden hebben. Een voorbeeld hiervan is Nederlandse broedvogels voor eerstejaars en Nederlandse broedvogels voor tweedejaars studenten. Ook hierin zit een efficiëntie voordeel. Als je meerdere examens of examens voor meerdere doelgroepen wilt opnemen, kun je de examenvragen labelen, zodat de toetssoftware alleen examenvragen voor dat examen/die doelgroep selecteert.

In paragraaf 2.2 en 2.3 vind je meer informatie over at random versus vaste toetsversies.

Aantal vragen

Een veel gehoorde vraag is hoeveel vragen je nu minimaal in een vragenbank moet opnemen. Het antwoord hangt af van een aantal factoren:

- De 'slijtage' van vragen; worden de vragen goed afgeschermd en hoe vaak worden ze gesteld? Bij een x-aantal keer kunnen ze bekend worden en zijn ze niet meer goed bruikbaar. Hoe groter de itembank, hoe kleiner de kans op slijtage.

- De normen die worden gehanteerd om over voldoende examenvragen te beschikken, ook in geval van herexamens en als slechte examenvragen uit de database worden verwijderd. Bij een net opgestarte vragenbank kun je bijvoorbeeld uitgaan van drie keer een examen trekken. Bij versie-examens zou je uit veiligheidsoverwegingen kunnen kiezen voor vier keer het examen kunnen trekken. Als je een examen met 30 vragen hebt, heb je dus minimaal 90 tot 120 vragen nodig, om mee te beginnen. Voor meer informatie over starten met een beperkte omvang van de database zie paragraaf 2.3.
- De normen die gehanteerd worden voor wat betreft de dekking van de leerstof versus de omvang van de leerstof. Bij uitgebreide leerstof moeten wel alle relevante aspecten in de database terugkomen. Dat betekent dat over alle relevante aspecten vragen moeten kunnen worden gesteld. Kandidaten mogen namelijk niet het idee krijgen dat bepaalde onderdelen van de leerstof toch nooit aan de orde komen.
- De onderwerpen die altijd in een examen terug horen te komen. Natuurlijk moet een examen qua inhoud valide zijn. Als maar een zeer klein deel van de te toetsen stof in het examen aan de orde komt, is dat niet het geval. Hoe uitgebreider de te toetsen leerstof, des te meer vragen zijn nodig in het examen. Hoe groter het examen, des te meer vragen in de database.
- Het aantal meetpunten dat de toets minimaal hoort te bevatten. Een globale richtlijn is minimaal 30 meetpunten, dat wil zeggen punten waarop een kandidaat kan scoren. Dit aantal gaat echter omhoog, wanneer de gokkansen stijgen. Bij een aantal ja/nee vragen in de toets (gokkans 50%) hoort het aantal meetpunten ook hoger te zijn. En hoe meer meetpunten, des te meer vragen. Stel een toets heeft bijvoorbeeld open vragen met 2 punten per vraag, dan zal de toets minimaal 15 vragen bevatten. Een itembank met alleen

open vragen, met elk 2 punten per vraag, zou dan minimaal 45 tot 60 vragen moeten bevatten. Deze norm ten aanzien van het aantal vragen in de database ligt bij gesloten vragen een stuk hoger.

Cesuur

Bij toetsing bepaal je altijd de cesuur: de grens tussen wel of niet slagen voor een examen (voldoende - onvoldoende). Soms doe je dat vooraf, soms achteraf. Bij digitaal toetsen vraagt de meeste toetssoftware om vooraf de cesuur vast te leggen. Er zijn veel manieren om de cesuur te bepalen. Een gemakkelijke manier is: de helft moet goed zijn. Een andere manier is: de gokkans te verrekenen in de cesuur. De gokkans is de kans dat de student een goed antwoord geeft, ook als hij het antwoord niet weet. Bij meerkeuzevragen met vier antwoordalternatieven is dat bijvoorbeeld 25%. Wanneer bij elke vraag gegokt wordt is er een kans dat 25% van de vragen correct beantwoord zijn. Voeg hierbij wat een student minimaal moet weten. Dat is meestal 50%. In dit geval is de cesuur kortgezegd: $50 + 25 = 75\%$. Doorgaans zijn cesuurformules van toepassing, waarin het aantal vragen en het aantal verschillende alternatieven, of de grootste gemene deler daarvan, wordt meegenomen.

Opmaak van de toets

Iedereen die examens maakt, weet het: het opmaken van een examen is een tijdrovend en intensief proces en bovendien nauwkeurig werk. Je wilt dat het examen er representatief uitziet en dat kandidaten niet afgeleid worden door fouten of onregelmatigheden in de opmaak. Ook bij digitale toetsing is er een opmaak, per vraag en voor de toets als geheel. Bij de opmaak van de examenvragen kun je alles instellen wat je ook met een tekstverwerker instelt. Sterker nog, de meeste systemen beschikken over een editor waarmee je examenvragen kunt opmaken. Het voordeel van digitaal examineren is dat je de opmaak van het examen als geheel in

één keer instelt, voor alle examens. Je kunt ook per examen afwijken, als je dat wil. Zo heb je altijd een examen in je eigen 'look and feel' en met minder inspanning dan bij schriftelijke examens.

Inschrijven

Het inschrijven van kandidaten bij digitale toetsing verloopt vergelijkbaar als bij schriftelijke examens. Daarbij is het van belang om te benoemen dat toetssoftware soms al ingebouwde modules heeft rondom inschrijving en betaling of met deze modules is uit te breiden. Ook dit gedeelte verloopt met digitale toetsing zeer efficiënt. Kandidaten kunnen zich online en individueel inschrijven, waarbij ze zelf een beschikbaar moment kiezen. Daarnaast is het mogelijk om groepen kandidaten in te schrijven, wat handig kan zijn wanneer het examen op één vast moment plaatsvindt. Lees meer hierover onder 'vast moment of flexibel'.

Vast moment of flexibel

Voor de start van het examen kies je of je het examen op een vast moment of flexibel gaat afnemen. Een vast moment wordt gekozen, als je wil dat iedereen tegelijk examen doet. Een reden hiervoor kan zijn dat je op een eigen locatie wil toetsen, waarbij er niet doorlopend gelegenheid is om kandidaten te ontvangen. Laat je niet in de war brengen: een vast toetsmoment valt niet persé samen met het gebruik van vaste versies in je vragenbank. Je kunt op een vast moment toetsen en de vragenbank at random inrichten of omgekeerd: flexibel toetsen met vaste versies in de examenbank. Een vast toetsmoment met vaste versies in de vragenbank lijkt wel het meest op schriftelijke toetsing.

Examenafname

Summatieve toetsing vindt gecontroleerd plaats, dat betekent onder andere dat er toezicht is. Dat toezicht kan geregeld worden op een eigen of een gehuurde toetslocatie. Ook is er de mogelijkheid om 'op afstand' onder toezicht te toetsen via online proctoring. Online proctoring is een afnamemethode, waarbij iedere kandidaat op een eigen locatie examen doet. Dat kan bijvoorbeeld thuis of op de werkplek zijn. De kandidaat kan daarmee in zijn eigen vertrouwde omgeving examen doen, vaak op een vooraf afgesproken moment. De instelling of organisatie bespaart op kosten van de afnamelocatie. Het toezicht op de examenafname wordt op afstand live of achteraf gedaan door een proctor.

Bij digitale toetsing kent het toetssysteem vaak allerlei middelen om de examenafname te vergemakkelijken, zoals verschillende rekenmachines, automatische weergave van 'tijd over', de voortgang tijdens het examen en de optie om tijdens het examen tijd bij te geven (bij calamiteiten). Bij de examenafname kun je ook allerlei zaken instellen voor speciale doelgroepen, zoals dyslectici en blinden of slechthzienden. Denk daarbij aan extra beschikbare tijd, een grotere schermweergave of ondersteuning middels verklanking.

Na het examen

Typerend voor digitale toetsing is de mogelijkheid om automatisch gesloten vragen na te kijken, feedback te geven en de uitslag te verstrekken. Je bepaalt zelf wat er al dan niet automatisch verloopt.

We gaan hieronder kort in op de volgende onderwerpen die van toepassing zijn op het proces na de afname van het examen:

- Correctie;
- Resultaat en analyse;
- Feedback bij summatieve toetsing;
- Feedback bij formatieve toetsing.

Correctie

Digitaal toetsen vergemakkelijkt de correctie, of je nu gesloten vragen gebruikt die automatisch worden nagekeken, of open vragen waar een corrector aan te pas komt. In beide gevallen heb je de zekerheid dat altijd het juiste antwoordmodel wordt gehanteerd, omdat het systeem alleen het meest actuele antwoordmodel laat zien. Je kunt bij open vragen bovendien één of meerdere correctoren per examenwerk instellen. Het systeem stuurt dan automatisch mailtjes aan de betreffende corrector(en), bijvoorbeeld dat er een examenwerk voor hem klaar staat om nagekeken te worden. De corrector corrigeert het examenwerk digitaal, waarbij beoordelingen worden opgeslagen in het systeem. Zo zijn ze altijd weer op te vragen en in te kijken.

Resultaat en analyse

Bij automatisch na te kijken vragen zijn de resultaten direct beschikbaar voor de kandidaat. Een kandidaat weet meteen of hij al dan niet geslaagd is. Vaak is ook de mogelijkheid aanwezig om een voorlopige uitslag te genereren. Vanuit de toetssoftware kun je bovendien informatie genereren over de moeilijkheidsgraad en de kwaliteit van de vragen. Daarnaast is het mogelijk om analyses te maken per opleider/corrector en zicht te krijgen op verschillen. Door naar de analyses van je toets te kijken kun je zien of en hoe het examen verbeterd kan worden. Natuurlijk is deze informatie ook te genereren bij schriftelijke examens, alleen is dat relatief veel werk. De vragen worden immers eenmaal gebruikt en analyse verloopt niet automatisch.

Feedback bij summatieve toetsing

Soms is het wenselijk dat bij summatieve toetsing feedback gegeven wordt, bijvoorbeeld wanneer een inzage niet tot de mogelijkheden behoort. In dat geval is een kandidaat geholpen met bijvoorbeeld feedback op het niveau van eindtermen. Zonder dat de kandidaat inzage krijgt in de vragen krijgt hij wel een studieadvies voor de herkansing.

Feedback bij formatieve toetsing

Het tussentijds toetsen is formatief en telt niet mee voor het eindcijfer. Omdat je hierbij leert van de toets moet je feedback krijgen. Digitale toetsing vergemakkelijkt het geven van feedback.

Wanneer je digitaal toetst in plaats van schriftelijk is formatieve toetsing een fluitje van een cent. Kwalitatief goede feedback is echter essentieel. Door te toetsen worden kandidaten bewust van datgene wat ze niet kennen (en kunnen). Ze worden uitgedaagd aan het werk te gaan.

Voor kandidaten wordt door het gebruik van formatieve toetsing veel duidelijker waar ze staan. Toetsen is een prettige, effectieve manier om te leren. Bijkomend voordeel: faalangstige kandidaten gaan beter presteren wanneer ze regelmatig een goede formatieve toets gemaakt hebben. Ze worden bevestigd door de feedback die ze krijgen en hoeven niet meer bang te zijn voor een toets. Ze weten immers dat ze de stof beheersen.

Vier tips voor het geven van feedback bij formatieve, digitale toetsing:

1. Bedenk van te voren op welk niveau je feedback wilt geven. Het geven van feedback kan op verschillende niveaus: op het niveau van eindtermen of onderwerpen, op het niveau van vragen of op het niveau van antwoorden. De verschillende mogelijkheden worden misschien beperkt door de gebruikte toetssoftware. Hattie & Timperley (2007) geven aan dat feedback op verschillende niveaus gegeven kan worden. Zowel op het niveau van juist/onjuist als bij het in gang zetten van reflectie op het denkproces bij de kandidaat;
2. Denk na over de manier van presenteren van de feedback. Wanneer feedback op het niveau van antwoordalternatieven gegeven wordt, ontstaat het risico dat de feedback negatief geformuleerd wordt. Probeer opbouwend te zijn in het geven van feedback;
3. Bedenk wanneer je feedback wilt geven:
 - Direct na het invullen van een vraag of oefening;
 - Direct na afsluiten van een toets of;
 - Wanneer het oordeel definitief is (nabespreken of inzage);

Mogelijkheden van feedback geven hangen ook af van het toetssysteem. Soms kan een kandidaat direct het antwoord aanpassen (tweede kans); daar past ook directe terugkoppeling bij. Uit recent onderzoek komt naar voren dat uitgestelde feedback mogelijk beter werkt dan directe feedback tijdens de oefening (Van der Kleij, 2017).

4. Probeer feedback zo onderhoudsarm mogelijk te geven. Daarmee wordt bedoeld: verwijs niet naar een bladzijde in een boek wanneer dat boek ieder jaar een herziene uitgave heeft. Dan moet je jaarlijks je feedback bijwerken. Beter is te verwijzen naar het onderwerp wat beter bestudeerd moet worden.

3.2 Stappen naar digitale toetsing

Op basis van bovenstaande informatie kun je al concluderen welke stappen je moet zetten om van schriftelijke naar digitale toetsing te komen. In deze paragraaf lichten we deze stappen, ofwel keuzes, toe. Achter in dit boekje vind je tevens een praktische checklist die je kunt gebruiken wanneer je de overgang van papieren naar digitale toetsing vorm geeft.

Uitgangspunt is dat de volgende documenten beschikbaar zijn:

- leerdoelen, eindtermen en/of toetstermen;
- een globale toetsmatrijs met de verdeling van de leerstof over de vragen en weging;
- vastgesteld uitgangspunt over de cesuur.

Belangrijke aspecten die nog wat nadere toelichting kunnen gebruiken voordat je aan de slag kunt met de checklist zijn:

- De inrichting van vragenbanken;
- At random versus vaste versies;
- Het ontwikkelproces van examenvragen.

De inrichting van vragenbanken

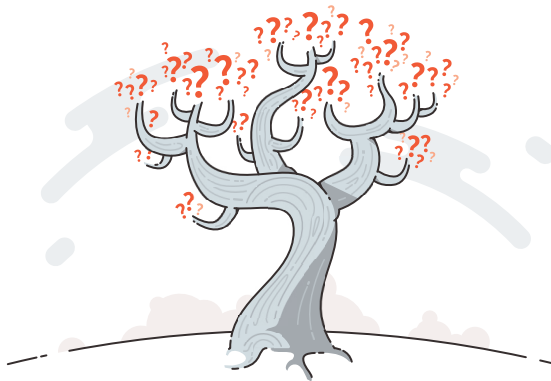
De inrichting van vragenbanken is specifiek voor digitale toetsing en neemt een belangrijke plek in. Een groot deel van de checklist richt zich dan ook hierop. Hierbij gaan we ervan uit dat het mogelijk moet zijn om flexibel te toetsen, dus toetsen op verschillende momenten. Waarom is het voor de checklist relevant om deze aanname te doen? Eigenlijk maar om één reden: Flexibele toetsing stelt eisen aan de gelijkwaardigheid van de examens. Een kandidaat die een week eerder of later komt, mag niet een heel andere toets dan zijn medecursist

krijgen. In paragraaf 2.3 over fasering bij de invoering van digitale toetsing gaan we dieper in op flexibel versus vast toetsen.

At random versus vaste versies

In het eerste deel van de checklist zie je over welke punten je doorgaans beslist bij het inrichten van een itembank. Daarbij maken we onderscheid tussen at random toetsen en toetsen met vaste versies. Je kunt uit de checklist opmaken dat het toetsen met vaste versies meer handmatige controle vergt, maar dat je vooraf minder hoeft te beslissen aangaande de structuur van de vragenbank. Voor meer informatie over de voor- en nadelen van at random versus vaste versies, zie de tabel aan het einde van deze paragraaf.

De structuur van de vragenbank met vaste versies bestaat alleen uit complete examens (zeg versie-examen 1 tot en met 10). De structuur van de vragenbank bij at random toetsen wordt apart ontwikkeld en wordt weergegeven in een vragenboom.



Een vragenboom is een op de toetsmatrijs gebaseerde, gedetailleerde weergave van de structuur van de vragenbank. Een vragenboom bevat informatie over:

- hoeveel vragen minimaal én maximaal worden gesteld per onderwerp;
- het aantal punten per vraag (of de range van punten per vraag of vragenset);
- het beheersingsniveau van de vraag (begrip, kennis of toepassing, zie Bloom, 1956);
- of een toets al dan niet casussen bevat.

Uit de checklist (zie bijlage achterin) kun je opmaken dat de toetsmatrijs overloopt in de vragenboom. Het verschil tussen een toetsmatrijs en een vragenboom zit in de toepassing ervan en de mate van detaillering. Een toetsmatrijs wordt breed en bij een veelheid aan toetsvormen toegepast, zowel bij schriftelijke als bij digitale toetsen en geeft de opbouw van de toets aan. Een vragenboom wordt toegepast bij digitale toetsing van theorie en geeft de structuur van de database aan.

Het ontwikkelproces van examenvragen

Tot slot merken we op dat het ontwikkelproces van de examens bij een at random bank anders verloopt dan bij een examenbank met vaste versies. Bij vaste versies ontwikkel je doorgaans per examen, dat wil zeggen dat één of twee auteurs gezamenlijk een complete toets maken. Bij at random toetsen ontwikkelen verschillende auteurs setjes vragen voor de vragenbank. Dat is het grote voordeel wanneer er moet worden ontwikkeld onder tijdsdruk; vele handen maken licht werk.

Tegelijkertijd maakt dit ontwikkelproces je minder kwetsbaar. Bij het afkeuren van vragen, vervalt namelijk niet de complete toets. Het ontwikkelproces wordt in de checklist buiten beschouwing gelaten. Er zijn allerlei stappen die je kunt zetten en die anders zijn dan bij schriftelijke toetsing, maar deze zijn ook organisatie- en projectgebonden.

Hieronder vind je een opsomming van de voordelen van at random versus vaste versies:

At random	Vaste versies
✓ Gemak bij het automatisch klaarzetten van de toets.	x Handmatig samenstellen van de toets.
✓ Overzichtelijke structuur van de database.	x Geen overzichtelijke structuur van de database.
✓ De toets voldoet automatisch aan de toetsmatrijs.	x Handmatig controleren of de toets aan de toetsmatrijs voldoet.
✓ De toetsconstructie is flexibeler, mogelijkheid om losse opdrachten onder verschillende toetsconstructeurs uit te zetten.	x Een of twee toetsconstructeurs werken aan één versie-examen. Wanneer vlak voor afname blijkt dat het examen niet voldoet, kan er een kwaliteitsprobleem ontstaan.
✓ Alle vragen komen maar één keer in de database voor, waardoor onderhoud efficiënt verloopt.	x Vragen kunnen op meerdere plaatsen voorkomen in de database, waardoor onderhoud inefficiënter is.

✓ Minder vragen nodig bij de start, bijvoorbeeld drie keer het examen kunnen trekken.	✗ Meer vragen nodig bij de start, bijvoorbeeld vier versie-examens.
✓ Vragen minder snel herkenbaar of minder gemakkelijk te onthouden, doordat ze automatisch gehusseld worden.	✗ Vragen sneller herkenbaar, gemakkelijker te onthouden doordat ze in een vaste versie worden aangeboden. (Bij voldoende aantal versies valt dit nadeel weg).
✓ Terugkoppeling op de kwaliteit van toetstermen; doordat de vragenbank (deels) op toetsterm is ingedeeld en per toetsterm wordt ontwikkeld kun je snel zien hoe het met de kwaliteit van de toetstermen gesteld is.	✗ Doordat de vragenbank is ingedeeld per examen en niet per toetsterm en er zo ook niet wordt ontwikkeld, kun je niet direct zien hoe het met de kwaliteit van de toetstermen is gesteld.
✗ Bij de overgang naar digitale toetsing hoeft vooraf geen structuur van de vragenbank te worden bepaald. Dit scheelt tijd bij de opstart. Daarnaast biedt deze werkwijze meer ruimte aan toetsconstructeurs voor wat betreft de opbouw van de toets. Deze ligt immers niet vast.	✓ Bij de start van digitale toetsing is het nodig om eerst aandacht te besteden aan het maken van keuzes rondom de inrichting van de vragenbank.

Afbeelding: at random versus vaste versies

Voor meer informatie over keuzes bij de inrichting van vragenbanken, zie de checklist achterin het boekje.

3.3 Digitaal toetsen gefaseerd invoeren

Hoe zorg je voor een naadloze overgang van schriftelijk naar digitaal toetsen? Je kunt het als een project zien, je kunt de werkprocessen helemaal veranderen of je kunt 'klein' beginnen. Hoe werkt dat dan, gefaseerd starten met digitale toetsing?

We behandelen hieronder enkele manieren om 'klein' te starten, om de hoeveelheid werk en benodigde tijd om 'digitaal te gaan' (in eerste instantie) te beperken:

Impact op bestaande processen van digitale toetsing

↑ Weinig impact / veel impact ↓	Afname	Inrichting database
	• Flexibel toetsmoment • Vast toetsmoment • Schriftelijke afname	• At random toetsen • Beperkte vragenbank • 'Sobere' inhoud van de vragenbank • Vaste versies

Afbeelding: schematische weergave impact digitale toetsing

Beperkte vragenbank

Een eenvoudige manier om het 'klein' te houden is om met een kleine vragenbank te starten. Daarbij maken we onderscheid tussen at random toetsen en toetsen met vaste versies. In een later stadium, als volgende stap, kun je dan de vragenbank eenvoudig uitbreiden.

Bij at random toetsen

Uitgangspunt kan bijvoorbeeld zijn dat ieder examen drie keer moet kunnen worden getrokken met verschillende vragen. Waarom drie keer? Bij 'drie keer trekken' kan een kandidaat twee keer herkansen met iedere keer compleet nieuwe vragen. Mocht hij dan een derde keer herkansen, krijgt hij wel vragen die hij al heeft gehad, maar dan in een compleet andere volgorde en 'gehusseld'. Het voordeel hiervan is dat de herkenbaarheid bij een eventuele derde herkansing laag is.

Bij toetsen met vaste versies

Uitgangspunt kan bijvoorbeeld zijn dat je start met vier examens, een beetje meer dan bij at random examens. Waarom zou je de norm van het aantal examenvragen ophogen? Omdat er iets niet in orde kan zijn met één van de examens, waardoor dat examen tijdelijk niet kan worden ingezet. Bovendien geeft het de kandidaat bij een eventuele derde herkansing toch een uniek examen.

'Sobere' inhoud van de bank

Mocht je bij de start van digitale toetsing opzien tegen het gebruik van allerlei verschillende vraagsoorten, kun je natuurlijk prima besluiten deze pas in een later stadium te gaan gebruiken. Start dan met alleen de vraagvormen die je ook gebruikt bij schriftelijke toetsing, bijvoorbeeld open vragen of een combinatie van open vragen en meerkeuzevragen.

Vaste versies

Bij toetsen met vaste versies is de omslag naar digitaal toetsen relatief klein. Dat komt doordat je de structuur van de database nog niet hoeft te hebben en dat je het ontwikkelproces ongewijzigd kunt laten. Zie ook de checklist in de bijlage. Het toetsen met vaste versies is eigenlijk vergelijkbaar met het werken met schriftelijke toetsen.

Vast moment

Wil je de overgang van schriftelijk naar digitaal nog kleiner maken, dan kun je kiezen voor toetsen op een vast toetsmoment. Vooral voor het verstrekken van de uitslag en het analyseren van resultaten is dat gemakkelijk. Procedures rondom uitslagbepaling en -verstrekking kunnen dan onveranderd blijven. Als je de uitslag wil veranderen naar aanleiding van de afname, kan dat, net zoals bij schriftelijke toetsing, groepsgewijs. Bij flexibele afname kun je ook uitslagen aanpassen, maar dat gebeurt dan vaak pas nadat kandidaten hun uitslag al hebben ontvangen. Laat je overigens niet in de war brengen: Toetsen op een vast moment is iets anders dan toetsen met vaste versies. Je kunt ook toetsen op een vast moment met een at random opgestelde vragenbank.

Schriftelijke uitdraai

Als je nu wilt dat de examenafname nog meer lijkt op de schriftelijke variant, kun je (een deel van) de examens schriftelijk afnemen. Dat kan door een uitdraai te maken van de digitale examens. Het grote voordeel is dat je dezelfde locaties als bij schriftelijke toetsing kunt gebruiken. Er zijn immers geen computers of een stabiele online verbinding nodig. Je kunt

dan profiteren van de voordelen van digitale toetsing, zonder dat je de processen rondom afname hoeft te veranderen. Een bijkomend voordeel van digitale toetsing is dat, ook na schriftelijke afname, digitale verwerking van de antwoorden en analyse van de vragen kan plaatsvinden.

Zorg voor een soepele overgang van papier naar digitaal.

Pilots

Bij de overgang naar digitale toetsing is het goed denkbaar dat er nieuwe vragen worden ontwikkeld of zelfs een compleet nieuw examen wordt ontwikkeld. In ieder geval kan het qua 'look and feel' voor kandidaten als nieuw overkomen. Hoe zorg je nu voor een soepele overgang? Hiervoor onderscheiden we de volgende manieren:

- pilot door vaste toetsmomenten;
- pilot met proefperiode (flexibele toetsing);
- pilot voorafgaand aan openstelling.

Pilot door vaste toetsmomenten

Hoe ziet een dergelijke pilot eruit? Dat kan op verschillende manieren. De eerste, eenvoudige manier, is te starten met digitaal toetsen op vaste momenten. Je organiseert bijvoorbeeld 2 of 3 vaste toetsmomenten. Na ieder moment doe je een vragenanalyse en stel je uitslagen bij, indien nodig. Zo test je meteen de kwaliteit van de vragen en beperk je eventuele onrust onder kandidaten en andere stakeholders.

Pilot met proefperiode

Een andere vorm van pilots draaien, toe te passen bij flexibele toetsing, is het instellen van een proefperiode van bijvoorbeeld drie maanden. In die drie maanden begin je met een lagere cesuur. Je communiceert dat ook aan kandidaten. Daarna evalueer je hoe het gesteld is met de slagingspercentages en ga je eventueel toe naar een wat hogere cesuur. Deze cesuurverhoging kun je gradueel opbouwen, in stappen. Dit vergt wel de nodige tact en duidelijkheid, om onrust te voorkomen.

Pilot voorafgaand aan openstelling

Stel je wilt voorafgaand aan het openstellen van de bank zeker weten hoe de vragen scoren, dan kun je kandidaten uitnodigen om examen te doen. Heldere communicatie is wel een vereiste. Wat is de status als ze slagen? Wordt er examengeld terug gestort als ze niet slagen? Etcetera.

Kwaliteit van examen- vragen

- De feitelijke inhoud van de vragenbank, de examenvragen, is van essentieel belang voor het succes van digitale toetsing. Daarom gaan we in dit hoofdstuk in op verschillende kwaliteitscriteria, de analyse van examenvragen en vraagsoorten met bijbehorende scoringsmogelijkheden.

kwali

4.1 Kwaliteitscriteria examen vragen

Wanneer heeft een toets een goede kwaliteit? De termen validiteit en betrouwbaarheid komen aan de orde bij het in kaart brengen van de kwaliteit van de toets. Achter deze termen gaat een hele wereld schuil, maar hier beperken we ons tot de grootste algemene deler. Voor meer informatie over de begrippen betrouwbaarheid en validiteit verwijzen we naar de toelichting van Sluiter, Hemker en Eggen (2018). Een toets is betrouwbaar als het resultaat ervan niet berust op toeval. Dat is bijvoorbeeld wel het geval als er storende factoren zijn bij de afname of als er fouten in vragen dan wel in de beoordeling daarvan zitten. Een valide toets meet wat deze verondersteld wordt te meten en sluit aan op het doel waarvoor de toetsresultaten worden gebruikt. Een voorbeeld van de toepassing van dit criterium is toetsvragen die aansluiten bij specifiek geformuleerde leer- of toetsdoelen en bij het niveau waarop wordt getoetst. Klachten over de validiteit van een toets hangen vaak samen met de betekenis die stakeholders verlenen aan de resultaten van een toets.

Kortom: een goede kwaliteit van toetsing houdt in dat een kandidaat die ook werkelijk de leerstof beheerst, slaagt en dat stakeholders de beleving hebben dat de toetsing eerlijk is. Met examen vragen die aan kwaliteitscriteria voldoen is doorgaans grote winst te behalen en het voldoen aan deze criteria is meestal ook vrij gemakkelijk te realiseren. Bewustwording van deze criteria en het kunnen toepassen ervan is dan van belang. De toetstechnische criteria die worden gehanteerd hebben betrekking op:

1. Relevantie;
2. Objectiviteit;
3. Specificiteit;
4. Efficiëntie.

1. Relevantie

Hoort de vraag bij de leer- en toetsdoelen in het examenprogramma? Betreft het dan ook relevante kennis of gaat het over zulke details dat ofwel niemand dit kan weten ofwel niemand deze kennis ooit zal gebruiken? De vraag moet gaan over vakinhoud waar een beroepsbeoefenaar iets aan heeft. Een voorbeeld: In een examen over productkennis in de levensmiddelenhandel is het wellicht minder van belang om de voedingswaarde van pindakaas uit je hoofd te kennen. Hiervan is immers het nodige te lezen op het etiket.

2. Objectiviteit

Is het juiste antwoord bij deze vragen altijd goed of zijn er ook situaties denkbaar waarin het 'juiste' antwoord feitelijk niet juist is? Kunnen er ook andere antwoorden goed gerekend worden? Een objectieve vraag leidt doorgaans niet tot discussie. Zie onderstaand voorbeeld van een niet-objectieve vraag.

Welke kleuren heeft de Nederlandse vlag? Kruis alle juiste antwoorden aan.

- A. Rood
- B. Blauw
- C. Wit
- D. Oranje

Juiste antwoord: A, B en C

De vraag is of antwoord D, oranje, ook niet goed gerekend moet worden. Stel dat de vlag een wimpel heeft, dan is deze oranje. Antwoord D is wellicht niet het meest relevante

antwoord, maar echt fout is het ook niet. In ieder geval kan de vraag tot discussie leiden.

3. Specificiteit

Een vraag dient zo gericht gesteld te zijn dat iemand die de stof beheerst, de vraag correct zou moeten kunnen beantwoorden en iemand die de stof niet beheerst niet. Een specifieke vraag maakt dus onderscheid in goede en slechte kandidaten. Zie hieronder voor een voorbeeld van een niet-specifieke (open) vraag, die de meeste volwassenen wel kunnen beantwoorden.

Beschrijf de leiderschapsstijlen van een veel toegepaste managementtheorie.

Antwoord:

In de theorie van Hersey en Blanchard worden vier stijlen beschreven.

Beschrijving:

- Delegeren: het overlaten van taken aan medewerkers, weinig sturing en weinig ondersteuning;
- Ondersteunen, overleggen: het helpen van medewerkers, weinig sturing;
- Overtuigen, motiveren: veel taakgerichte sturing en veel ondersteuning;
- Opdragen, instrueren: veel sturing, weinig ondersteuning.

Andere antwoorden ter beoordeling van de corrector.

Het probleem met deze vraag is dat deze weinig gericht is; er zijn verschillende managementtheorieën en -modellen die vaak worden toegepast. Daarnaast wordt niet vermeld aan welke eisen de beschrijving moet voldoen. Er zijn op deze manier erg veel antwoorden die goed gerekend zouden moeten worden.

4. Efficiëntie

Om aan het criterium van efficiëntie te voldoen is het van belang de informatie in de vraag te beperken tot alleen die informatie die nodig is voor de beantwoording van de vraag. Een voorbeeld van wat we vaak zien is dat een casustekst het hele artikel van een dagblad bevat als achtergrondinformatie. Het advies is dan om dergelijke informatie niet in het examen maar in de lesstof op te nemen. Een ander voorbeeld is dat de kandidaat, door het opnemen van een dubbele ontkenning, de vraag meerdere keren moet lezen om deze goed te begrijpen. Ontkenningen zou je het best vet of cursief kunnen maken zodat ze de aandacht trekken. Feitelijk vallen taalfouten en ingewikkelde taalconstructies onder het criterium 'efficiëntie'.

| 4.2 Analyse van examenvragen

Daar waar kwaliteitscriteria gebruikt worden bij de toetsconstructie, gebruik je analysegegevens achteraf. Hier belichten we kort enkele waarden, inclusief betekenis, die toetssoftware automatisch levert:

- p-waarde;
- a-waarde;
- rit-waarde.

De p-waarde is het percentage van de kandidaten dat de vraag goed beantwoord heeft.

Dat wil zeggen dat een hoge p-waarde, op een relatief eenvoudige vraag wijst. Van deze groep kandidaten heeft bijna iedereen de vraag goed. Te moeilijke en te makkelijke vragen zijn niet wenselijk in een toets. Te makkelijke vragen leveren geen bijdrage aan een toets en leiden de goede kandidaat af. De kandidaat denkt dat het zo eenvoudig niet kan zijn. Voor een item is een p-waarde tussen 0,3 en 0,8 optimaal.

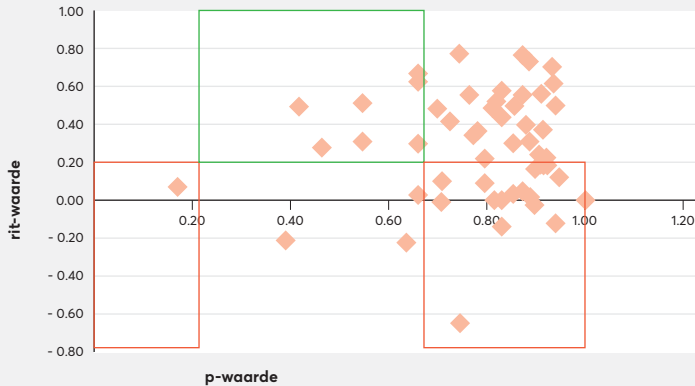
De a-waarde verwijst naar de afleider en geeft het percentage van de kandidaten wat voor deze afleider gekozen heeft.

Het doel van de afleiders is een plausibel antwoord te geven voor die kandidaten die de stof niet beheersen. Door een afleideranalyse te doen zie je meestal erg snel dat meerkeuze vragen slechts drie antwoordalternatieven hoeven te hebben. Bij meer dan twee afleiders (twee afleiders plus het goede antwoord zijn drie keuzealternatieven) zie je vaak dat enkele afleiders helemaal niet gekozen worden. Je kunt besluiten om die niet gekozen alternatieven niet meer aan te bieden.

De rit-waarde geeft aan hoe goed het item afzonderlijk, hetzelfde meet als de toets in het geheel. Oftewel hoe goed het betreffende item in de toets past.

De rit-waarde laat zien hoe het item verschil maakt tussen kandidaten met hoge scores op de toets en kandidaten met lage scores op de toets. De rit-waarde geeft het onderscheidend vermogen van het item aan.

Gecombineerde vraaganalyse (p en rit)



Afbeelding: Gecombineerde vraaganalyse

Wanneer de p-waarde en de rit-waarde tegen elkaar uitgezet worden in een grafiek is overzichtelijk te zien welke items kritisch beoordeeld moeten worden. De items in de rode vlakken moeten in ieder geval onder de loep genomen worden. Hier valt op dat er wel erg veel makkelijke items zijn (hoge p-waarden). De norm voor een acceptabele rit-waarde varieert. In de literatuur wordt over het algemeen de volgende normering aangehouden (Veldhuijzen, Goldebelld & Sanders, 1993):

- 0.40 en hoger zeer goed
- 0.30 – 0.39 goed
- 0.20 – 0.29 twijfelachtig
- 0.19 en lager slecht

4.3 Vraagsoorten

Bij digitale toetsing zijn veel vraagsoorten mogelijk. Verkijk je niet op de toetssoftware leverancier die honderd vraagsoorten aanbiedt. De kans is klein dat je ze alle honderd gaat gebruiken. Met acht vraagsoorten kun je al behoorlijk variëren. Onderstaand een aantal omschrijvingen van vraagsoorten die in een digitale toets gebruikt kunnen worden.

Vragen die nagekeken moeten worden of in de correctielogistiek (indien aanwezig) komen:

1. Open vraag

Hierbij moet de kandidaat zijn antwoord definiëren in het antwoordveld. Naar keuze van de toetsconstructeur is een editor beschikbaar voor de kandidaat.

2. Upload vraag

Hierbij moet de kandidaat aanpassingen doen in een bijlage en deze bijlage opslaan en uploaden. Soms kan de software verbinding maken met externe applicaties zoals Excel of een ander softwareprogramma. Dat gebruik je met name bij een praktijktoets. Niet iedere toetssoftware biedt echter deze mogelijkheid.

Vragen die automatisch nagekeken worden:

3. Multiplechoicevraag

Hier moet één antwoord geselecteerd worden uit de aangeboden antwoorden.

4. Multiple select vraag

Hier kunnen uit de aangeboden opties meerdere antwoorden geselecteerd worden.

5. Hotspotvraag enkelvoudig

Er kan een antwoord worden ingevoerd door op de juiste locatie binnen de afbeelding te klikken, daar wordt een pijltje (marker) geplaatst. Daarna kan het antwoord gewijzigd worden door het pijltje (marker) te slepen met de muis. Let op: dit is geen multiplechoicevraag, tenzij alleen voorgedefinieerde, zichtbare hotspots aangewezen worden. Met de moderne technieken is dat echter niet nodig en kun je een hotspot willekeurig in het plaatje definiëren.

6. Hotspotvraag meervoudig

Hier kunnen meerdere antwoorden worden ingevoerd door op de juiste locaties binnen de afbeelding te klikken. Daarna kan het antwoord desgewenst gewijzigd worden door het pijltje (marker) te slepen met de muis.

7. Video-hotspotvraag

Deze vraagsoort bevat een filmpje waar op het juiste moment geklikt moet worden. Dat is bijvoorbeeld te gebruiken wanneer het doel is om te testen of de kandidaat een bepaald moment in gesprekssituaties kan aangeven. Het antwoord kan vervolgens gewijzigd worden door eerst op het kruisje en dan op een nieuw punt in het filmpje te drukken.

8. Rangschikvraag (ordeningsvraag)

De antwoordmogelijkheden (dit kan tekst zijn, maar ook plaatjes zijn mogelijk) moeten in een juiste volgorde gezet worden.

9. Sleepvraag

Objecten (tekst of plaatjes) moeten op de juiste plaatsen binnen een afbeelding worden gesleept.

10. Matching/matrixvraag

Deze vraag bestaat uit een aantal rijen en kolommen, waarbij de student in de juiste rij/kolom kruisjes zet. Dit is een passende vraagsoort voor het toetsen van kennis waarbij relaties tussen twee variabelen moeten worden gelegd. Bijvoorbeeld welke struiken en bloemen horen bij welke soort vegetatie?

De volgende vragen kennen in de correctie verschillende vooraf gedefinieerde correcte antwoorden, en worden automatisch nagekeken:

11. Numerieke vraag

Er kan een antwoord ingevoerd worden in de vorm van een getal. Soms is het ook mogelijk vragen te klonen met behulp van variabelen. De som (vraag) heeft dan verschillende verschijningsvormen, je toetst wel steeds dezelfde som.

12. Invulvraag enkelvoudig

In het invoerveld kan een antwoord ingegeven worden, in de vorm van een korte tekst of een enkel woord. In het onderliggende correctiemodel bij enkelvoudige invulvraag zie je dat meerdere antwoorden als correct beoordeeld kunnen worden. Hierdoor kunnen veel voorkomende spelfouten toch automatisch goed gerekend worden.

13. Invulvraag meervoudig

Er kunnen meerdere antwoorden ingevoerd worden in de aangeboden invoervelden, zowel tekstueel als numeriek.

Door gebruik te maken van moderne technieken is het mogelijk formuleren, filmpjes en audiobestanden op te nemen in een vraag.

Tip: Media die je voor onderwijs hebt gebruikt kun je hergebruiken voor een toets. Houd er wel rekening mee dat je in dat geval bij de afname zorgt voor koptelefoons.

Tip: Denk aan het auteursrecht bij het gebruik van media. Wanneer je twijfelt of je media mag gebruiken vraag dan toestemming aan de maker.

4.4 Scoring van verschillende vragen

Nu we gekozen hebben welke vraagsoorten gebruikt gaan worden in de toets moet ook nagedacht worden over de waarde van de vraag in de toets. Met andere woorden: hoe waardeer je de antwoorden? Bij een multiplechoicevraag is scoring vrij eenvoudig: het antwoord is goed of fout. We noemen dat dichotoom. Bij een polytome scoring kan de vraag echter ook gedeeltelijk goed zijn. Bij een multiple select vraag kan het gebeuren dat de kandidaat maar één van de twee correcte alternatieven kiest. Hoe beoordeel je dat?

Scoring van vragen:

- dichotoom: kandidaat krijgt alleen (alle) punten als hij de vraag helemaal juist beantwoordt. Bij een foutief antwoord krijgt de kandidaat geen punten.
- polytoom: per juist antwoord krijgt de kandidaat een deel van de maximale vraagwaardering. Het kiezen van een afleider (fout antwoord) kan leiden tot puntenaftrek. De ondergrens van de vraag is nul.

De scoring wordt als volgt berekend:

(Percentage gekozen juiste antwoorden – percentage gekozen afleiders) * maximale vraagwaardering

Een veel gebruikte vorm van polytome scoring is de productregel. In de productregel wordt rekening gehouden met de mogelijkheid dat een kandidaat alle antwoorden aankruist en daarmee al een deel van de punten kan halen. Het bijzondere aan de productregel is dat de score van de kandidaat voor die vraag omlaag gaat wanneer deze onterecht een fout antwoord als 'goed' aankruist.

Voor de scoring van de rangschikvraag en de sleepvraag wordt meestal de quotiëntregel gebruikt. De score wordt berekend volgens een ingebouwde formule op basis van het aantal mogelijke volgordes van steeds twee objecten:

(Als er in totaal 4 objecten zijn: A,B,C,D dan zijn er 6 te testen volgordes: A boven B, A boven C, A boven D, B boven C, B boven D, C boven D).

Hierbij geldt dat je tweederde van de antwoorden goed moet hebben om de helft van de punten te krijgen.

Bij digitale toetsing wordt vaak voorafgaand aan het ontwikkelen van toetsen gekozen welke manier van scoren van toepassing is. De belangrijkste vraag daarbij is of er deelpunten mogen worden gegeven, met andere woorden of polytome scoring toegepast wordt. Dit vergt wel wat uitleg richting betrokken stakeholders, maar is wel een eerlijke manier van beoordelen, omdat een kandidaat voor alle goede antwoorden of aspecten daarvan (deel)punten ontvangt. Digitale toetsing wordt nog wel eens tegengehouden omdat het scoren van digitale toetsen oneerlijk zou zijn, je wordt afgerekend op

goed of fout. Uit de verscheidenheid van bovenstaande scoringsregels, zie je dat dit onterecht is. Sterker nog, digitale toetsing maakt het scoren van antwoorden objectiever doordat een itemauteur vooraf goed na moet denken over het instellen van de juiste scoringsregel waarmee recht gedaan wordt aan de beantwoording van de vraag (Examens, NR3,2010). Let op: Niet alle toetssoftware beschikt echter over de uitgebreide scoringsregels.

De keuze van toetssoftware

- De keuze van toetssoftware is een lastige. Er zijn vele pakketten met een uiteenlopend kostenplaatje en het is niet altijd even gemakkelijk om in te schatten met welke toekomstige wensen je rekening moet houden.

keuze

De keuze voor toetssoftware is er één die je voor de langere termijn maakt. Kijk dan ook naar de specifieke eisen waaraan het pakket nu en in de toekomst zou moeten voldoen. Wellicht gaat het kandidatenaantal of het aantal gebruikers of toetslocaties als gevolg van interne veranderingen toenemen. Bijvoorbeeld: in een grote internationaal opererende organisatie worden examens nu nog op een zelfde tijd en een aantal centrale plaatsen afgenomen. In de toekomst is het plan om toetsing tijds- en plaatsonafhankelijk te laten plaatsvinden. Dat stelt eisen aan de toetssoftware. Dit is maar één voorbeeld, maar er zijn vaak een groot aantal onzekere factoren.

Om je te helpen daarover na te denken schetsen we een aantal aandachtspunten:

1. Techniek
2. Veiligheid (AVG)
3. Rollen en rechten
4. Modulaire opbouw
5. Boeken en betalen
6. Gebruiksvriendelijkheid
7. Open vragen of alleen gesloten?
8. Beschikbaarheid statistische gegevens
9. Multi-media toepassingen
10. Koppeling met andere software
11. Afname opties
12. Maatwerk

1. Techniek

Gaan er meerdere mensen, verspreid over verschillende locaties gebruik van de toetssoftware maken? Denk daarbij aan toetsconstructeurs en beoordelaars van vragen. Dat vraagt om een webbased systeem, waarbij opslag van vragen op

een centrale en goed beveiligde plaats gebeurt. Daarnaast is het van belang vragen te stellen over de beveiliging van het systeem. Is deze goed geregeld? Zeker wanneer het summatieve toetsing betreft, is dat van belang. Bij systemen die alleen geschikt zijn voor formatief toetsen is deze beveiliging niet altijd optimaal. Tot slot is het van belang om te onderzoeken hoe de software zich gedraagt, wanneer er een grote hoeveelheid kandidaten deelneemt. Zijn de gebruikte servers daarvoor geschikt?

2. Veiligheid (AVG)

In hoeverre is het systeem toegerust om aan de eisen die de AVG stelt te voldoen? De beveiliging, de opslag en het gebruik van persoonsgegevens en de mogelijkheid om deze te verwijderen en aan te passen, zijn van belang.

3. Rollen en rechten

Bij de inrichting van examenprocessen zijn er doorgaans verschillende verantwoordelijkheden en bevoegdheden. In hoeverre kun je met de digitale toetssoftware daarbij aansluiten. Wie is uiteindelijk verantwoordelijk voor de inhoud van de examenbank? Wie mag vragen invoeren, controleren en reviseren? Wie mag aanpassingen doen aan de gebruikte sjablonen en aanpassingen doen op het niveau van de hele toets etc.? Bij het onderzoek naar passende toetssoftware is het van belang te weten welke eisen je stelt en te achterhalen in hoeverre het systeem daaraan voldoet.

4. Modulaire opbouw

Hoe is het pakket opgebouwd en is het ook mogelijk om het pakket in delen te gebruiken? Denk hierbij bijvoorbeeld aan een module voor afname op locatie, voor toezicht, voor inschrijving en registratie en een module om gegevens te analyseren en rapporten te distilleren. Is er wellicht ook een mogelijkheid om een 'light' versie te gebruiken, om bijvoorbeeld formatief te toetsen?

5. Boeken en betalen

Hoe gemakkelijk is het voor kandidaten om examens te boeken, inschrijvingen te wijzigen of te betalen? Zijn er mogelijkheden om een koppeling te maken tussen de website en de betreffende toetssoftware? En op welke wijze is communicatie met kandidaten vervolgens mogelijk? Is er een mailmodule en hoe eenvoudig is het om daar maatwerk van te maken?

Als gebruiker heb je baat bij een overzichtelijke structuur.

6. Gebruiksvriendelijkheid en kwaliteit

Hoe eenvoudig is het om vragen in te voeren? Kan dat gebeuren door relatief ongetrainde mensen? Het kan dan

heel relevant zijn om het toetsstelsel toegankelijk te maken voor toetsconstructeurs en beoordelaars van vragen en de verschillende groepen verschillende rechten te geven voor invoer, revisie en vaststelling. Daarnaast kan het van belang zijn om in het stelsel zichtbaar te maken welke vragen al zijn vastgesteld, welke niet en welke reeds gereviseerd zijn. Hoe wordt de status van vragen zichtbaar?

Verder is het een belangrijk punt van aandacht om te bekijken welke mogelijkheid tot analyseren van de resultaten het stelsel heeft. Bevat het stelsel al formules waarmee p- en rit-waarden kunnen worden berekend, ook van andere vraagsoorten dan meerkeuzevragen? Sommige systemen vragen om zelfwerkzaamheid van de gebruiker op dit gebied, terwijl andere systemen deze optie standaard geïntegreerd hebben.

Tot slot heb je als gebruiker veel baat bij een overzichtelijke structuur. Hoe zijn de vragen in de database geordend? Zie je direct de structuur of lijkt het eerder een puzzel?

7. Open of alleen gesloten vragen?

Hoe kan de leerstof het beste worden getoetst, met open of gesloten vragen? Open vragen zijn bijvoorbeeld nodig, wanneer getoetst moet worden of een student zelf de antwoorden kan formuleren of wanneer het gaat om een 'hogere orde' vaardigheid, zoals analyseren. Op het moment dat dat aan de orde is, zoek dan naar een stelsel dat deze mogelijkheid biedt. Daarnaast is er ook de vraag in hoeverre de manier van corrigeren ruimte laat: is het mogelijk om meerdere correctoren te hanteren? En zijn er diverse mogelijkheden om tot het cijfer te komen?

8. Beschikbaarheid statistische gegevens

Niet ieder toetssoftware systeem beschikt over dezelfde mogelijkheden om statistische gegevens te distilleren. Denk daarbij aan p-, rit- en a-waarden maar ook aan gegevens over scoring door correctoren en een weergave van de gokkans per vraag. Soms heb je ook informatie nodig over welke kandidaten een bepaalde examenvraag in hun examen hebben gehad. Dat is bijvoorbeeld het geval als je gebruikt maakt van flexibele toetsing en je achteraf één of meerdere vragen van het examen wil uitsluiten.

9. Multimedia toepassingen

Bij verschillende vraagsoorten kan het nuttig zijn om beeldmateriaal, in de vorm van illustraties en video's in te zetten. Hoe eenvoudig is het daarvan gebruik te maken in de software? Welke vereisten, in termen van beeldformaten en omvang van bestanden, zijn daarop van toepassing?

10. Koppeling met andere software

Is het te verwachten dat een koppeling met andere software nodig is? Dan kan aan de orde zijn wanneer er tijdens het examen gebruik moet worden gemaakt van andere programma's, bijvoorbeeld Excel of een ander boekhoudprogramma. Sommige systemen bieden daartoe de mogelijkheid.

11. Afname opties

Bij schriftelijke afname gelden heel andere mogelijkheden dan bij digitale afname. In dit boekje hebben we er diverse behandeld. In hoeverre is het nodig om tijd- en/of plaatsongebonden te toetsen? Denk daarbij wel aan de toekomstbestendigheid van het systeem. Wat als kandidaten over het hele land of wellicht in het buitenland wonen? Dan is het wel verstandig daar bij de start rekening mee te houden. Kijk bij de afname opties ook naar de mogelijkheden voor dyslectici, visueel gehandicapten etc. Of naar de mogelijkheid om een examen offline af te leggen; er zijn wellicht situaties denkbaar waarin er geen internetverbinding is.

12. Maatwerk

Hoe prettig is het om de herkenbaarheid voor gebruikers en kandidaten te vergroten door een examen te creëren in de eigen stijl en met het eigen logo? Diverse softwarepakketten bieden deze mogelijkheid. Zijn er maatwerkaanpassingen in het systeem mogelijk? Is de leverancier bereid om aanpassingen te doen naar aanleiding van specifieke wensen. Staat deze open voor innovatieve ideeën en suggesties voor ontwikkeling?

Ook al is de keuze complex, er zijn verschillende aspecten te benoemen die van belang zijn. In vogelvlucht zijn deze benoemd, maar per systeem is een gedetailleerde invulling van toepassing. Deze aspecten kunnen je helpen om de juiste vragen te stellen aan toetssoftware leveranciers.

Innovatie in digitale toetsing

- Mensen verlangen steeds meer en steeds vaker dat gebruik gemaakt kan worden van digitale hulpmiddelen. In de opleidingssituatie zie je dit ook terug. Veel lesmateriaal wordt digitaal aangeboden, sommige scholen werken al met laptops en/of tablets. Het gevolg is dat mensen steeds meer gewend zijn aan het gebruik van digitale tools. Dat pleit ervoor om ook toetsen zoveel mogelijk digitaal aan te bieden.

Innovat

Het is goed om je te realiseren dat op dit moment de beoordeling van open vragen nog vrijwel altijd door menselijke beoordelaars plaatsvindt. Gesloten vragen worden automatisch nagekeken. Menselijke beoordeling heeft een grotere foutmarge dan automatische beoordeling. Verschillende soorten vooroordelen kunnen een rol spelen bij menselijke beoordeling. In de toekomst is het niet ondenkbaar dat met behulp van Artificial Intelligence (AI) open vragen automatisch kunnen worden beoordeeld.

Hieronder benoemen we nog enkele andere innovaties binnen digitaal toetsen:

- Vaardigheidstoetsing
- Online proctoring
- Fraudepreventie

Vaardigheidstoetsing

Kijkend naar digitale toepassingen op het gebied van vaardigheidsonderwijs en -toetsing hebben de laatste jaren relevante ontwikkelingen, zoals serious gaming en virtual reality, een vlucht genomen. Serious gaming is een vorm van onderwijs en toetsing die verder gaat dan een spel spelen voor het plezier. Het leereffect van het gebruik van serious gaming is al aangetoond middels onderzoek. Van serious gaming kun je meer leren dan van e-learning (bron: itpedia.nl). Virtual reality in het onderwijs is eveneens relevant, met name als het gaat om de beleving van studenten en kandidaten. Doordat je met virtual reality het gevoel hebt dat je deel uitmaakt van de situatie, is het gemakkelijker om te leren maar ook om de juiste antwoorden te geven. Vooral het trainen van meer complexe vaardigheden via serious gaming neemt een enorme vlucht. Ook toetsing, in de zin van predictieve toetsing ten behoeve

van het bedrijfsleven vindt in serious gaming en virtual reality een toepassing. Denk daarbij bijvoorbeeld aan toetsing voor luchtverkeersleiders. Middels predictieve toetsing met serious gaming wordt de geschiktheid van kandidaten voor de functie van luchtverkeersleider getoetst. Daarnaast valt te denken aan het predictief en summatief toetsen van het leidinggevend inzicht van kandidaten middels virtual reality. Toetsing middels serious gaming en virtual reality zorgt doorgaans voor een hoge face validity (de beleefde validiteit). De relevantie voor de beroepspraktijk van deze toetsvormen, zoals deze door kandidaten wordt ervaren, is hoog.

Er kan op afstand meegekeken worden op het scherm van de kandidaat.

Online proctoring

Een van de manieren van afname die sterk in ontwikkeling is en een aanzienlijke verhoging van de flexibiliteit voor kandidaten inhoudt, is online proctoring. Deze optie is al genoemd in paragraaf 2.1., maar deze behandelen we hieronder nogmaals, om inzicht te geven in hoe online proctoring precies werkt. De kandidaat maakt opnames van zijn omgeving en identificeert zichzelf, met behulp van een webcam en een mobiele telefoon. Ook kan er meegekeken worden op het scherm van de kandidaat. De proctor controleert zodoende de omgeving van de kandidaat op de aanwezigheid van eventuele

onregelmatigheden of hulpbronnen. Er is een variant waarbij de proctor direct contact heeft met de kandidaat en een variant waarbij de proctor achteraf controleert op onregelmatigheden. De laatste variant geeft meer vrijheden voor wat betreft het moment van examenafname, maar er kan dan niet tijdens de afname worden ingegrepen, bijvoorbeeld als een camera niet goed staat. In beide gevallen is het van belang om vooraf maatregelen te nemen ter bescherming van persoonsgegevens en kandidaten om toestemming te vragen voor de toepassing van online proctoring. Bij het verzamelen, registreren en opslaan van beeldmateriaal is dat in het bijzonder van belang.

Fraudepreventie

Tot slot zijn er methoden voor fraudepreventie in ontwikkeling om onverwachte antwoordpatronen en tijdsbesteding per vraag op het spoor te komen. Op basis van data analyse kan dan mogelijke fraude worden opgespoord. Dat kan bijvoorbeeld het geval zijn wanneer een kandidaat bij de beantwoording van moeilijke vragen geen bedenktijd nodig heeft én goed scoort en bij makkelijke vragen een relatief lange bedenktijd neemt. Deze techniek is nog in ontwikkeling en de bevindingen die hieruit komen kunnen niet voor honderd procent voor 'waar' worden aangenomen. Deze techniek wordt dan ook nog niet losstaand van andere bevindingen toegepast, zoals concrete waarnemingen op de toetslocatie. Als een kandidaat zich bijvoorbeeld verdacht gedraagt, wordt in tweede instantie gekeken naar indicaties in de data analyse.

Kortom: digitale toetsing is op veel vlakken in ontwikkeling.

7

Literatuur

Bronne

Literatuur

- Eggen, T.J.H.M., & Sanders, P. F. (Eds.) (1993), *Psychometrie in de praktijk*, Arnhem: Cito.
- Hattie, J., & Timperley, H. (2007). *The power of feedback*. Review of Educational Research, 77(1), 81-112.
- Jordan, Sally and Mitchell, Tom (2009). *e-Assessment for learning? The potential of short-answer freetext questions with tailored feedback*. British Journal of Educational Technology, 40(2), pp. 371-385.
- Kaap, A. van der (2018), *Formatieve evaluatie, werken aan groei*, Enschede: SLO.
- Karpicke, J. D., & Blunt, J. R. (2011). *Retrieval practice produces more learning than elaborative studying with concept mapping*. Science, 331, 772-775.
- Kleij, F.M. van der (2017), *Leeropbrengst van feedback in computergestuurde toetsen; wat is effectief?*, Arnhem: CITO.
- McKenna, C., & Bull, J. (1999). *Designing effective objective test questions: an introductory workshop*, CAA Centre, Loughborough University.
- Sluijsmans, D., Peeters, A., Jakobs, L., Weijzen, S. (2012). *De kwaliteit van toetsing onder de loep*.
- Sluiter, C., Hemker, B., en Eggen, T. (2018). *Beoordelen van de kwaliteit van toetsen en examens*. Arnhem: Cito.
- Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. Londen: Routledge.

Veldhuijzen, N.H. , Goldebeld, P., & Sanders, P.F., (1993). *Klassieke testtheorie en generaliseerbaarheidstheorie*. In T.J.H.M. Eggen & P. F. Sanders (Eds.) , Psychometrie in de praktijk (pp33-82), Arnhem: Cito.

Vrieze, de J., *Test uzelf en haal een hoger cijfer*, NRC, 25-1-2011.

Wilbrink, B. (2005). *Toetsuragen Schrijven*, Het Spectrum.

(Digitale) toetsing algemeen

http://nl.wikibooks.org/wiki/Onderwijstechnologie/Klastechnologie/Digitaal_toetsen

<https://wij-leren.nl/taxonomie-van-bloom.php>

<https://www.utwente.nl/en/examination/>

Vraagsoorten

<http://www.open.ac.uk/blogs/SallyJordan/?p=215>

Leren en toetsen met serious gaming

<https://www.itpedia.nl/2017/09/01/serious-games-zijn-effectiever-dan-e-learning/>

<https://www.ou.nl/welten-serious-gaming>

<https://www.brainsfirst.com/in-het-nieuws/>

Toetsinhoudelijke blogs

<http://ict-en-onderwijs.blogspot.nl/>

<http://draaijeronassessmentandtesting.wordpress.com>

<http://www.boektweepunt.nl/>

<https://www.optimumassessment.com/actueel/wat-zijn-de-voordelen-van-at-random-toetsen/>

<https://www.optimumassessment.com/actueel/hoe-ga-je-examenfraude-tegen/>

Download whitepaper toetsmatrijs:

<https://www.optimumassessment.com/downloads/toetsmatrijs/>

Bijlage

Checklis

Checklist: stappen naar digitale toetsing

De opmerkingen met een asteriks (*) in de tabel geven de gevolgen aan voor toetsen met vaste versies.

Inrichting vragenbanken

Checklist	Toelichting
1. Wordt er at random of met vaste versies getoetst?	Voordeel van at random is efficiëntie, overzicht en gemak bij het ontwikkelproces, het klaarzetten van de toets, het onderhoud en de beperkte omvang van de database. * Voordeel bij vaste versies is dat de structuur van de vragenbank niet verder hoeft te worden uitgewerkt.
2. Bevat ieder examen een zelfde aantal punten/vragen of mag er variatie in zitten?	Is zowel bij at random als vaste toetsversies van belang. Als je dit niet gelijk houdt in een vragenbank, kan dat oneerlijk zijn.
3. Hoeveel vragen zitten er in één examen?	
4. Hoeveel vragen moeten er in eerste instantie worden ontwikkeld? Hoeveel in tweede instantie?	Afhankelijk van de gestelde normen. Bijvoorbeeld een examen in eerste instantie drie keer kunnen trekken. * Bij vaste versies veiligheidshalve minimaal vier versie-examens.

5. Hoeveel vragen mogen er maximaal in de database actief zijn?	Bij relatief kleine aantallen kandidaten (bijvoorbeeld 150 per jaar) kan het lang duren voordat de vragen voldoende vaak zijn gesteld om analyses uit te kunnen voeren. Dan is het goed om een maximum te stellen aan het aantal actieve vragen in de vragenbank. Denk dan bijvoorbeeld aan twee keer de omvang van het examen.
6. Hoeveel punten zijn er per vraag? 7. Hoeveel vragen per onderwerp zijn er nodig?	Aantal punten per vraag en aantal vragen per onderwerp corresponderen met het gewicht van ieder onderwerp in de toetsmatrijs. Bij een at random examen stel je dit één keer. * <i>Bij een versie-examen bekijk je steeds handmatig of het examen voldoet aan de verdeling.</i>
8. Welke onderwerpen zijn zo belangrijk dat ze in ieder examen aan de orde moeten komen?	Neem deze onderwerpen op als vast onderdeel in de vragenboom. * <i>Of noteer deze handmatig in geval van versie-examens.</i>
9. Welk exameneisen of onderwerpen worden (bijna) altijd samen getoetst? 10. Zijn er exameneisen waar met vervolgvragen het juiste niveau van diepgang wordt bereikt?	Beide aspecten vormen aanleiding tot het opnemen van casussen in de vragenboom en vragenbank. * <i>Bij versie-examens kun je wat losser omgaan met het gebruik van casussen.</i>

11. Wat is de verhouding tussen kennis, inzichtsen toepassingsvragen? (Bloom, 1956)

Staan deze niveaus in de toetsmatrijs en zo ja, zijn deze leidend? Je kunt er beargumenteerd van afwijken, maar dat moet dan wel duidelijk zijn uit de vragenboom. Richtlijn is dat de vragen binnen één vakje van de vragenboom allemaal gelijkwaardig zijn.

12. Op welk beheersingsniveau (kennis, begrip en/of toepassing, zie Bloom 1956) wordt wat getoetst?

• *Bij versie-examens kun je wat losser omgaan met deze niveaus. Je bekijkt dan of de toets in zijn geheel de juiste verhouding (kennis, begrip, toepassing) haalt.*

13. Welke exameneisen toetsen we met open en welke met gesloten vragen?

Per exameneis kijk je of deze het beste met een open of met een gesloten vraag kan worden getoetst. Dit doe je bij een at random examen in verband met gelijkwaardigheid van examens. Als bij een gesloten vraag al snel het antwoord weggegeven wordt, kun je beter voor een open vraag kiezen. Ook als de kandidaat zelf het antwoord moet formuleren.

• *Bij vaste versies kijk je naar het examen als geheel, of deze dezelfde verhouding open/ gesloten heeft als de toetsen die al ontwikkeld zijn. Dit doe je dan handmatig.*

14. Welke richtlijnen hanteren we voor het gebruik van verschillende vraagvormen? 15. Geven we deelpunten of alleen hele punten bij gesloten vragen?	<i>* Dit is zowel bij at random toetsen als toetsen met vaste versies van belang.</i> Hoe zorgen we dan dat er gelijkwaardige vragen in de vragenbank komen? Dit betekent dat er richtlijnen komen over de 'omvang' van de vragen. Hoeveel alternatieven zijn er maximaal per vraagvorm? Hiermee hangt samen of er deelpunten mogen worden gegeven volgens scoringsformules. Deelpunten zorgen voor een gelijkwaardiger scoring, maar vergen uitleg richting 'het veld'.
16. Waar ligt de cesuur?	Worden gokkansen meegewogen? Is er een reden om van de 'normale' norm af te wijken? Gebruiken we een oplopende cesuur, om kandidaten eerst aan het examen te kunnen laten 'wennen'?
17. Hoeveel tijd is er nodig per vraag?	Ook al is het niet nodig dit in het systeem aan te geven, het is wel goed om een inschatting te maken. Zie ook punt 24.
18. Welke taakverdeling is er?	Wie voert vragen in? Wie controleert de invoer?

'Look and feel' van het examen

Checklist	Toelichting
19. Hoe verloopt het inschrijfproces?	Kunnen kandidaten zelf inschrijven en betalen?
20. Welke mails ontvangen kandidaten bij inschrijving?	
21. Zijn er logo's nodig in het examen?	
22. Hoe ziet het start- en afsluitscherm van het examen eruit?	
23. Welke hulpmiddelen kunnen er worden gebruikt?	
24. Wat is de beschikbare tijd voor het examen?	
25. Hoeveel tijd wordt gegeven in geval van aanvragen voor 'extra tijd'?	
26. Welke middelen worden ingezet voor speciale doelgroepen, zoals dyslectici en slechtzienden?	

- 27.** Wordt er een schriftelijk examen uitgedraaid in geval van een storing?

Correctie, uitslag en analyse

Checklist	Toelichting
28. Hoeveel correctoren worden er ingezet per examenwerk?	Wanneer komt er een tweede en/of derde corrector aan te pas? Binnen welke cijferrange?
29. Hoe komt de score tot stand bij open vragen?	Is de score per vraag een gemiddelde van twee correctoren? Of is er een arbiter bij afwijkende scores? Of is het zo dat de tweede corrector feitelijk het cijfer bepaald?
30. Hoeveel dagen heeft een corrector om na te kijken?	Zit daar verschil in voor de eerste, tweede en eventueel derde corrector?
31. Binnen welke termijn heeft de kandidaat uiterlijk de uitslag?	
32. Kan de termijn van de uitslagverstrekking variëren?	Door snelle beschikbaarheid van correctoren kan ook de uitslag snel bekend zijn.

33. In geval van gesloten vragen, mag dan de uitslag onmiddellijk na de examenafname bekend worden?	Kandidaten willen dat vaak graag, maar het kan onrustig zijn tijdens de examenafname wanneer de uitslag onmiddellijk bekend wordt.
34. Hoe wordt feedback gegeven?	Bij de uitslagverstrekking of al eerder? En op welk niveau wordt feedback gegeven, per onderwerp of op vraagniveau?
35. Vindt er inzage plaats en wat is er in te zien?	Is er extra feedback zichtbaar bij de inzage, over goede en/of foute antwoorden? Zijn alle opgaven in te zien of alleen de fouten?
36. Zijn er situaties denkbaar dat uitslagen worden aangepast, individueel of collectief?	Wanneer is dat het geval?
37. Welke analysegegevens gebruiken we voor de vragenanalyse?	Hierin kun je zien welke vragen mogelijk te moeilijk zijn of niet goed zijn. (p-, a-, rit-waarden).
38. Wat doen we met niet goed scorende vragen?	Wordt er periodiek onderhoud gepland? Of vindt de eerste analyse plaats wanneer het statistisch gezien mogelijk is? Lees: wanneer de vragen vaak genoeg gesteld zijn.
39. Wie controleert de kwaliteit van vragen naar aanleiding van klachten?	

40. Wie controleert de instellingen in de toetssoftware naar aanleiding van klachten?	
41. Welke andere gegevens gebruiken we?	Bijvoorbeeld: • Scores per opleider; • Verschil in scoring door correctoren; • Gebruikte tijd per examen; • Slagingspercentages over een bepaalde periode.
42. Nemen we de kandidaten direct na het examen een korte enquête af? Of hebben ze de gelegenheid om direct feedback te geven?	Voorbeeld vragen: • Of ze het examen vinden aansluiten op de opleiding; • Of ze het te moeilijk/makkelijk vinden; • Of ze de inhoud herkenbaar vinden.

De checklist bevat geen informatie over organisatorische processen, omgang met (freelance) inhoudsdeskundigen en communicatie met kandidaten en overige stakeholders. Deze is namelijk erg context gebonden en valt daarmee buiten de scope. De meeste toetssoftware leveranciers zullen daarover wel kunnen adviseren, dus leg de vraag gerust aan hen voor.

Heb je vragen over digitaal toetsen?

Neem dan contact op met Optimum
Assessment:

T: 013 – 528 63 63

E: info@optimumassessment.com



optimum
assessment