

**Evaluatie pilot Naar Betrouwbare
Persoonsgegevens**

Stichting ICTU

AUTEURS:

Ron van Troost
Martine Bos-Horstink
Jolanda Soeting
Willemien Uil-Hoogerwaard

DATUM:

9 januari 2015

Inhoudsopgave

1.	Inleiding	1
2.	Kwalificatiestructuur	2
2.1	Kwalificatiedossiers opstellen	2
2.2	Toetstermen ontwikkelen	3
2.3	Toetsplannen samenstellen	5
2.4	Bevindingen van de kandidaten	6
2.5	Bevindingen van de opleiders	7
3.	Toetsontwikkeling	9
3.1	Deskundigheidsbevordering	9
3.2	Opdrachtverstrekking	10
3.3	Toetsvragen ontwikkelen	12
3.4	Toetsen samenstellen	14
3.5	Praktijkopdrachten ontwikkelen	15
3.6	Toetsvragen en praktijkopdrachten vaststellen	17
3.7	Vastgestelde toetsen en praktijkopdrachten onderhouden	19
4.	Toetsafname en verwerking	20
4.1	Toetsresultaten en analyses	20
4.1.1	Toetsresultaten	20
4.1.2	Toets- en itemanalyse	21
4.1.3	Bestede toetstijd	24
4.2	Resultaten praktijkopdrachten	26
4.3	Algemeen verloop afnames	32
4.4	Bevindingen van de kandidaten	33
4.5	Bevindingen van de beoordelaars	37
4.6	Bevindingen van de opleiders	38
4.7	Ingediende bezwaren	39
5.	Overige aandachtspunten	41
5.1	Implementatie van het toetsbeleid	41
5.2	Communicatie	42
5.3	Examencommissie	43
6.	Overzicht van de verbeter suggesties	45
6.1	Verbetersuggesties ICTU/NVVB	45
6.2	Verbetersuggesties Examenadviesburo	47
6.3	Verbetersuggesties inhoudsdeskundige	48
6.4	Verbetersuggesties toetsdeskundige	48
6.5	Verbetersuggesties inhoudsdeskundige/toetsdeskundige	50

1. Inleiding

Dit evaluatierapport is opgesteld in het kader van het project *Naar Betrouwbare Persoonsgegevens* (NBP). In het kader van dit project is een nieuwe kwalificatiestructuur ontwikkeld voor de uitstroomrichtingen Persoonsinformatievoorziening (PIV) en Identiteitsmanagement (IDM), bedoeld voor medewerkers PIV en IDM. Deze kwalificatiestructuur met bijbehorende examens (toetsen met gesloten en open vragen en praktijkopdrachten) is beproefd in een pilot bij diverse gemeenten in Nederland. Tijdens de pilot zijn bij de volgende acht modules examens ontwikkeld:

Module	Toets	Praktijk-opdracht
<i>MBO-2 Documenten en Verzoeken</i>	T1	-
<i>MBO-2 Migratie</i>		-
<i>MBO-4 Identiteit en Identiteitsdocumenten I</i>	T2	-
<i>MBO-4 Afstamming</i>	T3	-
<i>MBO-4 Nationaliteit</i>		-
<i>MBO-4 Migratie</i>	T4	-
<i>HBO Identiteit en Identiteitsdocumenten II</i>	T5	PO1
		PO2
<i>HBO Landelijke ontwikkelingen</i>	-	PO3

Voor deze evaluatie is onder meer de volgende informatie verzameld:

- resultaten op de examens;
- bevindingen van kandidaten die hebben deelgenomen aan de examens;
- bevindingen van opleiders, vraagconstructeurs, vaststellers en beoordelaars.

Daarnaast beschrijven we het ontwikkelproces van de kwalificatiestructuur en de examens. Op basis van deze informatie, doen we een aantal verbeter suggesties.

In 2014 zijn voor nog zes vervolgmodes toetsen en praktijkopdrachten ontwikkeld. Met ingang van 1 januari 2015 kunnen kandidaten aan deze examens deelnemen. Daarvan zijn dus nog geen resultaten beschikbaar. We nemen in dit evaluatierapport bevindingen mee die (ook) betrekking hebben op het ontwikkeltraject voor de vervolgmodes.

In deze inleiding willen we niet onvermeld laten dat het project in 2013 is bekroond met zowel de *NVVB-award* als de *Prijs voor Examens* van de Nederlandse Vereniging voor Examens.

2. Kwalificatiestructuur

2.1 Kwalificatiedossiers opstellen

We werkten binnen het project NBP toe naar een nieuwe kwalificatiestructuur waarin een duidelijk verband bestaat tussen (een gedeelte van) onderstaande onderdelen:

- werkprocessen;
- kerntaken;
- onderwerpen;
- modules;
- uitstroomrichtingen/specialisaties (PIV/IDM);
- niveaus.

Tijdens de voorbereiding op de eerste ontwikkelbijeenkomst in januari 2013 constateerden we dat de begrippen die werden gebruikt nog niet eenduidig waren en dat een goed fundament voor een dergelijke nieuwe structuur nog ontbrak. De bestaande lijst *Referentiemodellen werkprocessen* van de NVVB leek een goed uitgangspunt, omdat die lijst het hele inhoudelijke gebied van burgerzaken omvat. Daarboven hoort echter nog een structuurlaag te liggen. Een dergelijke structuur is vaak gekoppeld aan functies, maar in dit geval werden functies liever vermeden, omdat het een landelijke structuur moest worden.

Daarom is een structuur ontwikkeld met kerntaken en onderliggende werkprocessen, die nodig zijn om het werk te op de drie verschillende niveaus (MBO-2, MBO-4 en HBO) te kunnen uitvoeren. Vervolgens is binnen die structuur voor elk niveau een kwalificatiedossier ontwikkeld met een beschrijving van de beroepsgroep, een overzicht van de drie kwalificaties binnen de beroepsgroep, een omschrijving van elk van deze drie kwalificaties met bijbehorende werkcontext en niveau. De kwalificatiestructuur voor MBO-2 bestaat uit slechts twee modules, waarvoor één gezamenlijke toets wordt afgenomen. Voor het overgrote deel van de uit te voeren taken is dus minimaal niveau MBO-4 vereist.

Daarnaast zijn van elke kerntaak de onderliggende werkprocessen en een omschrijving opgesteld. Om een koppeling te leggen met het reguliere onderwijs en om duidelijk te maken welke competenties in deze structuur voornamelijk centraal staan, is een relatie gelegd tussen de kerntaken/werkprocessen en de competenties uit het *Competentiemodel Kenniscentra Beroepsopleiding Bedrijfsleven* (powered by SHL) in het MBO. De ontwikkeling van de kwalificatiedossiers is voornamelijk uitgevoerd door de projectgroep NBP, onder begeleiding van Teelen Kennismanagement. Ter inspiratie hebben we in dit proces het MBO-kwalificatiedossier *Juridisch medewerker* gebruikt, aangezien die kwalificatie relaties heeft met de MBO-kwalificaties in de beroepsgroep 'medewerkers burgerzaken'.

Voordelen van de gekozen werkwijze

- De kwalificatiestructuur is opgebouwd op basis van kerntaken en werkprocessen uit het dagelijkse werk van de medewerkers Burgerzaken. Daardoor is er een goede aansluiting op de praktijk.
- Voor de kwalificatiedossiers zijn voorstellen geschreven door enkele personen uit het beroepsveld. Vervolgens zijn de kwalificatiedossiers en de toetstermen door drie commissies van de NVVB gevalideerd en voorgelegd aan de beroepsgroep en aan de opleiders. Dit leverde weinig commentaar op, wat er op wijst dat er een relatief groot draagvlak is.
- Er is een duidelijk onderscheid gemaakt tussen MBO-2, MBO-4 en HBO niveau, los van functies in verschillende gemeenten. Hierdoor is de structuur landelijk bruikbaar, ook al gebruiken gemeenten verschillende functienamen.
- Er is een koppeling met het reguliere onderwijs. Dit maakt de waarde van de opleidingen en examens duidelijk en dit maakt aansluiting op reguliere vervolgoopleidingen mogelijk.
- Toetsing geschiedt geheel onafhankelijk van de opleidingen.

Nadelen van de gekozen werkwijze

- Er is in de kwalificatiedossiers een onderscheid gemaakt tussen de uitstroomrichtingen PIV en IDM. Er blijkt echter op kerntaak- en werkprocesniveau nauwelijks verschil te zijn tussen beide uitstroomrichtingen (één kerntaak op HBO-niveau verschilt). Dit verschil blijkt meer op modulenniveau te liggen. Achteraf gezien twijfelen we daarom aan het nut van het onderscheid tussen de twee uitstroomrichtingen in de kwalificatiedossiers. Dit wordt bevestigd door een onderzoek dat PBLQ uitvoerde onder een beperkt aantal gemeenten. Zij gaven eveneens aan dat dit onderscheid minder relevant is.

Verbeterpunten

- Onderzoeken of de kwalificatiestructuur voor MBO-2 opgeheven kan worden en de twee modules ondergebracht kunnen worden in de MBO-4 kwalificatiestructuur of als losstaande startmodules die buiten de kwalificatiestructuur vallen (eventueel inclusief de BOBZ-module *Hoe werkt de gemeente* en een algemene module *Burgerzaken*)¹.

2.2 Toetstermen ontwikkelen

De drie kwalificatiedossiers vormden een goede basis voor het ontwikkelen van de toetstermen tijdens verschillende ontwikkelbijeenkomsten met diverse experts. De toetstermen konden namelijk nu heel duidelijk gekoppeld worden aan het werk in de dagelijkse praktijk. De vraag die telkens werd gesteld, was 'Wat moeten medewerkers burgerzaken kennen/kunnen om dit werkproces te kunnen uitvoeren?'

¹ De algemene module *Burgerzaken* is verwerkt in de kwalificatiestructuur.

Tijdens verschillende werkbijeenkomsten zijn bestaande toetstermen vanuit de examenstructuur van NVVB hergebruikt, opnieuw ingedeeld en gekoppeld aan de werkprocessen in de nieuwe structuur. Daarnaast zijn er nieuwe toetstermen op de drie niveaus toegevoegd. Dit geheel van toetstermen is geclusterd in modules (2 op MBO-2 niveau, 14 op MBO-4 niveau en 8 op HBO-niveau).

De toetstermen zijn ontwikkeld door inhoudsdeskundigen, onder begeleiding van Teelen Kennismanagement. De inhoudsdeskundigen zijn werkzaam in het vakgebied van PIV en IDM. Zij werken bij verschillende organisaties, voornamelijk gemeenten en overheidsinstanties. Bij het ontwikkelen van toetstermen waren veel verschillende inhoudsdeskundigen betrokken, elk met hun eigen deskundigheid. Door een combinatie van verschillende specialismen is een compleet beeld van het vakgebied ontstaan. Daarnaast ontstond, door de betrokkenheid van een grote groep mensen uit het vakgebied, een breed gedragen basis voor de examens.

Tijdens het ontwikkelen van de toetsvragen en praktijkopdrachten bleken bepaalde toetstermen niet concreet en duidelijk (genoeg) omschreven te zijn. Bovendien lijken toetstermen soms te veel op elkaar, waardoor ook de vragen die erover worden ontwikkeld, veel van elkaar weg hebben. Daarnaast zijn er tijdens de pilotperiode ook wijzigingen in wet- en regelgeving geweest, die invloed hadden op de toetstermen. Het is in dit stadium van het project heel normaal dat geformuleerde toetstermen moeten worden gewijzigd. Alle opmerkingen over gewenste wijzigingen in de toetstermen zijn daarom opgenomen in een verzameldocument. Volgens het toetsbeleidsplan project NBP worden op één moment noodzakelijke wijzigingen in toetstermen doorgevoerd, namelijk na evaluatie van de pilotperiode. Een deel van de gewenste wijzigingen is echter al eerder doorgevoerd, omdat het wenselijk dat de opleidingen werden aangepast.

Omdat de kwalificatiedossiers nog onbekend en niet uitgekristalliseerd waren tijdens de ontwikkeling van de toetstermen, was het voor de inhoudsdeskundigen soms lastig om de koppeling te maken tussen toetstermen en werkprocessen. Het was beter geweest als de kwalificatiedossiers ruim voor de eerste ontwikkelbijeenkomst beschikbaar waren.

Tijdens de ontwikkelsessies voor toetstermen is steeds de opdracht gegeven om toetstermen toe te wijzen aan een of beide uitstroomrichtingen (PIV en/of IDM). In de toetsplannen is het onderscheid tussen PIV/IDM achterwege gelaten. Aangezien er uiteindelijk alleen in de diplomalijnen onderscheid op modulenniveau wordt gemaakt, adviseren we om dit onderscheid op toetstermniveau niet alsnog zichtbaar te maken.

Verbetersuggesties

- Verwerken van de resterende wijzigingen in het verzameldocument met toetstermen.
- Regelmatig onderhoud op toetstermen inplannen. Dit zou minimaal één keer per jaar moeten plaatsvinden en daarnaast waar nodig incidenteel, zoals bij de invoering van LO3.9 per 31 januari 2015. Bij elke wijziging moeten natuurlijk ook alle vastgestelde vragen en opdrachten (op het gebied van de wijziging) worden nagelopen.
- Toetstermen beoordelen op duidelijkheid, op (gedeeltelijke) overlap en op de mogelijkheid om gecombineerd te worden met inhoudelijk zeer gerelateerde andere toetstermen.
- Minder nadruk leggen op het onderscheid tussen PIV en IDM en in de diplomalijnen rekening houden met vrije keuzes in modules, die aansluiten bij de taken van een medewerker Burgerzaken binnen een grote gemeente (meer gericht op specialisaties) en binnen een kleine gemeente (meer gericht op allround inzetbaarheid).

2.3 Toetsplannen samenstellen

De toetsplannen zijn ontwikkeld door inhoudsdeskundigen, onder begeleiding van Teelen Kennismanagement. Bij de weging van de toetstermen en het bepalen van de toetsvormen zijn verschillende inhoudsdeskundigen betrokken geweest. In eerste instantie was het uitgangspunt om over elke toetsterm minimaal één vraag te ontwikkelen. Gezien de grote hoeveelheid toetstermen bleek dit, zowel financieel als praktisch, niet haalbaar. Er is vervolgens een prioritering in de toetstermen aangebracht, waardoor er clusters van toetstermen ontstonden waaruit één vraag kan worden gesteld.

De toegestane tijd, cesuur, scoringsregels, hulpmiddelen en instructies zijn door ICTU, in overleg met een toetsdeskundige, vastgesteld. Voor de cesuur geldt dat op het toetsplan staat vermeld 55% op basis van de kansscore. In de praktijk leidt dit voor de pilotmodules tot een cesuur tussen 63 en 71%. Voor de praktijkopdrachten geldt een vaste cesuur van 70%.

Tijdens de pilotperiode zijn de toetsplannen op verschillende momenten aangepast. Deze wijzigingen zijn in rood in de bestaande toetsplannen verwerkt en verstrekt aan alle betrokkenen, waarbij achter elke toetsterm de wijzigingsdatum is vermeld. In 2014 zijn alle toetstermen van de acht pilotmodules en de zes vervolgmodes nagelopen, bijvoorbeeld op gewijzigde wetgeving, en indien nodig geactualiseerd. Tijdens het ontwikkelen van vragen of opdrachten bij de pilotmodules bleek dat sommige toetstermen anders moesten worden geformuleerd. Deze opmerkingen zijn verzameld en meegenomen bij de actualisering. Er zijn tussentijds enkele toetstermen vervallen, omdat zij niet langer relevant of niet juist waren. In enkele gevallen zijn nieuwe toetstermen toegevoegd.

In deze ronde hebben opleiders de gelegenheid gehad om hun commentaar te geven. Ook hun commentaar is verwerkt. Daarna zijn de examenvragen op basis van de aangepaste toetstermen verbeterd of geactualiseerd.

Tijdens het ontwikkelen van toetsvragen voor de nieuwe modules in 2014 is voor minimaal één module bepaald dat de toets in twee delen moet worden gesplitst: een eerste deel waarbij **geen** hulpmiddelen mogen worden gebruikt en een tweede deel waarbij dit **wel** mag. Omdat vanaf januari 2015 de hulpmiddelen alleen nog maar digitaal worden toegestaan, kan dit via TestVision Online worden gecontroleerd.

Verbetersuggestie

- Beoordelen of meer modules moeten worden gesplitst in twee delen (zonder en met hulpmiddelen) en dit vervolgens realiseren in TestVision Online.

2.4 Bevindingen van de kandidaten

Ruim 60% van de 147 kandidaten die de enquête hebben ingevuld, vindt dat de examenstof voldoende aansluit op de praktijk. Degenen die dat niet vinden, geven onder andere de volgende argumenten (van meest genoemd naar minder genoemd):

- Het examen sloot niet goed aan bij de lesstof of de inhoud van de lessen. Dit wil eigenlijk zeggen dat de lesstof of de inhoud van de lessen niet aansloot bij de toetstermen. Zo werd er volgens verschillende kandidaten in de les te weinig aandacht gegeven aan persoonslijsten en het verwerken van brondocumenten. Ook geven enkele cursisten aan dat echtheidskenmerken juist wel (uitgebreid) behandeld zijn, maar dat de toets daar weinig vragen over bevatte.
- In de praktijk maak ik gebruik van hulpbronnen (handboeken, wetten, kennisbank, landeninformatie, collega's) of heb ik meer tijd om iets uit te zoeken.

Op de vraag of het examen aansloot op de gevolgde opleiding geeft vervolgens ruim 50% van de kandidaten aan dat dit zo is. Degenen die dat niet vinden, geven onder andere de volgende argumenten (van meest genoemd naar minder genoemd):

- Verschillende onderwerpen zijn niet of te weinig aan bod gekomen in de opleiding. Ook hier worden onderwerpen als persoonslijsten en het verwerken van brondocumenten genoemd. Het onderwerp echtheidskenmerken kwam juist weer te veel aan bod in de opleiding.
- Enkele kandidaten geven echter aan dat ze bepaalde vragen op basis van de toetstermen niet hadden verwacht. Dit gaat dus meer over de aansluiting examen-toetstermen en niet zozeer over de aansluiting examen-opleiding. Ook vinden enkele kandidaten dat er te veel vragen (over hetzelfde onderwerp) worden gesteld. Ook het totaal aantal vragen wordt als (te) veel ervaren.

- De opleiding was te kort voor een goede voorbereiding op de toetsen. Het examen bevat vragen over uitzonderingssituaties, bijvoorbeeld geprivilegieerden, geboortekte van een vondeling, faciliteitenpaspoort.
- Om de toets goed te kunnen maken, heb je (ook) voldoende werkervaring nodig.

Verbetersuggesties

- Opleiders attenderen op de door cursisten gesignaleerde discrepantie tussen de inhoud van de opleiding en de toetstermen.
- Toetstermen beoordelen op dubbeling en op de mogelijkheid om gecombineerd te worden met inhoudelijk zeer gerelateerde andere toetstermen, zodat bepaalde onderwerpen minder of juist meer aan bod komen. Hierbij kan ook worden gekeken of het totale aantal vragen per toets kan worden verlaagd.

2.5 Bevindingen van de opleiders

Vier opleiders hebben een evaluatieformulier ingevuld en drie opleiders waren vertegenwoordigd bij een evaluatiebijeenkomst op 28 november 2014. De opleiders hebben de volgende opmerkingen over de kwalificatiestructuur gemaakt:

- Men vindt de toetsplannen nog onvoldoende duidelijk. Vanuit het project NBP is in de loop van de proefperiode twee keer aangeboden de toetsplannen bij aanvang van de pilot te bespreken met docenten die lesmateriaal zouden ontwikkelen en lessen zouden verzorgen. Docenten en opleiders gaven destijds aan dat daar geen behoefte aan was. Wel vindt men dus dat de toetsplannen op sommige punten toelichting nodig hebben.
- Het project NBP heeft voldoende tussenevaluatiemomenten belegd.
- De stap van MBO-2 naar MBO-4 is te groot. Bij MBO-2 vindt men de structuur te smal. Men verwijst daarbij naar de BOBZ waarin ook een module *Hoe werkt de gemeente* is opgenomen en een algemene module *Burgerzaken*.
- De opleiders constateren dat de praktijk bij gemeenten kan afwijken van de theorie die de basis vormt van de kwalificatiestructuur.
- De opleiders geven aan behoefte te hebben aan duidelijkheid over het vervolg van de ontwikkeling van de kwalificatiestructuur (de volgende tien modules en de voucherregeling). Daarbij vraagt men ook aandacht voor het uitfaseren van de huidige diplomalijnen BOBZ/BOP en BABZ/MLOBZ.
- De opleiders geven aan meer ervaring te hebben met de modules en dus betere uitspraken te kunnen doen over het aantal studiebelastingsuren per module. Zij willen deze inspanning echter laten afhangen van een uitspraak over het vervolg van het verder vullen van de modulestructuur.
- De vijver waarin gevist kan worden voor docenten wordt steeds kleiner.

Verbetersuggesties

- Nader onderzoeken wat de opleiders exact onduidelijk vinden aan de toetsplannen en waar mogelijk de toetsplannen uitbreiden en verduidelijken.
- Onderzoeken of de kwalificatiestructuur voor MBO-2 opgeheven kan worden en de twee modules ondergebracht kunnen worden in de MBO-4 kwalificatiestructuur of als losstaande startmodules die buiten de kwalificatiestructuur vallen (eventueel inclusief de BOBZ-module *Hoe werkt de gemeente* en een algemene module *Burgerzaken*).
- Duidelijkheid verstrekken over de verdere ontwikkeling van de kwalificatiestructuur en over de uitfasering van de huidige diplomalijnen BOBZ en BABZ/MLOBZ.

3. Toetsontwikkeling

3.1 Deskundigheidsbevordering

Gedurende de pilot heeft Teelen Kennismanagement verschillende bijeenkomsten in het kader van deskundigheidsbevordering verzorgd voor vraagconstructeurs en beoordelaars:

- een (basis)cursus *Vraagconstructie* (5 keer);
- een instructiebijeenkomst voor beoordelaars (1 keer);
- een verdiepingscursus *Vraagconstructie* (1 keer);
- een startbijeenkomst *Praktijkopdrachten ontwikkelen* (2 keer).

In de basiscursus *Vraagconstructie* besteden we veel aandacht aan het ontwikkelen van kwalitatief goede toetsvragen bij de toetstermen uit de kwalificatiestructuur, maar ook aan de werkinstructies en (sinds 2014) aan het werken met TestVision Online. Uit de evaluatieformulieren blijkt dat de deelnemers aan de (basis)cursus *Vraagconstructie* (zeer) positief zijn over de training, de toegevoegde waarde en de praktische toepasbaarheid ervan. Ook het van en met elkaar leren in een groep wordt als positief ervaren. Wel is het programma van de cursus aan de volle kant en hadden sommige deelnemers van tevoren beter geïnformeerd willen worden over wat zij kunnen verwachten.

In de instructiebijeenkomst voor beoordelaars is aandacht voor het belang van een goede beoordeling en het gevaar van beoordelaarseffecten en oefenen de deelnemers met het beoordelen van open vragen en praktijkopdrachten. Ook de deelnemers aan deze instructiebijeenkomst zijn in de evaluatieformulieren positief over de training en de toepasbaarheid ervan in de praktijk.

In de verdiepingscursus *Vraagconstructie* is onder meer aandacht voor het ontwikkelen van toetsvragen op het juiste taxonomieniveau en voor het gebruiken van de verschillende vraagvormen in TestVision Online. Deze training is slechts door vier deelnemers gevolgd (terwijl er negen deelnemers waren uitgenodigd). De aanwezige deelnemers waren na afloop enthousiast over de training en de toegevoegde waarde ervan.

De startbijeenkomst *Praktijkopdrachten ontwikkelen* bestaat uit een korte inleiding, waarna groepjes onder begeleiding starten met het ontwikkelen van de praktijkopdrachten. Deze werkwijze leidt ertoe dat de constructeurs direct al een stuk in de goede richting op weg zijn na afloop van de bijeenkomst. In dezelfde groepjes worden de praktijkopdrachten vervolgens verder uitgewerkt.

Er zijn **tijdens de pilot** 35 personen actief geweest als vraagconstructeur. Deze hebben een enquête ingevuld over het proces van vragen ontwikkelen en vaststellen. Aan de pilot en vervolgmodes samen hebben 57² vraagconstructeurs meegewerkt. In totaal hebben 22 vraagconstructeurs (63%) de enquête ingevuld. We hebben de vraagconstructeurs in deze enquête nogmaals gevraagd naar de mate waarin ze het in de cursus geleerde gebruikt hebben bij het ontwikkelen van toetsen en praktijkopdrachten. Negentig procent van de respondenten geeft aan veel aan het geleerde te hebben (gehad). Wel geven enkele constructeurs aan dat meer aandacht voor het verschil tussen de niveaus (MBO-2, MBO-4 en HBO) gewenst is en ook zouden sommige deelnemers (nog) meer willen oefenen.

In enkele gevallen zijn vraagconstructeurs ingezet die vooraf geen basiscursus hadden gevolgd. In de meeste gevallen betekende dit extra werk in het proces van vragen redigeren en vaststellen, omdat men onvoldoende op de hoogte was van de kwaliteitseisen die worden gesteld aan goede vragen, maar ook van de werkinstructies en afspraken.

Verbetersuggesties

- Deelnemers van tevoren beter informeren over doel, inhoud en belang van de cursussen.
- Toekomstige vraagconstructeurs uitleggen hoe belangrijk het voor de kwaliteit van de examenvragen en de efficiëntie van het ontwikkeltraject is om de basiscursus *Vraagconstructie* te volgen indien men nog niet over de vereiste kennis en vaardigheden beschikt.
- In de basiscursus *Vraagconstructie* aandacht besteden aan het verschil tussen de niveaus door voorbeelden van toetstermen over hetzelfde onderwerp op verschillende niveaus te bespreken.
- In de uitnodiging beter duidelijk maken dat deelname niet vrijblijvend is, bijvoorbeeld door deelnemers de uitnodiging te laten bevestigen en erop aandringen dat deelnemers zich tijdig afmelden als ze niet kunnen.

3.2 Opdrachtverstrekking

De vraagconstructeurs krijgen hun opdrachtverstrekking na telefonisch contact via e-mail toegestuurd. De opdrachtverstrekking bestaat uit een overzicht van geselecteerde toetstermen en het aantal vragen dat bij elke toetsterm moet worden ontwikkeld. Daarnaast bevat de e-mail de volgende informatie:

- de werkinstructie voor het ontwikkelen van goede toetsvragen en de instructie voor het werken in TestVision Online;
- een verwijzing naar toetsplannen op de website van NVVB;
- een declaratieformulier en geheimhoudingsverklaring.

² We hebben de enquête niet naar alle vraagconstructeurs gestuurd, omdat een groot deel van de nieuwe constructeurs op dat moment nog maar weinig vragen hadden ontwikkeld en nog geen enkele vaststellingsvergadering hadden bijgewoond.

Alle respondenten geven aan dat ze de opdrachtverstrekking en de werkinstructies voldoende duidelijk vinden. Er worden een aantal opmerkingen gemaakt ten aanzien van de werkinstructie, zoals de behoefte aan een instructie voor het invoegen van een afbeelding (dit wordt gedemonstreerd in de basiscursus), aan een overzicht met definities en/of een schema hoe dingen samenhangen en de kanttekening dat de werkinstructie niet altijd actueel is.

We hebben tijdens het proces van het redigeren en vaststellen van de toetsvragen diverse keren geconstateerd dat de opdrachtverstrekking niet altijd optimaal is. Zo zijn er in een enkel geval toetsvragen ontwikkeld bij toetstermen die in de (vervolgens) aangepaste toetsplannen waren komen te vervallen. Bij sommige toetstermen zijn te veel vragen ontwikkeld, bijvoorbeeld doordat voor deze toetstermen relatief veel vragen uit de NVVB database konden worden geconverteerd.³ Bij andere toetstermen hebben we nog te weinig of zelfs helemaal geen vragen, bijvoorbeeld omdat tijdens vaststellingsvergaderingen bleek dat het niet mogelijk was om passende vragen te ontwikkelen. Ook mochten enkele vraagconstructeurs zelf uit een set toetstermen kiezen bij welke toetsterm ze een of meer vragen ontwikkelden, waardoor niet altijd voor elke afzonderlijke toetsterm binnen een cluster toetsvragen ontwikkeld zijn. Bij het uitzetten van dezelfde toetsterm bij verschillende vraagconstructeurs bestaat bovendien de kans dat de constructeurs (onafhankelijk van elkaar) vragen ontwikkelen die te veel op elkaar lijken.

Daarnaast hebben we in de praktijk gemerkt dat de vraagconstructeurs soms moeite hebben met het ontwikkelen op het juiste taxonomie-niveau (Kennis, Begrip, Reproductieve vaardigheden of Productieve vaardigheden) en met het correct opvolgen van de werkinstructies ten aanzien van de vraagnamen, de instructies bij de verschillende vraagvormen en de correcte instellingen van de eigenschappen van een vraag (vaste volgorde, kenmerken).

Verbetersuggesties

- Voorafgaand aan de opdrachtverstrekking een actueel overzicht maken met de benodigde toetsvragen per (cluster van) toetsterm(en).
- Bij de opdrachtverstrekking de toetstermen verdelen over zo min mogelijk verschillende vraagconstructeurs en de vraagconstructeurs geen eigen keuzen laten maken.
- Het opstellen van de opdrachtverstrekkingen door Teelen Kennismanagement laten uitvoeren.
- In de opdrachtverstrekking het belang van het ontwikkelen op het juiste taxonomie-niveau benadrukken (inclusief korte toelichting op de niveaus met voorbeelden).

³ Omdat er inmiddels voor is gekozen te gaan werken met at random samengestelde toetsen, worden alle vastgestelde toetsvragen gebruikt, ook bij een overschot.

- In de basiscursus *Vraagconstructie* meer oefenen met het ontwikkelen van vragen op het juiste taxonomie-niveau en het maken van vragen op RV- én B-niveau bij één toetsterm.
- De werkinstructies actualiseren en uitbreiden met de genoemde informatie.
- In de werkinstructie de informatie over de naamgeving van vragen en de iteminstellingen (vaste volgorde alternatieven, taxonomie-niveau, wel of geen gedeeltelijke scoring) meer benadrukken/uitlichten.

3.3 Toetsvragen ontwikkelen

Bij de eerste modules hebben de vraagconstructeurs de vragen ontwikkeld in een speciaal format in MS Word, waarna de vragen na vaststelling in TestVision Online werden geïmporteerd. Inmiddels werken de constructeurs direct in TestVision Online. Dit scheelt werk en stimuleert de constructeurs mogelijk meer gebruik te maken van verschillende vraagvormen. De meeste respondenten (95%) geven aan dat ze het werken in TestVision Online altijd of meestal plezierig vinden. Eén respondent (5%) meldt dit nooit het geval te vinden.

De vraagconstructeurs werken in een eigen ontwikkelmap in TestVision Online en kunnen dus geen andere vragen zien en/of bewerken. Dit betekent dat ze geen inspiratie kunnen opdoen door het bekijken van andere vragen. En bij het ontwikkelen van (parallele) vragen bij toetstermen waarbij al vragen beschikbaar zijn, kunnen ze niet eenvoudig kijken welke vragen er al zijn. Het is echter ook niet wenselijk om alle constructeurs toegang te geven tot alle vragen. Elke constructeur heeft wel altijd de beschikking over een voorbeeldvraag van elke vraagvorm.

Tijdens de pilotfase heeft een toetsdeskundige van Teelen Kennismanagement een groot aantal parallele vragen ontwikkeld. Dit heeft als voordeel dat deze vragen alleen inhoudelijk en niet toetstechnisch beoordeeld hoeven te worden.

Het grootste struikelblok bij het ontwikkelen van de toetsvragen is de formulering van de toetstermen. Men vindt de formulering te abstract of te breed, waardoor onvoldoende duidelijk is wat ermee wordt bedoeld. Tijdens vaststellingsvergaderingen worden nog veel discussies gevoerd over de strekking van de toetstermen. Mede daardoor is er behoefte aan een toelichting op de toetstermen of aan een overzicht van bij de toetstermen passende leerstof, documenten en/of procedures.

De vraagconstructeurs (17 respondenten) besteden gemiddeld 48 minuten aan het ontwikkelen van een gesloten toetsvraag. Er is één constructeur die aangeeft slechts 5 (!) minuten te besteden en twee constructeurs besteden maar liefst 120 minuten aan het ontwikkelen van een gesloten toetsvraag. Aan het ontwikkelen van open vragen besteden de constructeurs (10 respondenten) gemiddeld 78 minuten.

De minimaal bestede tijd bedraagt 20 minuten en één constructeur besteedt 180 minuten aan het ontwikkelen van een open vraag. De vraagconstructeurs geven hun eigen vragen bij aanlevering voor redactie en correctie gemiddeld een 7.

Tot en met begin november is (per vraagvorm) het volgende aantal toetsvragen voor de drie niveaus ontwikkeld:

Vraagvorm	MBO-2	MBO-4	HBO
Een-uit-meervraag	215	701	109
Meer-uit-meervraag	81	222	67
Rangschikvraag	1	3	0
Matchvraag gelijk	0	4	2
Matchvraag ongelijk	2	19	1
Hotspotvraag	0	4	0
Hotmatchvraag	0	0	1
Invulvraag	1	13	0
Invulvraag numeriek	2	13	2
Invulvraag meervoudig	0	1	1
Open vraag	0	28	43
Totaal	302	1008	226

Tabel 3.1: Aantal toetsvragen per vraagvorm en per niveau

Uit dit overzicht blijkt dat al redelijk veel gebruik wordt gemaakt van de meer-uit-meervraag, maar dat de andere vraagvormen nog slechts voorzichtig worden gebruikt. Open vragen worden in beperkte mate gebruikt en alleen in de modules van de pilot, zodat er zo min mogelijk achteraf met beoordelaars gewerkt hoeft te worden. Met de introductie van de meervoudige invulvraag in TestVision Online 5 is het mogelijk om een deel van de eerder ontwikkelde open vragen om te zetten naar een meervoudige invulvraag die automatisch kan worden beoordeeld.

Voor de meer-uit-meervraag geldt dat standaard de optie *gedeeltelijk scoren* aan staat, waardoor de kandidaat een deel van de score krijgt voor een gedeeltelijk goed antwoord. Bij het vaststellen van de vraag wordt beoordeeld of deze optie moet worden uitgezet. Dit maakt de vraag moeilijker en verlaagt de kansscore van de vraag. Uitgangspunt is dat *gedeeltelijk scoren* alleen wordt uitgezet als de vraag absoluut volledig correct moet worden beantwoord. Voor het niveau MBO-2 geldt dit voor 5% van de meer-uit-meervragen en voor de niveaus MBO-4 en HBO voor 15%.

Tijdens de pilot zijn ook enkele vragen met videofragmenten ontwikkeld. Vanaf februari 2015 wordt het in TestVision Online mogelijk om de vragen met videofragmenten daadwerkelijk in te zetten.

Verbetersuggesties

- De (te) abstracte of complexe toetstermen voorzien van een uitgebreidere toelichting of enkele voorbeeldvragen.
- Een overzicht maken van bij de toetstermen passende documenten, procedures of beschikbare leerstof en achtergrondinformatie en dit ter beschikking stellen aan de vraagconstructeurs.
- De vraagconstructeurs verder professionaliseren in het ontwikkelen van andere vraagvormen. De verdiepingscursus is hiervoor een zeer geschikte mogelijkheid.
- Bij het ontwikkelen van parallelle of extra vragen bij een toetsterm de oorspronkelijke vragen als PDF-document bij de opdracht verstrekken.
- De open vragen die ervoor geschikt zijn omzetten naar meervoudige invulvragen.
- Per situatie afwegen of parallelle vragen het best door een vraag constructeur of door een toetsdeskundige kunnen worden ontwikkeld.
- Bepalen of het zinvol is om bij bepaalde toetstermen (op toepassingsniveau) vragen met videofragmenten te laten ontwikkelen.

3.4 Toetsen samenstellen

In de pilot zijn per module drie of vier toetsvarianten ontwikkeld. Het werken met toetsvarianten in plaats van at random samenstellen, had als doel voldoende afnamegegevens te verzamelen om de kwaliteit van de toetsen en de vragen te kunnen beoordelen (zie paragraaf 4.1). Omdat het aantal kandidaten lager bleek dan van tevoren was ingeschat, zijn uiteindelijk slechts twee toetsvarianten per module daadwerkelijk afgenomen.

Het goed verdelen van de toetsvragen over de toetsversies is een lastige klus die relatief veel tijd in beslag neemt. Bovendien is het bij vast samengestelde toetsversies relatief veel werk om een vraag die is vervallen te vervangen en om nieuw ontwikkelde vragen toe te voegen. Daarnaast raken de vragen in vast samengestelde toetsen meestal sneller bekend onder de kandidaten dan bij at random samengestelde toetsen. Ten slotte blijkt dat ook bij twee toetsvarianten nog steeds onvoldoende resultaten beschikbaar zijn voor een goede toets en itemanalyse (zie paragraaf 4.1). Om deze redenen is er inmiddels voor gekozen om de toetsen at random samen te stellen. Dit betekent dat elke kandidaat een andere combinatie van toetsvragen krijgt voorgelegd, vanzelfsprekend wel conform de toetsmatrijs uit het toetsplan. Omdat er hierdoor nog maar één toetsversie per module is, hoeft het Examenadviesburo kandidaten niet (handmatig) meer over meer versies te verdelen.

Bij at random samengestelde toetsen kunnen minder uitgebreide toetsanalyses worden uitgevoerd, omdat de toets voor elke kandidaat anders is. Het onderscheidend vermogen van vragen kan alleen bij vaste toetsen worden bepaald, omdat de vragen dan in een vaste samenstelling met andere vragen aan dezelfde kandidaten worden voorgelegd. De kwaliteit van het item kan dan gerelateerd worden aan de andere items. Bij at random samengestelde toetsen kan alleen de moeilijkheid van de items worden onderzocht. Ook hiervoor geldt dat **een item** minstens 50 en liefst minstens 100 keer moet zijn afgenomen.

TestVision Online kan een cesuur berekenen op basis van de kansscore. Op dit moment wordt het cijfer van de kandidaat echter vastgesteld door EMO (het systeem van het Examenadviesburo) op basis van een vaste cesuur en een vaste maximale score. Dit systeem is niet in staat rekening te houden met de kansscore. Om die reden kunnen we op dit moment nog niet werken met een individuele cesuur per kandidaat, maar bepalen we voor elke module een vaste cesuur die zo dicht mogelijk ligt bij de cesuur met de laagste mogelijke kansscore. Dit betekent dat voor de meeste kandidaten de cesuur (iets) lager uitvalt dan wanneer de cesuur op basis van de kansscore van hun individuele toets zou zijn bepaald. Deze vaste cesuur wordt vanaf 2015 vermeld op de toetsplannen.

Omdat het cijfer in EMO wordt berekend op basis van een vaste maximale score, moet er bij het at random samenstellen van de toetsen voor worden gezorgd dat binnen een set open vragen waaruit wordt geselecteerd voor een specifieke toetssterm elke open vraag altijd hetzelfde aantal punten waard is.

Verbetersuggestie

- De mogelijkheid onderzoeken om het cijfer vanuit TestVision Online in EMO over te nemen, zodat bij de cesuur rekening kan worden gehouden met de kansscore en er geen vaste maximale score meer noodzakelijk is.

3.5 Praktijkopdrachten ontwikkelen

Praktijkopdrachten worden ontwikkeld in MS Word en pas nadat ze volledig zijn vastgesteld ingevoerd in TestVision Online voor afdruk.

Praktijkopdrachten worden meestal gezamenlijk ontwikkeld door twee vraagconstructeurs. Veertien vraagconstructeurs hebben praktijkopdrachten ontwikkeld. De 11 respondenten die praktijkopdrachten hebben ontwikkeld, besteedden daar gemiddeld ruim 11 uur aan. De tijdsbesteding loopt uiteen van 2 uur tot 30 uur. Dit verschil in tijdsbesteding zal mogelijk samenhangen met de rol in het duo en met het onderscheid tussen unieke en parallele praktijkopdrachten.

Elke praktijkopdracht bestaat uit een instructie voor een kandidaat (eventueel met bijlagen) en een beoordelingsformulier. Soms zijn er aanvullende documenten met achtergrondinformatie opgenomen voor de beoordelaars.

Er is een MS Word-sjabloon gebruikt, waardoor de praktijkopdracht steeds dezelfde opbouw heeft en (waar mogelijk) dezelfde formuleringen zijn gebruikt. Uiteraard kostte dit bij de ontwikkeling van de eerste praktijkopdrachten meer tijd en gaat dit sneller nu er meer uitgewerkte praktijkopdrachten beschikbaar zijn.

In een bijeenkomst worden de eerste ideeën voor de inhoud en vorm van een praktijkopdracht ontwikkeld, waarbij een toetsdeskundige van Teelen Kennismanagement meedenkt over de haalbaarheid en geschiktheid. De praktijkopdrachten komen vervolgens via meerdere ontwikkelslagen tot stand, waarbij de toetsdeskundigen de aangeleverde versies redigeren en verbeteruggesties geven.

De constructeurs vonden het in het begin lastig dat zij geen beeld hadden van de eindproducten die zij moesten opleveren. Het formuleren van objectieve beoordelingscriteria bleek voor hen soms moeilijk. De toetsdeskundigen hebben daarom veel voorbeelden gegeven, meegedacht en behoorlijk aan de beoordelingsformulieren gesleuteld. Door de inmiddels beschikbare voorbeelden zijn de constructeurs nu beter in staat om zelf geschikte beoordelingscriteria te formuleren. Zij ervaren het ontwikkelen van praktijkopdrachten nu minder als een enorme opgave, wat in het begin zeker wel zo was.

Tijdens de pilot constateerden we dat het bij sommige praktijkopdrachten mogelijk is om eenvoudig een parallelle versie te ontwikkelen, waarbij de beoordelingscriteria grotendeels of zelfs volledig overeenkomen. Bij een tweede ontwikkelronde hebben we tijdens de eerste ontwikkelbijeenkomst de mogelijkheden voor parallelle versies met de constructeurs besproken. Na het ontwikkelen van de eerste unieke versies blijkt het soms toch niet mogelijk om ook een parallelle versie te ontwikkelen. Verder blijkt het bij sommige toetstermen heel lastig om concrete en voldoende uitdagende praktijkopdrachten te ontwikkelen.

Voordat de eerste afnames plaatsvonden, is besloten dat de administratie rond de praktijkopdrachten grotendeels in TestVision Online plaatsvindt. De instructie van praktijkopdrachten die thuis of op de werkplek worden uitgewerkt, worden via EMO aan de kandidaten verstrekt. Kandidaten moeten hun uitwerking wel in TestVision Online uploaden. ICTU overweegt in de toekomst Moodle te gebruiken voor het afnemen van de praktijkopdrachten thuis of op de werkplek, omdat dit mogelijk het inschrijfproces en het toewijzen van het beoordelen eenvoudiger maakt.

Het is goed om dit nader te onderzoeken, omdat in TestVision Online 6 veel nieuwe functionaliteiten bij het toewijzen van beoordelaars beschikbaar komen.

Verbetersuggesties

- Het MS Word sjabloon actualiseren en opsplitsen in twee varianten: één voor de praktijkopdrachten op examenlocaties en één voor praktijkopdrachten die thuis of op de werkplek worden gemaakt.
- Vooraf beter nagaan of het mogelijk is om bij (set van) toetstermen geschikte praktijkopdrachten en eventuele parallelle versies van praktijkopdrachten te ontwikkelen.
- Onderzoeken wat de beste manier is om de praktijkopdrachten aan te bieden aan de kandidaten en het proces van beoordelen te optimaliseren.

3.6 Toetsvragen en praktijkopdrachten vaststellen

Alle toetsvragen en praktijkopdrachten zijn stuk voor stuk vastgesteld tijdens vaststellingsvergaderingen. In eerste instantie waren bij een vaststellingsvergadering twee inhoudelijke vaststellers (inhoudsdeskundigen), de vraagconstructeur en een toetsdeskundige van Teelen Kennismanagement aanwezig. Na verloop van tijd is er voor gekozen om bij het vaststellen van toetsvragen slechts één inhoudelijke vaststeller aanwezig te laten zijn. Dit scheelt veel capaciteit en is beter in te plannen. De aanwezigheid van twee inhoudsdeskundigen (vaststeller en vraagconstructeur) bleek ruim voldoende. Bij het vaststellen van de praktijkopdrachten zijn nog altijd twee vaststellers aanwezig (met uitzondering van enkele praktijkopdrachten die op zeer korte termijn moesten worden vastgesteld). In die gevallen heeft de tweede vaststeller wel voorafgaand aan de vergadering schriftelijk feedback gegeven.

Voorafgaand aan een vergadering beoordeelt de toetsdeskundige de vragen en praktijkopdrachten toetstechnisch. De bevindingen worden teruggekoppeld aan de vraagconstructeur en de vaststeller, indien mogelijk uiterlijk vijf werkdagen voordat de vergadering plaatsvindt. De inhoudelijke vaststeller geeft zijn opmerkingen tijdens de vergadering zelf. Dit zou ook van tevoren schriftelijk kunnen worden aangeleverd. Sommige vaststellers doen dit al.

Bij het beoordelen van de toetsvragen constateren we een duidelijke toename in de kwaliteit van de ontwikkelde vragen. Veel vraagconstructeurs worden steeds vaardiger in het ontwikkelen van goede toetsvragen die goed aansluiten bij (het taxonomie-niveau) van de toetsterm. Als bij het redigeren blijkt dat een toetsvraag onvoldoende aansluit op de toetsterm, krijgt de vraagconstructeur het verzoek voorafgaand aan de vaststellingsvergadering de vraag aan te passen.

De respondenten vinden in meerderheid (90%) dat er altijd of meestal voldoende tijd beschikbaar was voor het vaststellen van de vragen en/of praktijkopdrachten. Ook is men in meerderheid geheel of grotendeels tevreden over de rol van de inhoudelijke vaststeller (95%) en over de begeleiding door de toetsdeskundige van Teelen Kennismanagement (100%). De respondenten beoordelen hun vragen en opdrachten na vaststelling met gemiddeld een 7,8. Dit is dus een verbetering ten opzichte van het cijfer voor vaststelling.

Het bijwonen van een vaststellingsvergadering doet een groot beroep op de beschikbare tijd van de deelnemers, ook omdat ze soms ver moeten reizen. Daarom zijn de vraagconstructeurs niet altijd persoonlijk aanwezig bij de vergadering maar nemen ze soms telefonisch deel. Hier kiezen we uitsluitend voor als het gaat om relatief weinig en kwalitatief al goede vragen. De vaststeller is altijd persoonlijk bij de vergadering aanwezig. In een enkel geval zijn vragen via een telefonische vergadering vastgesteld.

Eén vaststeller geeft aan dat er veel verschil is tussen de kwaliteit van de beginnende en ervaren vraagconstructeurs en dat er daarom soms tijdens de vaststellingsvergadering nog veel aan een vraag moet worden verbeterd. Een andere vaststeller plaatst vraagtekens bij de inhoudelijke kennis van sommige constructeurs in relatie tot de toegewezen onderwerpen. Ten slotte geven enkele vaststellers aan dat ze bij de vraag vastgelegd willen zien waarop het correcte antwoord is gebaseerd (de onderbouwing, bijvoorbeeld de betreffende wettekst). Dit is, zeker bij vragen waarbij het correcte antwoord ook voor de inhoudsdeskundige niet in alle gevallen direct evident is, een goede suggestie. Het gaat dan nadrukkelijk **niet** om vragen waarbij het goede antwoord discutabel is (in dat geval moet de vraag worden aangepast of komen te vervallen) maar om vragen die complexe en/of weinig voorkomende situaties behandelen.

Verbetersuggesties

- Indien een constructeur of vaststeller onverhoopt niet aanwezig kan zijn bij een vaststellingsvergadering: afspraken dat Skype wordt ingezet, waarbij het beeldscherm van de toetsdeskundige wordt gedeeld, zodat de persoon op afstand mee kan kijken tijdens het bewerken van de toetsvragen of praktijkopdrachten.
- Tijdens vaststellingsvergaderingen bij vragen waarbij het juiste antwoord niet voor elke inhoudsdeskundige evident is, in een notitie bij de vraag de onderbouwing van het goede antwoord kort vastleggen.
- Indien mogelijk ook de inhoudelijk vaststeller voorafgaand aan de vergadering zijn commentaar schriftelijk laten geven.
- Bij de opdrachtverstrekking (nog) beter nagaan of de vraag constructeur voldoende inhoudelijke kennis van de toetstermen heeft.

3.7 Vastgestelde toetsen en praktijkopdrachten onderhouden

In de praktijk blijken toetsvragen en (in mindere mate) praktijkopdrachten regelmatig aangepast te moeten worden. Dat heeft verschillende oorzaken:

- Een vastgestelde toetsvraag of praktijkopdracht blijkt bij nader inzien onvoldoende duidelijk te zijn geformuleerd of zelfs niet (meer) correct. Dit soort gewenste aanpassingen komt vaak pas aan het licht na afname, doordat een kandidaat een bezwaar indient (zie ook paragraaf 4.7) of een beoordelaar wijzigingen voorstelt.
- Door een verandering in de wet- of regelgeving zijn een of enkele toetsvragen of praktijkopdrachten niet meer actueel.
- Door een verandering in de wet- of regelgeving is een grote verzameling toetsvragen of praktijkopdrachten niet meer actueel. Deze verandering is vaak al langere tijd van te voren bekend.

De vragen kunnen eenvoudig worden aangepast in TestVision Online. Daarbij moet er wel rekening mee worden gehouden dat, zodra een toets wordt gesynchroniseerd, alle wijzigingen direct worden doorgevoerd in de toets. Het is dus niet mogelijk een wijziging tot een bepaalde datum 'uit te stellen'.

Het is ook voorgekomen dat een vraag in het geheel niet meer mag worden afgenomen. De vraag moet dan uit de vragenset worden verwijderd. Omdat we inmiddels werken met at random samengestelde toetsen, hoeft de vraag niet meer (direct) te worden vervangen door een andere vraag. Indien de vulgraad van de betreffende toetsterm onvoldoende is geworden (minder dan drie vragen), zal er wel een nieuwe (parallele) vraag moeten worden ontwikkeld.

Een goed voorbeeld van een ingrijpende actualisering is de invoering van de wet BRP. Dit had grote invloed op de ontwikkelde vragen. Net samengestelde (nog niet afgenomen) toetsen moesten daardoor worden gewijzigd. Het was gezien het geringe aantal afnames efficiënter geweest het samenstellen van de toetsen een maand uit te stellen.

Verbetersuggesties

- Een of twee keer per jaar de vulgraad van de itembank met toetsvragen controleren en waar nodig nieuwe (parallele) toetsvragen ontwikkelen.
- Bij grootschalige actualisering de aanpassingen zo kort mogelijk van te voren in de toetsvragen en/of praktijkopdrachten doorvoeren en in de planning van het samenstellen van toetsen hiermee rekening houden.
- Bij het aanpassen van een individuele toetsvraag of praktijkopdracht in de notities vastleggen waarom een aanpassing is doorgevoerd, zodat dit op een later moment snel terug te vinden is.

4. Toetsafname en verwerking

4.1 Toetsresultaten en analyses

4.1.1 Toetsresultaten

Zoals beschreven in paragraaf 3.4, zijn bij elke module twee toetsversies afgenomen. Onderstaande tabel bevat gegevens over de resultaten op de toetsen.

Toetsnaam	aantal kandidaten	% geslaagd	max. score	kansscore	cesuur-score	gem. score	min. behaald	max. behaald
<i>MBO-2 Documenten/Verzoeken en Migratie versie A</i>	37	84%	60	21,51	42,60	45,60	10,90	55,07
<i>MBO-2 Documenten/Verzoeken en Migratie versie B</i>	47	91%	60	22,81	42,60	46,39	40,87	53,23
<i>MBO-4 Identiteit en Identiteitsdocumenten I versie A</i>	37	87%	50	18,80	35,50	37,29	27,32	41,42
<i>MBO-4 Identiteit en Identiteitsdocumenten I versie B</i>	50	82%	50	18,57	35,50	39,22	30,74	47,33
<i>MBO-4 Afstamming en Nationaliteit versie B</i>	43	63%	74	20,82	50,32	48,65	33,52	63,65
<i>MBO-4 Afstamming en Nationaliteit versie C</i>	51	55%	71	20,40	48,28	48,17	29,63	60,52
<i>MBO-4 Migratie versie B</i>	51	57%	30	9,55	21,00	20,15	13,81	28,31
<i>MBO-4 Migratie versie C</i>	45	73%	30	9,76	21,00	22,82	15,05	29,50
<i>HBO Identiteit en Identiteitsdocumenten II versie A</i>	22	32%	30	5,39	18,90	17,01	1,00	23,25
<i>HBO Identiteit en Identiteitsdocumenten II versie B</i>	47	68%	29	6,18	18,27	19,39	8,27	25,07

Tabel 4.1 Resultaten op de toetsen

Tot en met 31 december 2014 zijn in totaal 430 geldige toetsresultaten geregistreerd voor de in de tabel vermelde toetsversies. Deze resultaten vormen het uitgangspunt voor deze versie van het evaluatierapport. Daarnaast hebben 44 kandidaten een tijdelijke toetsversie gemaakt die op een later moment is uitgezet. Die resultaten zijn dus niet in deze evaluatie verwerkt.⁴

⁴ Het aantal toetsmomenten op toetsen in TestVision Online zal nog wat hoger liggen dan 474. Zo heeft een aantal kandidaten een toets opnieuw gemaakt na een onvoldoende resultaat en zijn de resultaten van enkele kandidaten in Bonaire niet in deze analyses meegenomen.

Wat opvalt, is dat de slaagpercentages op beide versies van de toets *MBO-2 Documenten/Verzoeken en Migratie* en de toets *MBO-4 Identiteit en Identiteitsdocumenten II* hoog zijn. Het slaagpercentage op versie A van *HBO Identiteit en Identiteitsdocumenten II* is juist vrij laag (32%).

Voor de meeste toetsen zijn de resultaten op beide toetsversies vergelijkbaar. De slaagpercentages liggen minder dan 10% uiteen en de gemiddelde scores liggen redelijk dicht bij elkaar. Bij *MBO-4 Identiteit en Identiteitsdocumenten I* valt op dat voor versie B het slaagpercentage lager is dan voor versie A, terwijl de gemiddelde score hoger is. De kandidaten die voor versie B zijn gezakt, scoren overwegend nét onder de cesuur en de geslaagde kandidaten scoren relatief hoger dan bij versie A. Voor de toetsversies van *MBO-4 Migratie* liggen de slaagpercentages wel vrij ver uit elkaar (57% versus 73%). Ook de gemiddelde scores wijken ruim twee punten van elkaar af. Hetzelfde geldt voor de toetsversies van *HBO Identiteit en Identiteitsdocumenten II* (32% versus 68%). Gezien het relatief kleine aantal kandidaten, is niet met zekerheid vast te stellen dat versies (veel) in moeilijkheid verschillen. Aangezien met ingang van januari 2015 at random samengestelde toetsen worden aangeboden en het onderscheid tussen vaste versies dan dus verdwijnt, vinden wij het niet zinvol om hier nader onderzoek naar te doen.

De cesuur voor de toetsen is gebaseerd op een cesuur van 55% rekening houdend met de kansscore van de toetsen. Om één cesuur te kunnen hanteren voor alle toetsversies per module, is de cesuur gebaseerd op de variant met de **laagste** kansscore. Bij de beide toetsversies van *MBO-4 Afstamming en Nationaliteit*, versie B van *MBO-4 Migratie* en versie A van *HBO Identiteit en Identiteitsdocumenten II*, ligt de gemiddelde score (net) onder de cesuurscore. Voor de andere toetsen ligt de gemiddelde score er (ruim) boven. Voor *MBO-2 Documenten/Verzoeken en Migratie* en *MBO-4 Identiteit en Identiteitsdocumenten II* is dit verschil ruimschoots één maal de standaardmeetfout. Gezien de slaagpercentages, lijkt de cesuur in de meeste gevallen goed gekozen. Er is dan ook geen reden geweest om de cesuur achter af aan te passen, hoewel het slaagpercentage van versie A van *HBO Identiteit en Identiteitsdocumenten II* wel onder de signaleringsnorm uit het toetsbeleid ligt (50% of minder). Ook hier geldt dat het aantal kandidaten te klein is om conclusies te verbinden aan het slagingspercentage.

4.1.2 Toets- en itemanalyse

We hebben een beknopte toets- en itemanalyse uitgevoerd op de beschikbare gegevens. Er mag geen te grote waarde aan deze informatie worden gehecht, omdat het aantal kandidaten per toets relatief klein is. Voor een betrouwbare toets- en itemanalyse, zijn minimaal 50 en liefst meer dan 100 afnamegegevens noodzakelijk. Slechts voor vier van de tien toetsversies wordt dit aantal benaderd.

In een aantal toetsversies, zoals beide versies van *MBO-4 Identiteit en Identiteitsdocumenten I* zijn relatief veel tussentijdse wijzigingen aangebracht, waardoor de analyse over een kleiner deel van de beschikbare afnames kon worden uitgevoerd. In onderstaande tabel met de gegevens van de **toetsanalyse** staat naast het aantal afnames en het aantal vragen, de gemiddelde moeilijkheid (p) en de betrouwbaarheid (KR-20/ α) van elke toets vermeld.

Toetsnaam	n	aantal vragen	p	KR-20/ α
<i>MBO-2 Documenten/Verzoeken en Migratie versie A</i>	36	60	.78	.60
<i>MBO-2 Documenten/Verzoeken en Migratie versie B</i>	31	60	.77	.16
<i>MBO-4 Identiteit en Identiteitsdocumenten I versie A</i>	15	50	.73	.35
<i>MBO-4 Identiteit en Identiteitsdocumenten I versie B</i>	22	50	.83	.54
<i>MBO-4 Afstamming en Nationaliteit versie B</i>	35	60	.70	.80
<i>MBO-4 Afstamming en Nationaliteit versie C</i>	44	60	.70	.76
<i>MBO-4 Migratie versie B</i>	49	30	.66	.51
<i>MBO-4 Migratie versie C</i>	45	30	.75	.70
<i>HBO Identiteit en Identiteitsdocumenten II versie A</i>	14	22	.66	.36
<i>HBO Identiteit en Identiteitsdocumenten II versie B</i>	47	22	.69	.49

Tabel 4.2 Toetsanalysegegevens

De ideale gemiddelde moeilijkheid van een toets hangt af van de gebruikte vraagvormen en het aantal alternatieven van de een-uit-meervragen. Gezien de mix aan vraagvormen in de onderzochte toetsversies, ligt de ideale moeilijkheid tussen de .63 en .70. De meeste p -waarden vallen binnen deze marge. De *MBO-2 toetsen* zijn, zoals al eerder vastgesteld, relatief makkelijk. Ook de toetsversies van *MBO-4 Identiteit en Identiteitsdocumenten I* zijn relatief makkelijk. De geanalyseerde versie A van *HBO Identiteit en Identiteitsdocumenten II* heeft een acceptabele moeilijkheid. Een eerdere gepubliceerde versie van de toets heeft een p van .51 (gebaseerd op een afname bij 8 kandidaten).

De ondergrens voor betrouwbaarheid wordt normaal gesproken gelegd bij .60, en voor een toets met een certificerend karakter, zoals deze toetsen, zelfs bij .80. Het merendeel van de toetsen voldoet niet aan deze ondergrens, maar zoals gezegd, kunnen er gezien het relatief lage aantal kandidaten geen conclusies aan de gegevens worden verbonden. Ook hebben toetsen met relatief weinig vragen meestal een lagere betrouwbaarheid dan toetsen met meer vragen. Dit verklaart mede de lage betrouwbaarheid van *HBO Identiteit en Identiteitsdocumenten II versie B*, aangezien het aantal afnamegegevens daar wel bijna 50 is.

Op het niveau van **itemanalyse** zijn vooral de volgende aspecten interessant om te onderzoeken:

- het aantal vragen dat door alle kandidaten correct is beantwoord (p -waarde = 1 en r_{it} -waarde = 0);
- het aantal vragen met een te hoge p -waarde: deze zijn te makkelijk;
- het aantal vragen met een te lage p -waarde: deze zijn te moeilijk;
- het aantal vragen met een negatieve r_{it} -waarde: deze vragen hebben een sterk negatief effect op de betrouwbaarheid van de toets: ze worden fout beantwoord door relatief goede kandidaten;
- het aantal vragen met een erg lage r_{it} -waarde: deze vragen dragen weinig bij aan de betrouwbaarheid.

Overigens gaan een hoge p -waarde en een lage r_{it} -waarde vaak samen. Vragen die door (bijna) iedereen correct worden beantwoord, hebben altijd een laag onderscheidend vermogen. De door iedereen correct beantwoorde vragen zijn **niet** opnieuw meegeteld in de kolommen over de p -waarden en de r_{it} -waarden.

Toetsnaam	# allen correct	# $p > .90$	# $p < .30$	# $r_{it} < 0$	# $0 < r_{it} < .05$
<i>MBO-2 Documenten/Verzoeken en Migratie versie A</i>	7	16	2	4	6
<i>MBO-2 Documenten/Verzoeken en Migratie versie B</i>	10	14	1	9	1
<i>MBO-4 Identiteit en Identiteitsdocumenten I versie A</i>	6	6	1	9	3
<i>MBO-4 Identiteit en Identiteitsdocumenten I versie B</i>	11	14	2	5	3
<i>MBO-4 Afstamming en Nationaliteit versie B</i>	3	6	2	7	2
<i>MBO-4 Afstamming en Nationaliteit versie C</i>	1	8	1	6	5
<i>MBO-4 Migratie versie B</i>	1	5	2	2	2
<i>MBO-4 Migratie versie C</i>	1	9	2	2	0
<i>HBO Identiteit en Identiteitsdocumenten II versie A</i>	1	3	1	6	1
<i>HBO Identiteit en Identiteitsdocumenten II versie B</i>	0	1	1	0	1

Tabel 4.3 Itemanalysegegevens

Zoals te verwachten op basis van de toetsresultaten, komen (erg) makkelijke vragen vooral voor in de toetsen *MBO-2 Documenten/Verzoeken en Migratie* en *MBO-4 Identiteit en Identiteitsdocumenten I*. Omdat er inmiddels at random wordt getoetst, is het wenselijk om te controleren of er binnen de vraagclusters waaruit vragen worden geselecteerd geen (te) grote verschillen in moeilijkheid bestaan. Het aantal vragen met een lage p -waarde is klein, er is geen aanleiding deze vragen nader te onderzoeken, tenzij de lage p -waarde wordt gecombineerd met een negatieve r_{it} -waarde.

Het aantal vragen met een negatieve r_{it} -waarde is iets groter. Omdat deze vragen de kwaliteit van een toets sterk negatief beïnvloeden, is het zeer wenselijk deze vragen nader te onderzoeken. Als de vulgraad van de itembank dit toelaat, kunnen de vragen het best worden afgeschreven. Indien dit niet het geval is, is het verstandig om te onderzoeken of de vragen kunnen worden aangepast.

Twee toetsen bevatten open vragen. Bij *MBO-4 Afstamming en Nationaliteit versie B* hebben de meeste open vragen goede r_{it} -waarden (hoger dan 0,50). Hierop is echter een uitzondering: één open vraag heeft een negatieve r_{it} -waarde. Bij *MBO-4 Afstamming en Nationaliteit versie C* hebben alle open vragen goede r_{it} -waarden. Ook bij *HBO Identiteit en Identiteitsdocumenten II versie A* hebben 4 van de 5 open vragen (zeer) hoge r_{it} -waarden. De vijfde open vraag heeft een r_{it} -waarde van 0,18, wat voor een open vraag aan de lage kant is. Gezien het relatief lage aantal kandidaten in de analyse, is er geen aanleiding hier nader onderzoek naar te doen. Dezelfde vraag doet het bovendien in de analyse van een eerdere gepubliceerde versie (veel) beter. Bij *HBO Identiteit en Identiteitsdocumenten II versie B* functioneren alle open vragen goed. In het algemeen kan dus worden gesteld dat het gebruik van open vragen de betrouwbaarheid en daarmee de kwaliteit van de toets ten goede komt.

Naast de bovengenoemde gegevens, kan de informatie over de a -waarden eventueel worden gebruikt om slecht functionerende afleiders aan te passen of zelfs te verwijderen. Het gaat dan echter om een groot aantal afleiders en het verwijderen daarvan zal grote gevolgen hebben voor de kansscore en daarmee voor de cesuur. Het lijkt dan ook beter om op dit gebied voorlopig geen aanpassingen door te voeren.

Verbetersuggesties

- Controleren of er binnen de vraagclusters waaruit vragen worden geselecteerd geen (te) grote verschillen in de vraagmoeilijkheid bestaan.
- Als de vulgraad van de itembank dit toelaat, de vragen met een negatieve r_{it} -waarde afschrijven. Indien dit niet het geval is, is het verstandig te onderzoeken of de vragen kunnen worden aangepast.

4.1.3 Bestede toetstijd

Ten slotte is voor elke toets geïnventariseerd hoeveel tijd de kandidaten er gemiddeld aan hebben besteed, maar ook hoeveel tijd er minimaal en maximaal is besteed. De resultaten hiervan zijn vermeld in onderstaande tabel. De maximaal beschikbare tijd is gebaseerd op het volgende uitgangspunt: 1,5 minuut per gesloten vraag en 7,5 minuut per open vraag plus 20% marge. De berekende tijdsduur wordt vervolgens omhoog afgerond op halve uren. Dit is een zodanig ruime tijdsduur dat er geen verlenging voor dyslectische kandidaten nodig is.

Verbetersuggesties

- De tijdsduur van *MBO-2 Documenten/Verzoeken en Migratie* verlagen naar 100 minuten.
- De verlaagde tijdsduur van *HBO Identiteit en Identiteitsdocumenten II* heroverwogen: 120 minuten lijkt passender dan 90 minuten.

Toetsnaam	bestede tijd (in minuten)			tijd max beschikbaar (in minuten)
	min	max	gem	
<i>MBO-2 Documenten/Verzoeken en Migratie versie A</i>	9	120	66	120
<i>MBO-2 Documenten/Verzoeken en Migratie versie B</i>	28	120	62	120
<i>MBO-4 Identiteit en Identiteitsdocumenten I versie A</i>	21	90	56	90
<i>MBO-4 Identiteit en Identiteitsdocumenten I versie B</i>	21	90	50	90
<i>MBO-4 Afstamming en Nationaliteit versie B</i>	60	150	137	150
<i>MBO-4 Afstamming en Nationaliteit versie C</i>	55	150	128	150
<i>MBO-4 Migratie versie B</i>	20	60	49	60
<i>MBO-4 Migratie versie C</i>	23	60	50	60
<i>HBO Identiteit en Identiteitsdocumenten II versie A</i>	7	147	87	150 ⁵
<i>HBO Identiteit en Identiteitsdocumenten II versie B</i>	9	145	79	150 ⁵

Tabel 4.4 Bestede tijd en maximaal beschikbare tijd per toets

Wat opvalt, is dat er altijd wel minimaal één kandidaat is die de maximale toetsduur (vrijwel) volledig benut. In het algemeen lijkt de maximale toetsduur goed gekozen. Alleen voor *MBO-2 Documenten/Verzoeken en Migratie* zou de maximale toetsduur wellicht wat kunnen verlaagd. De inmiddels doorgevoerde verlaging bij de toets *HBO Identiteit en Identiteitsdocumenten II* zou op basis van deze gegevens heroverwogen kunnen worden, aangezien de gemiddelde benodigde tijd vrijwel gelijk is aan de nieuwe maximumtijd. Om die reden zou 120 minuten wellicht een betere maximale toetsduur voor die toets zijn. Ook valt de lange gemiddelde tijdsduur bij *MBO 4 Afstamming en Nationaliteit* op. Deze lange tijdsduur wordt veroorzaakt door de lengte van de toets en doordat er zes open vragen in deze toets zijn opgenomen.

4.1.4 Controle van de beoordelingen van open vragen

We hebben een steekproefsgewijze controle op de beoordeling van open vragen uitgevoerd, te weten twee beoordelingen van de open vragen van *MBO-4 Nationaliteit en Afstamming versie B* en drie beoordelingen van de open vragen van *HBO Identiteit en Identiteitsdocumenten II versie A*, zowel

⁵ De maximale tijdsduur van deze toets is inmiddels verlaagd van 150 naar 90 minuten, omdat het aantal open vragen is teruggebracht van 12 naar 5.

gezakte als geslaagde kandidaten. Daaruit blijkt dat de beoordelaars nauwgezet en streng beoordelen. Zo resulteert bijvoorbeeld het ontbreken van de laatste e in de samenstelling Heeze-Leende bij een beoordelingsaspect dat uit nog drie andere goed beantwoorde onderdelen bestaat in de score 0 punten (in plaats van 1). Daarnaast zien we dat de antwoorden van een overall minder goede kandidaat soms wellicht iets te streng worden beoordeeld (het zogenaamde horn-effect).

We zien een verbeterpunt voor de beoordelingsvoorschriften ten aanzien van het invullen van een datum. Bij de HBO-toets blijkt daarnaast dat kandidaten de vraagstelling soms niet (volledig) interpreteren zoals bedoeld door de vraagconstructeur.

Verbetersuggesties

- Bij het beoordelen van een datum in een open vraag duidelijk vermelden of er wel of niet mag worden afgeweken van het format uit het antwoordmodel (bijvoorbeeld: 7 november 2013 mag ook worden ingevuld als 7-11-2013 of dit is juist niet toegestaan).
- De vraagstelling van de open vragen in *HBO Identiteit en Identiteitsbewijzen II* aanscherpen op basis van de door kandidaten gegeven antwoorden. Dit kan het best worden uitgevoerd nadat TestVision Online 6 beschikbaar is, omdat in die versie de antwoorden makkelijker kunnen worden afgedrukt in combinatie met het beoordelingsvoorschrift. Een mogelijke verbetering is een paar correcte antwoorden als voorbeeld in de stam opnemen (bij vragen waar het aantal gevraagde antwoorden veel groter is dan het aantal mogelijke antwoorden).

4.2 Resultaten praktijkopdrachten en analyses

4.2.1 Resultaten praktijkopdrachten

Op HBO-niveau is een aantal praktijkopdrachten afgenomen. Het betreft voor de module *HBO Identiteit en Identiteitsdocumenten II* vier versies af te leggen op de examenlocatie en drie versies die thuis of op de werkplek mogen worden gemaakt. Voor de module *HBO Landelijke ontwikkelingen* zijn er thuis of op de werkplek drie versies van een praktijkopdracht afgenomen. Onderstaande tabel bevat de resultaten van de diverse praktijkopdrachten.

Praktijkopdracht	n	% geslaagd	max. score	cesuur-score	gem. score	min. behaald	max. behaald
<i>Identiteit en Identiteitsdocumenten II Onderdeel B versie A1</i>	14	43%	16	11,20	9,57	0,00	14,00
<i>Identiteit en Identiteitsdocumenten II Onderdeel B versie A2</i>	17	6%	16	11,20	8,94	4,00	15,00
<i>Identiteit en Identiteitsdocumenten II Onderdeel B versie B1</i>	29	52%	31	21,70	19,55	0,00	30,00
<i>Identiteit en Identiteitsdocumenten II Onderdeel B versie B2</i>	13	46%	31	21,70	19,54	11,00	29,00
<i>Identiteit en Identiteitsdocumenten II Onderdeel C versie A</i>	5	20%	35	24,50	21,81	14,00	34,00
<i>Identiteit en Identiteitsdocumenten II Onderdeel C versie B</i>	38	37%	39	25,74 ⁶	22,95	11,00	35,00
<i>Identiteit en Identiteitsdocumenten II Onderdeel C versie C1</i>	24	29%	34	21,08 ⁷	17,03	0,00	30,22
<i>Landelijke ontwikkelingen versie A</i>	8	63%	72	50,40	45,75	26,00	54,00
<i>Landelijke ontwikkelingen versie B</i>	24	29%	72	48,24 ⁸	45,50	19,00	67,00
<i>Landelijke ontwikkelingen versie C</i>	21	67%	37	25,90	25,55	9,50	34,50

Tabel 4.5 Resultaten op de praktijkopdrachten

Tot en met 31 december 2014 zijn er in totaal 73 praktijkopdrachten op een examenlocatie (*HBO Identiteit en Identiteitsdocumenten Onderdeel B*) en 120 praktijkopdrachten thuis of op de werkplek (*HBO Identiteit en Identiteitsdocumenten Onderdeel C* en *HBO Landelijke ontwikkelingen*) gemaakt en beoordeeld.

⁶ De cesuur is verlaagd naar 66%, deze was oorspronkelijk 70% (27,30 punten).

⁷ De cesuur is verlaagd naar 62%, deze was oorspronkelijk 70% (23,80 punten).

⁸ De cesuur is verlaagd naar 67%, deze was oorspronkelijk 70% (50,40 punten).

Het slaagpercentage van *HBO Identiteit en Identiteitsdocumenten Onderdeel B versie A2* is opvallend laag (6%). De gemiddelde score wijkt niet erg af van versie A1. De gezakte kandidaten op versie A2 scoren gemiddeld namelijk wel beter dan de gezakte kandidaten op versie A1.

De opdrachten die thuis of op de werkplek zijn gemaakt, wijken wat betreft slaagpercentage niet veel af van de opdrachten op de examenlocatie. Er zijn geen aanwijzingen dat op grote schaal misbruik wordt gemaakt van het ontbreken van toezicht. De slaagpercentages voor de versies die thuis of op de werkplek zijn gemaakt, lopen nogal uiteen, maar gezien het soms kleine aantal kandidaten kunnen hieraan geen conclusies worden verbonden.

Op drie versies van de praktijkopdrachten is een score van 0 toegekend. Dit betekent dat de betreffende kandidaten niet hebben voldaan aan de cruciale beoordelingscriteria. Deze criteria worden dus daadwerkelijk op de gewenste manier toegepast.

De praktijkopdracht *HBO Identiteit en Identiteitsdocumenten II Onderdeel C versie B* is twee maal afgenomen, in juli 2014 en in november 2014. Bij deze versie heeft met instemming van de examencommissie een cesuurverlaging plaatsgevonden, omdat het aantal geslaagde kandidaten erg laag was. Bij de afname in juli is 36% van de kandidaten geslaagd, ook na de cesuurverlaging. Bij de afname in november was bij de oorspronkelijke cesuur (70%) 25% van de kandidaten geslaagd en bij de verlaagde cesuur (66%) werd dit 37% van de kandidaten geslaagd.

De gemiddelde scoringspercentages bleken bij de tweede afname gemiddeld lager (58%) te liggen dan die bij de eerste afname (61%). Dat betekent dat het nogmaals inzetten van een praktijkopdracht die thuis of op de werkplek mag worden gemaakt mogelijk is zonder dat dit enkel geslaagde kandidaten oplevert.

Bij de praktijkopdracht *HBO Identiteit en Identiteitsdocumenten II Onderdeel C versie C1* leverde de oorspronkelijke cesuur (70%) een slagingspercentage van 4% op. Naar aanleiding van dit slagingspercentage en de opmerkingen van de beoordelaars, is in detail gekeken naar de beoordelingscriteria en naar de uitwerkingen van de kandidaten. Met de constructeur van de praktijkopdracht en een inhoudelijke vaststeller is uiteindelijk besloten om een aantal beoordelingscriteria weg te laten. Geen van de kandidaten had de in de betreffende criteria genoemde punten in zijn uitwerking genoemd en bij nader inzien kwamen deze punten niet duidelijk genoeg naar voren in de opdrachtinstructie. Verder zijn enkele criteria samengevoegd tot een algemener criterium, waardoor bepaalde (slecht gemaakte) onderdelen minder zwaar wogen in de beoordeling van de praktijkopdracht.

Het maximale aantal te behalen punten is door deze ingrepen verlaagd van 35 naar 27 punten. Elke kandidaat behield echter het door hem of haar behaalde aantal punten om te voorkomen dat hij of zij werd benadeeld door het vervallen of samenvoegen van beoordelingscriteria. De scores werden omgerekend naar de scores die in verhouding bij 35 punten gehaald zouden zijn. Bij het gewijzigde beoordelingsformulier en de oorspronkelijke cesuur (70%) zou 13% van de kandidaten geslaagd zijn. De cesuur is bovendien verlaagd naar 62%, waardoor 31% van de kandidaten is geslaagd.

Bij *HBO Landelijke ontwikkelingen* is naar aanleiding van de tweede afname van versie B en de afname van versie C onderzoek gedaan naar de noodzaak om de cesuur en/of de resultaten aan te passen. Bij versie C was geen aanleiding voor het doorvoeren van aanpassingen. Bij versie B heeft de examencommissie besloten de cesuur te verlagen naar 67%. Dit leverde één extra geslaagde kandidaat op.

Voor alle praktijkopdrachten waarin achteraf aanpassingen zijn doorgevoerd, moeten de wijzigingen ook worden verwerkt in de opdracht zelf voor toekomstige afnames.

Gezien de aanpassingen in de later afgenomen versies was het wellicht rechtvaardiger geweest om ook de relatief lage slagingspercentages bij *HBO Identiteit en Identiteitsdocumenten II Onderdeel C versie A2* (6% bij 17 kandidaten) en *HBO Landelijke ontwikkelingen versie A* (20% bij 5 kandidaten) nader te onderzoeken. Het is verstandig de cesuur van 70% voor de praktijkopdrachten ter heroverwegen of altijd voor het bekendmaken van de resultaten te beoordelen of er aanleiding is tot aanpassingen.

Omdat de praktijkopdrachten als één item in TestVision zijn gescoord, is het niet mogelijk een toets- en itemanalyse uit te voeren. Gezien de relatief kleine aantallen kandidaten per opdracht zou een dergelijke analyse bovendien geen betrouwbare informatie opleveren.

Verbetersuggesties

- Doorvoeren van de achteraf bepaalde aanpassingen in de cesuur en criteria in de betreffende praktijkopdrachten voor toekomstige gebruik.
- Heroverwegen van de cesuur van 70% voor de praktijkopdracht of altijd voor het bekendmaken van de resultaten beoordelen of er aanleiding is tot aanpassingen.

4.2.2 Bestede tijd aan de praktijkopdrachten

Ook de bestede tijd aan praktijkopdrachten op examenlocatie is in kaart gebracht, zie onderstaande tabel.

Praktijkopdracht	bestede tijd			tijd max beschikbaar
	min	max	gem	
<i>Identiteit en Identiteitsdocumenten II Onderdeel B versie A1</i>	14	179	120	180
<i>Identiteit en Identiteitsdocumenten II Onderdeel B versie A2</i>	70	178	135	180
<i>Identiteit en Identiteitsdocumenten II Onderdeel B versie B1</i>	15	180	130	180
<i>Identiteit en Identiteitsdocumenten II Onderdeel B versie B2</i>	51	169	132	180

Tabel 4.6 Bestede tijd per praktijkopdracht

Gezien de gemiddelde tijdsduur van de praktijkopdrachten op locatie, is de periode van drie uur ruimschoots voldoende, maar er is geen reden om de maximale tijdsduur te verlagen. Ook hier zien we dat er in de meeste gevallen minimaal één kandidaat de tijd vrijwel volledig heeft benut. Bij praktijkopdrachten op examenlocatie kan worden overwogen te werken met meer dan een korte praktijkopdrachten in plaats van een lange praktijkopdracht. Op die manier kunnen meer toetstermen aan bod komen.

Voor het uploaden van de praktijkopdracht die thuis of op de werkplek is gemaakt, geldt dat dit in vrijwel alle gevallen binnen 1 à 2 minuten heeft plaatsgevonden. Wij hebben uiteraard geen zicht op de tijd die de kandidaten hebben besteed aan het uitwerken van een praktijkopdracht.

Verbetersuggesties

- Overwogen of de lange praktijkopdrachten op de examenlocatie kunnen worden vervangen door meer kortere opdrachten.

4.2.3 *Controle van de beoordelingen*

We hebben een steekproefsgewijze controle op de beoordeling van praktijkopdrachten uitgevoerd, te weten drie beoordelingen van de praktijkopdrachten bij *HBO Identiteit en Identiteitsdocumenten II – Onderdeel B*, drie beoordelingen van de praktijkopdrachten bij *HBO Identiteit en Identiteitsdocumenten II – Onderdeel C* en drie beoordelingen van de praktijkopdracht *HBO Landelijke ontwikkelingen*.

Een enkele keer heeft de beoordelaar bij de cruciale criteria (*HBO Landelijke Ontwikkelingen versie B*) geen Ja en/of Nee ingevuld, maar wel het eindoordeel Voldoende. Ook de omgekeerde werkwijze kwam voor: er was wel Ja en/of Nee, ingevuld, maar geen eindoordeel.

Eén beoordelaar heeft zijn puntenaantallen vermeld in de ruimte voor opmerkingen en niet in de kolommen met Ja en Nee (waarin het aantal te behalen punten in de kolom Ja en het cijfer 0 in de kolom Nee zijn voorgedrukt). Blijkbaar was het voor deze beoordelaar onvoldoende duidelijk hoe het beoordelingsformulier moet worden ingevuld.

De bekeken beoordelingsformulieren bevatten geen optelfouten en de subtotaal per onderdeel zijn goed overgenomen in de eindbeoordeling. Het totale aantal punten is ook in alle gecontroleerde gevallen goed overgenomen in TestVision Online. Hieruit kunnen we opmaken dat de beoordelaars zorgvuldig hebben gehandeld.

Twee kandidaten uit de steekproef haalden één punt minder dan het benodigde aantal punten voor een voldoende resultaat. Dit gegeven, samen met de relatief lage slagingspercentages, wijst erop dat de beoordelaars zeker niet soepel zijn omgegaan met het toekennen van punten.

De beoordelaars maken bij ongeveer 25% van de beoordelingscriteria gebruik van het opmerkingenveld om hun beoordeling toe te lichten. Een enkele beoordelaar geeft bij elk beoordelingscriterium een (korte) toelichting. De toelichting bestaat meestal uit maximaal één zin, of bij voorbeeld de melding 'Niet beschreven'. Het formulier lijkt dus voldoende ruimte te bieden voor opmerkingen.

De opmerkingen van de beoordelaars zijn soms waardevol voor eventuele volgende beoordelingen en zouden dus aan het beoordelingsformulier kunnen worden toegevoegd. Het is bijvoorbeeld makkelijk voor volgende beoordelaars als zij weten welk antwoord niet goed gerekend moet worden bij een bepaald beoordelingscriterium. Dit geldt echter vooral voor de praktijkopdrachten die op examenlocaties (en daardoor meerdere malen) worden gemaakt. Voor praktijkopdrachten die thuis of op de werkplek worden gemaakt lijkt dit minder zinvol, omdat zij over het algemeen slechts één tot twee keer zullen worden ingezet. Als praktijkopdrachten bekend raken, is het namelijk niet verstandig om ze in precies dezelfde vorm in te zetten, zelfs niet wanneer het beoordelingsmodel niet wordt verspreid.

De beoordelaars van praktijkopdrachten zijn veelal zelf ook constructeur. Het valt op dat zij soms uitwerkingen op de door henzelf ontwikkelde praktijkopdrachten beoordelen, terwijl dit volgens het toetsbeleidsplan niet is toegestaan. Eén beoordelaar merkte ook op het leuk te vinden om de uitwerkingen op zijn eigen praktijkopdrachten te zien. Een andere beoordelaar wilde dit echter liever niet. Het is wellicht een voordeel dat de beoordelaar de praktijkopdracht goed kent en weet wat met de beoordelingscriteria wordt bedoeld en er lijkt geen sprake te zijn van invloed op de kwaliteit van de beoordeling.

Verbetersuggesties:

- Een instructie voor beoordelaars opstellen met algemene beoordelingsregels en een voorbeeld van een op de juiste manier ingevuld beoordelingsformulier.
- Een procedure ontwikkelen voor het aanvullen van de beoordelingsformulieren bij praktijkopdrachten die op examenlocaties worden gemaakt op basis van opmerkingen van beoordelaars.
- Opnieuw overwegen of het toegestaan is dat een beoordelaar de uitwerkingen op de door hemzelf ontwikkelde praktijkopdracht beoordeelt.

4.3 Algemeen verloop afnames

De toetsafnames zijn in het algemeen goed verlopen. Zestien procent van de kandidaten die de enquête hebben ingevuld, heeft te maken gehad met technische problemen op de afnamelocatie. Een enkele kandidaat geeft aan dat er sprake was van een 'grote storing', waardoor de starttijd (veel) later was. Ook het inloggen verliep niet altijd probleemloos, maar dit is eigenlijk altijd binnen een half uur verholpen. Sommige afnamelocaties worden beschreven als krap en onrustig (Nieuwegein). Ook wordt geklaagd over het feit dat enkele locaties zijn gesloten waardoor de reisafstanden langer zijn geworden. Ten slotte waren er in enkele gevallen problemen met de instelling van het toetsenbord.

Acht procent van de kandidaten heeft problemen ervaren bij het werken met TestVision Online. We zien de volgende aandachtspunten:

- Het openen van extra informatie via een link (als PDF-document of afbeelding) vanuit de vraag wordt door 29% van de kandidaten als onprettig ervaren. In sommige gevallen is het zelfs in het geheel niet gelukt de hulpmiddelen te openen. Ook lijkt er een maximum te gelden voor het aantal keer dat een kandidaat een link kan openen.
- Het beantwoorden van de (ongelijke) matchvraag levert soms problemen op, vooral als de vraag te groot is voor het scherm en er gescrold moet worden.
- Bij het invullen van de praktijkopdracht levert het gebruik van de delete- en backspacetoets voor enkele kandidaten problemen op.
- Voor één kandidaat was het maximaal aantal tekens (ongeveer 15.000) bij de praktijkopdrachten te klein, waardoor een fout optrad. Deze kandidaat had een stuk tekst uit een ander document in het invoergeedeelte geplakt.
- Na het afronden van een praktijkopdracht, verschijnt een standaardmelding dat nog niet alle vragen zijn **beoordeeld**. Kandidaten lezen dit soms als beantwoord en raken daardoor in de war.
- Het uploaden van de praktijkopdrachten (thuis of op de werkplek) leidde tot diverse vragen aan het Examenadviesburo. In een aantal gevallen heeft het Examenadviesburo het document voor de kandidaat geüpload. Ook kregen sommige kandidaten de melding dat het bestandstype niet juist was.

Verbetersuggesties

- De opgetreden technische problemen nader onderzoeken en waar mogelijk verhelpen.
- In de instructie voor de praktijkopdrachten die thuis of op de werkplek worden gemaakt, vermelden welke bestandsformaten geüpload kunnen worden.
- Een instructie voor afnameleiders op de afnamelocaties ontwikkelen met tips over het oplossen van problemen bij het beantwoorden van sleepvragen, het openen van hulpmiddelen en andere vaker voorkomende probleemsituaties.
- Het gebruik van links met PDF-documenten en afbeeldingen zo veel mogelijk beperken. Vanaf TestVision Online versie 6 (beschikbaar in februari 2015) is het mogelijk voor kandidaten om afbeeldingen direct in de vraag te vergroten. Een aantal afbeeldingen kan daardoor alsnog direct aan de vraag worden toegevoegd.
- Het beschikbaar zijn van de hulpmiddelen nader onderzoeken/testen, in het bijzonder op locaties waar de links niet of een beperkt aantal keer geopend kunnen worden.
- De bug met de delete- en backspacetoets is gemeld bij de ontwikkelaars van TestVision Online en is vanaf versie 6 opgelost.
- In de instructie voor de praktijkopdrachten op de examenlocatie een indicatie geven voor de gewenste lengte van de uitwerking en de woordenteller bij de praktijkopdrachten aanzetten.
- In de instructie voor de praktijkopdrachten op de examenlocatie vermelden dat kandidaten geen teksten uit hulpmiddelen mogen knippen.

4.4 Bevindingen van de kandidaten

Deze paragraaf is gebaseerd op de bevindingen van de kandidaten over de **toetsen** (264 respondenten op de onder alle kandidaten uitgezette enquête) en op de **praktijkopdrachten** (41 respondenten). Hierbij moet wel de kanttekening worden gemaakt dat sommige respondenten op de enquête over toetsen ook opmerkingen maken over de praktijkopdrachten.

Bevindingen over de toetsen

Negentig procent van de kandidaten geeft aan goed geïnformeerd te zijn over de examens en 90% vindt de instructie bij de toetsen duidelijk. Degenen die dat niet vinden, geven aan dat ze geen uitleg over de toetstermen hebben gekregen, dat zij de toetsplannen niet goed leesbaar vinden (te kleine letters, betekenis kruisjes onduidelijk) en dat de opleiders zelf ook geen (goed) idee hebben waar het examen over ging en/of hoe het eruit zou zien.

Verschillende kandidaten geven aan behoefte te hebben aan proefexamens. Kennelijk zijn de oefentoetsen op de website van de NVVB niet bij alle kandidaten bekend.

Tijdens de pilotperiode bleek dat de cesuur (55%, rekening houdend met de kansscore) bij sommige kandidaten vragen oplevert. Het is voor hen niet altijd direct duidelijk hoeveel punten zij voor een toets moeten halen. Inmiddels is hierover een uitleg beschikbaar op de website van de NVVB.

Sommige kandidaten waren niet op de hoogte van het feit dat er open vragen in de toets zijn opgenomen (waarbij een deel van de kandidaten de invulvraag ook als open vraag beschouwt). Daarnaast geeft een kandidaat aan het niet prettig te vinden dat de open en gesloten vragen door elkaar worden aangeboden.

Veertien procent van de kandidaten vindt dat er te weinig tijd beschikbaar was voor de toets, 70% vindt de tijdsduur goed en 16% vindt de tijdsduur te lang. Dit sluit aan bij de gegevens over de tijdsduur uit de voorgaande paragrafen. Opvallend is dat een kandidaat (verontwaardigd) meldt dat zijn examen langer duurde dan dat van een andere kandidaat. Kennelijk is hij of zij zich er niet van bewust dat de verschillende examens een verschillende tijdsduur hebben. Daarnaast merken enkele kandidaten op dat de examenduur van *MBO-4 Afstamming en Nationaliteit* wel erg lang is (2,5 uur).

Negenzestig procent van de kandidaten vindt de vraagstelling in het algemeen duidelijk, 31% vindt dat dit niet het geval is. Er wordt onder meer kritiek geleverd op het feit dat vragen op meer dan een manier geïnterpreteerd kunnen worden. Enkele kandidaten klagen over het grote aantal spelfouten en –taalfouten. Alle vragen zijn echter taalkundig geredigeerd, dus er is zeker geen sprake van veel spel- en taalfouten. Bovendien lijkt het in een later commentaar niet zo zeer te gaan over de taalkundige kwaliteit van de examens, maar van de leermiddelen.

De vraagstelling van de open vraag leidt soms tot verwarring en ook de manier waarop antwoord moet worden gegeven, wordt niet altijd gewaardeerd. Een deel van deze klachten zal worden opgelost als de betreffende open vragen zijn omgezet naar numerieke invulvragen. Bij gesloten vragen zouden kandidaten juist graag een toelichting willen kunnen geven. Het lijkt erop dat dit laatste vooral geldt voor de meer-uit-meervragen. Ook merken verschillende kandidaten op dat ze moeite hebben met de interpretatie van de term *moeten*. De term *moeten* is echter vaak bewust gebruikt, omdat *kunnen* te veel onduidelijkheid oplevert.

Zevenenzestig procent van de kandidaten vindt de moeilijkheid van de toetsen goed. Twee procent vindt het examen te makkelijk en 31% vindt de toetsen te moeilijk.

Vijfenvierzig procent van de kandidaten heeft hulpmiddelen meegenomen naar het examen. Vierenzestig procent van de kandidaten vond de hulpmiddelen toereikend om het examen af te leggen.

Verscheidende andere kandidaten hadden echter graag de Wet BRP, de PUN en/of de RWN gebruikt als hulpmiddel, omdat deze ook op de werkplek worden geraadpleegd. De HUP en het LO mogen bij een deel van de examens worden gebruikt. Deze hulpmiddelen zijn door een deel van de kandidaten inderdaad gebruikt, maar een ander deel van de kandidaten vond juist deze hulpmiddelen niet relevant.

Ook merken enkele kandidaten op dat ze er helemaal niet mee bekend waren dat er hulpmiddelen mogen worden gebruikt. In de verschillende communicatie-uitingen (toetsplan, indicatie, bevestiging) wordt hierover tegenstrijdige informatie gegeven. Vanaf januari 2015 worden de hulpmiddelen uitsluitend digitaal aangeboden. Dit komt de duidelijkheid ten goede, maar nog steeds moeten kandidaten hierover goed worden geïnformeerd. Het bleek voor verschillende kandidaten ook verwarrend dat in het toetsplan vermeld stond dat kladpapier was toegestaan, maar dat dat alleen kladpapier van de examenlocatie betrof en geen eigen kladpapier.

Diverse kandidaten hebben de examenorganisatie gebeld, voornamelijk over de planning van het examen. Maar ook blijkt het verstrekken van de resultaten onduidelijkheid op te leveren. De melding van het systeem was niet optimaal te zijn bij toetsen met open vragen (dit is inmiddels naar aanleiding van de opmerkingen hierover al aangepast door het Examenadviesburo) en enkele kandidaten merken op dat de uitslag bij de ene kandidaat eerder/snelser wordt verstrekt dan bij de andere. Ook merken enkele kandidaten op dat ze het afgelegde examen graag op papier mee zouden krijgen.

In het algemeen is men tevreden over de informatievoorziening (ruim twee derde van de kandidaten), maar er zijn natuurlijk kandidaten die hierover klachten hebben, bijvoorbeeld op de informatievoorziening over de inzageprocedure. Gelukkig geven andere kandidaten aan dat ze vriendelijk en goed te woord zijn gestaan.

Ten slotte willen we hier graag vermelden dat 72% van de kandidaten aangeeft aan het examen deel te hebben genomen met als reden zelfontwikkeling, 38% nam (ook) deel op verzoek van de werkgever, 9% om de carrièremogelijkheden binnen de organisatie te vergroten en 15% om de kansen bij een nieuwe werkgever te vergroten. In de toelichting noemen veel kandidaten dat ze een bijdrage willen leveren aan de pilot en aan kwaliteitsverbetering.

Bevindingen over de praktijkopdrachten

Vijfentachtig procent van de respondenten vindt de instructie van de praktijkopdracht duidelijk. Twee derde van de respondenten had voldoende tijd voor het maken van de praktijkopdracht, 30% vindt de tijd te kort en 5% te lang.

Wat betreft de inhoud van de opdracht vindt 75% deze duidelijk en 25% niet. Over de moeilijkheidsgraad is 72% tevreden, 2% vindt de opdracht te makkelijk en 26% te moeilijk.

Twee derde van de kandidaten vindt dat de praktijkopdracht goed aansluit op de praktijk. De kandidaten die dit niet vonden, hadden geen tot weinig ervaring met het opstellen van een adviesnota. Eén kandidaat geeft wel aan er veel van geleerd te hebben en het interessant te hebben gevonden.

Van de 16 kandidaten die documenten in een tweede venster konden bekijken, is 75% tevreden over de werking daarvan. De overige kandidaten hadden hier (veel) moeite mee.

Onderdeel C is door 77% thuis gemaakt en door 23% op de werkplek. In 70% van de gevallen ging het uploaden zonder problemen. De overige kandidaten hadden problemen die in overleg met het Examenadviesburo zijn opgelost (zie ook paragraaf 4.3). Enkele respondenten geven aan graag een bevestiging van ontvangst te krijgen.

Diverse kandidaten hebben de examenorganisatie gebeld, onder meer over het aanmelden voor het examen. In het algemeen is men tevreden over de informatievoorziening (twee derde van de kandidaten). Wel heeft een aantal kandidaten klachten over de wijzigingen in de planning van de praktijkopdrachten, vooral omdat de afnameperiode daardoor in een vakantieperiode viel.

Bij de praktijkopdrachten heeft 90% deelgenomen met als reden zelfontwikkeling, 10% (ook) op verzoek van de werkgever, 5% om de carrièremogelijkheden binnen de organisatie te vergroten en 10% om de kansen bij een nieuwe werkgever te vergroten.

Verbetersuggesties

- De oefentoetsen op de website (nog) beter onder de aandacht brengen van de kandidaten.
- In de informatie (nog) duidelijker vermelden welke toetsen open vragen bevatten.
- De oefentoets voorzien van de genoemde 'probleemvragen' zodat de kandidaten beter zicht krijgen op het waarom van goede antwoorden.
- Enkele voorbeelden van open vragen inclusief modelantwoord in de vorm van een PDF-document op de website van NVVB plaatsen.
- De open vragen waar mogelijk omzetten naar meervoudige invulvragen, waardoor het beantwoorden beter wordt ondersteund (zie ook paragraaf 3.3).
- Eenduidig en duidelijk communiceren over het beschikbaar zijn van hulpmiddelen tijdens het examen en over het niet toegestaan zijn van eigen kladpapier.

- Voor alle modules nogmaals controleren of voldoende en de juiste hulpmiddelen worden beschikbaar gesteld.
- Duidelijker informeren over de verstrekking van de uitslag, toelichten waarom de vragen niet achteraf meegenomen mogen worden (geheimhouding) en de mogelijkheid tot inzage benoemen.
- De leesbaarheid van de toetsplannen verbeteren.

4.5 Bevindingen van de beoordelaars

Beoordelaars moeten twee soorten producten van kandidaten beoordelen:

- open vragen;
- praktijkopdrachten.

De open vragen en de praktijkopdrachten op de examenlocatie zijn door de kandidaat in TestVision Online ingevoerd en kunnen direct worden beoordeeld. De praktijkopdracht die thuis of op de werkplek wordt gemaakt, wordt door de kandidaat geüpload en vervolgens door de beoordelaar gedownload.

Tijdens een bespreking zijn de bevindingen van de beoordelaars geïnventariseerd. In het algemeen zijn de beoordelaars positief over het werken met het beoordelingsmodel en de cruciale criteria. We zien de volgende aandachtspunten:

- Het valt soms niet mee om de opdracht te beoordelen, vooral als het antwoord minder goed overeenkomt met de modeluitwerking.
- Er is behoefte aan meer beoordelaarsinstructie over het 'doorwerk-effect' van gemaakte fouten.
- De beoordelaars zouden in het beoordelingsmodel meer aandacht willen zien voor schrijfvaardigheden, zoals het kunnen opstellen van een adviesnota, collegevoorstel, plan van aanpak of brief, inclusief vorm- en taalaspecten.
- Men vindt het niet gewenst dat men ook bij de opdrachten voor andere beoordelaars kan. Vanaf TestVision Online 6 kan het beoordelen zo worden ingericht dat dit niet meer mogelijk is.

Tot december 2014 was er geen enkele aanwijzing dat kandidaten fraude of plagiaat hadden gepleegd. In december 2014 is bij *HBO Landelijke ontwikkelingen versie C* door een beoordelaar geconstateerd dat de uitwerkingen van twee kandidaten te veel overeenkomen om dit toe te schrijven aan toeval. Een nadere vergelijking van de uitwerkingen door een onderwijskundige leidde ook tot deze conclusie. Beide kandidaten zijn uitgenodigd voor een gesprek, waarvan de uitkomst op dit moment nog niet bekend is. Op basis van dit ene geval is er nog onvoldoende reden om nu te gaan werken met plagiaatsoftware.

Verbetersuggesties

- Formuleren van toetstermen en beoordelingscriteria voor schrijfvaardigheid.
- Vanaf TestVision Online versie 6 werken met het handmatig toewijzen van beoordelaars in TestVision Online, zodat elke beoordelaar alleen zijn eigen kandidaten kan zien. Daarbij kan ook voor anoniem beoordelen worden gekozen.
- Vanaf TestVision Online versie 6 herbeoordelingen laten uitvoeren voor kandidaten van wie de uitslag minder dan 10% afwijkt van de cesuur.
- Bij de praktijkopdrachten die thuis of op de werkplek mogen worden gemaakt de instructiezin over het plegen van fraude uitbreiden met de informatie dat het beschikbaar stellen van de eigen uitwerking aan andere kandidaten ook als fraude wordt beschouwd.

4.6 Bevindingen van de opleiders

Vier opleiders hebben een evaluatieformulier ingevuld en drie opleiders waren vertegenwoordigd bij een evaluatiebijeenkomst op 28 november 2014. De opleiders hebben de volgende opmerkingen rond de examens gemaakt:

- De informatievoorziening voor het inschrijven van examens zou verbeterd kunnen worden (hiervoor hebben we in paragraaf 4.4 al een aantal verbetersuggesties gegeven).
- De opleiders zien inmiddels de kracht van het gebruik van gesloten vragen, ook voor het toetsen van reproductieve vaardigheden.

De opleiders zijn relatief onbekend met (de inhoud van) de examens. Ze krijgen de informatie over de examens voornamelijk via de kandidaten. Sommige kandidaten geven in de enquête zelfs aan dat de opleiders hen verzoeken de inhoud van de vragen te onthouden en aan ze door te geven. Het lijkt daarom goed om de opleiders beter inzicht in de examens te geven, waarbij het natuurlijk niet wenselijk is dat men zich in de opleiding uitsluitend gaat richten op het voorbereiden op de examens.

Verbetersuggesties

- De opleiders informeren over de inhoud van de examens, bijvoorbeeld door hen de mogelijkheid te bieden een examen af te leggen of een informatiebijeenkomst over examenvragen te beleggen, waarin ook aandacht kan worden besteed aan (concept)toetstermen en het ontwikkelen van oefenvragen.
- Opleiders er op wijzen dat het ongewenst is kandidaten te vragen om de inhoud van vragen door te geven.

4.7 Ingediende bezwaren

Kandidaten kunnen een bezwaarschrift indienen over (de inhoud van) het examen en hun resultaat. Een bezwaarschrift wordt behandeld door de adviescommissie *Bezwaarschriften van de NVVB Examen Organisatie*. Afhankelijk van het oordeel van de adviescommissie, wordt de uitslag van de kandidaat aangepast, meestal doordat een of enkele vragen goed worden gerekend of doordat de score op een open vraag wordt aangepast. Er zijn gedurende de pilotperiode twaalf bezwaarschriften ingediend, waarvan er drie (gedeeltelijk) zijn gehonoreerd.

Op dit moment wordt een wijziging meestal eerst verwerkt in TestVision Online. Daarna wordt het resultaat in EMO aangepast, waarbij het cijfer handmatig wordt herberekend. Dit heeft twee nadelen: er is sprake van een dubbele handeling én het is foutgevoelig. De dubbele handeling kan worden voorkomen door de wijziging alleen in EMO door te voeren en niet in TestVision Online. Dit heeft opnieuw twee nadelen. Ten eerste komen de gegevens in TestVision Online niet meer overeen met de werkelijkheid, waardoor je niet kunt vertrouwen op bijvoorbeeld de review op kandidaatniveau, de informatie over slaagpercentages en andere rapportages uit TestVision Online. Daarnaast is de kans op fouten bij het verwerken van de wijziging nog groter dan in de situatie waarin de nieuwe score uit TestVision Online wordt overgenomen. Om die reden kan de dubbele registratie het best gehandhaafd blijven. Ook synchronisatie van de cijfers zou een goed alternatief zijn. Daarbij zou EMO de cijfers uit TestVision Online moeten overnemen. Daartoe zal EMO echter moeten worden aangepast.

Naast het aanpassen van het resultaat van de individuele kandidaat, kan de verwerking van een bezwaarschrift ook leiden tot een aanpassing op vraagniveau, bijvoorbeeld het aanpassen van de vraagstelling, van een afleider om soms zelfs van de sleutel.

Naast de officiële bezwaarschriften, staan er soms in EMO registraties van door kandidaten gemaakte opmerkingen over de toets en/of de toetsvragen of over andere (technische) problemen tijdens de afname. Er wordt een overzicht bijgehouden van deze opmerkingen. Wat betreft inhoudelijke opmerkingen kunnen kandidaten het best overwegen of ze een officieel bezwaar willen indienen. Hetzelfde geldt als de kandidaat van mening is dat hij of zij is benadeeld door de opgetreden problemen.

Verbetersuggesties

- In een procedure duidelijk beschrijven wie waarvoor verantwoordelijk is bij het verwerken van bezwaarschriften en de verwerking ervan volgens deze procedure afhandelen.
- De werkwijze voor het afhandelen van niet officieel ingediende klachten en opmerkingen opnemen in het toetsbeleidsplan.

- De mogelijkheid onderzoeken om het cijfer vanuit TestVision Online in EMO over te nemen, zodat ook bij een wijziging het cijfer uit TestVision Online kan worden overgenomen en niet (handmatig) hoeft te worden herberekend.

5. Overige aandachtspunten

5.1 Implementatie van het toetsbeleid

We kunnen concluderen dat de werkwijze die in het toetsbeleidsplan is beschreven, grotendeels wordt nageleefd. Hieronder vermelden we de punten die afwijken van het toetsbeleidsplan of die nog niet zijn uitgevoerd of geïmplementeerd:

- De praktijkopdrachten die op de examenlocatie worden gemaakt, hoeven niet direct na de afname te worden ingeleverd bij de examenleider. De afname vindt immers volledig digitaal plaats.
- Er is gedurende de pilot een extra (derde) toetsmoment geweest voor de praktijkopdrachten die thuis of op de werkplek worden uitgewerkt, terwijl in het toetsbeleidsplan wordt gesproken over twee afnamemomenten.
- In het toetsbeleidsplan wordt nog uitgegaan van twee inhoudsdeskundige vaststellers. Een groot deel van de toetsvragen is met één inhoudelijke vaststeller en de (inhoudelijk deskundige) constructeur vastgesteld.
- De toetsplannen bevatten (iets) andere en meer informatie dan was voorzien. Er is geen informatie over de vulgraad van de itembank in de toetsplannen vermeld. Dit past ook niet in het toetsplan, maar hoort in het toetsbeleid thuis. Het advies is om voor elke toetsterm of elk cluster van toetstermen **minimaal** drie toetsvragen beschikbaar te hebben. Het is verstandig om in de loop van de tijd het aantal vragen uit te breiden naar vijf vragen per toetsterm of cluster en 'slechte' vragen af te schrijven.
- In het toetsbeleidsplan staat expliciet vermeld dat beoordelaars nooit open vragen en praktijkopdrachten die ze zelf hebben ontwikkeld mogen beoordelen. Dit gebeurt in de praktijk wel (zie ook paragraaf 4.2).
- Er is een beperkte tweede beoordeling (door toetsdeskundigen, niet door beoordelaars) uitgevoerd. In het toetsbeleidsplan was dit ruimer voorzien.
- In de oefentoetsen zijn geen open vragen opgenomen, omdat deze bij digitale afname door een beoordelaar beoordeeld zouden moeten worden en dat is bij oefentoetsen niet wenselijk.
- In het toetsbeleidsplan staat dat voor **elke** module voorbeeldvragen beschikbaar zijn. Voor de pilotmodules is dat inderdaad het geval. Voor de vervolgmodes zal dat waarschijnlijk niet zo zijn, tenzij ervoor wordt gekozen om bijvoorbeeld enkele parallelle vragen te laten ontwikkelen. De tekst in het toetsbeleidsplan moet daar wellicht op worden aangepast.
- De vaststelling van het toetsplan vindt niet gecombineerd plaats met de vaststelling van de toetsen. In de praktijk wordt het toetsplan eerder vastgesteld.
- Voor praktijkopdrachten die op een examenlocatie worden gemaakt, gelden geen vastgestelde toetsmomenten. Geregistreerde kandidaten kunnen zich hiervoor dus inschrijven op een tijdstip en locatie naar keuze.
- Volgens het toetsbeleidsplan moet ICTU bij extreem hoge of lage scores (conform zogenaamde signaleringsnormen) vaststellen of er maatregelen (zoals het aanpassen van de cesuur) moeten worden genomen. Kandidaten

krijgen alleen een voorlopig resultaat. In de praktijk is het lastig om bij flexibele (doorlopende afnames) een moment te kiezen om de scores te beoordelen. Bovendien is het aantal kandidaten (te) klein om zoveel waarde te hechten aan de slaagpercentages. Deze werkwijze wordt dus in de praktijk niet uitgevoerd.

- De stuurgroep NBP heeft besloten om de inschaling door het NCP NLQF achterwege te laten.
- De stuurgroep NBP moet nog een keuze maken over het vastleggen van de examenresultaten in een landelijk diplomaregister.
- In het toetsbeleidsplan staat dat de NVVB aan het einde van het project NBP een overgangsregeling bepaalt. De NVVB moet nog bepalen hoe deze overgangsregeling eruit zal zien.
- De NVVB/BZK moet nog een keuze maken voor extern toezicht (bijvoorbeeld door Stichting Examenkamer), waarmee de examenorganisatie laat zien dat ze voldoet aan de *Algemene beginselen van behoorlijk examineren*.
- Tijdens de stuurgroepbijeenkomst op 8 december 2014 wordt een voorstel voor de in te voeren diplomalijnen besproken. De diplomalijnen zijn dus nog niet definitief.

Verbetersuggesties

- Actualiseren van het toetsbeleid, bijvoorbeeld ten aanzien van het inleveren van praktijkopdrachten, de inhoud en vaststelling van de toetsplannen, de vulgraad van de itembank, inschrijving praktijkopdrachten en beschikbaarheid oefenvragen.
- Enkele voorbeelden van open vragen inclusief modelantwoord in de vorm van een PDF-document op de website van NVVB plaatsen.
- Opnieuw overwegen of het toegestaan is dat een beoordelaar de uitwerkingen op de door hemzelf ontwikkelde praktijkopdracht beoordeelt en in een breder kader: bepalen of een systeem van functiescheiding ontwikkeld moet worden.
- Weglaten van de signaleringsnormen voor de cesuur in het toetsbeleid, omdat zij weinig nut hebben gezien de kleine aantallen kandidaten en de flexibele afnameperiode.

5.2 Communicatie

Tussen de projectleider vanuit ICTU en Teelen Kennismanagement is veelvuldig telefonisch en via e-mail contact geweest. In voorbereiding op de eerste ontwikkelsessies heeft ook overleg in Wilp plaatsgevonden. Ook tussentijds werd enkele keren ergens voor overleg afgesproken. Binnen Teelen Kennismanagement vond regelmatig een intern projectteamoverleg plaats, waarbij de projectleider vanuit ICTU enkele malen aanwezig was.

In eerste instantie zou het platform PLEIO ingezet worden voor de communicatie over het project. Al vrij snel bleek dat niet te werken, omdat mensen eerst moesten inloggen, het niet overzichtelijk was en mensen

tegelijk in documenten aan de slag gingen. Uiteindelijk is veel communicatie tijdens de pilot via e-mail verlopen.

Voor het plannen van vaststellingsvergaderingen is regelmatig gebruik gemaakt van www.datumprikker.nl. Een voordeel daarvan is dat de beschikbaarheid van een groot aantal personen kan worden geïnventariseerd. Een nadeel is dat deze datumprikker eigenlijk bedoeld is om één afspraak te plannen. Over de geplande vaststellingsvergaderingen moest dus alsnog via e-mail worden gecommuniceerd.

Tussen Teelen Kennismanagement en Examenadviesburo is regelmatig contact geweest via telefoon en e-mail. Beide partijen reageren snel op elkaars vragen en verzoeken, wat prettig werkt.

Over het project zijn een aantal artikelen geschreven in de Betrouwbare Courant: de communicatie rondom het project kreeg dus voldoende aandacht. Verder vond er communicatie plaats via nieuwsbrieven van de NVVB en de website van de NVVB. Het project NBP is ook enkele keren gepubliceerd in de 'Van GBA- tot Z krant van BZK. Bovendien heeft in het tijdschrift voor Examens een publicatie over dit project bestaan.

Vrij laat in de pilot werd bepaald dat kandidaten hun cijfers op de onderdelen van de HBO-modules door cijfers van andere onderdelen kunnen compenseren (de kandidaat krijgt een eindcijfer voor de gehele module). Hierover was eerder nog niet gecommuniceerd, hoewel de weging tussen de onderdelen wel in het toetsplan was vastgelegd.

Verbetersuggesties

- In de toetsplannen van de modules met meer onderdelen vermelden dat het eindcijfer is gebaseerd op het gewogen gemiddelde en dat dus niet altijd alle onderdelen met een voldoende hoeven te worden afgerond.
- Kandidaten nog meer wijzen op de informatie die op de NVVB-site staat, om vragen te voorkomen waarop het antwoord nu al te vinden is.

5.3 Examencommissie

De Examencommissie bestaat uit de volgende leden:

- Frans Korpel (voorzitter en werkzaam bij het Centraal Bureau rijvaardigheidsbewijzen)
- Geert Deenen (deskundige kwalificatiestructuur NBP)
- Harry Molkenboer (toetsdeskundige)
- Klaasje van der Haar (praktijkdeskundige gemeente Zwolle)
- Paul Hoogzaad (praktijkdeskundige Agentschap BPR)
- Ron van Troost (projectleider NBP)
- een verslaglegger namens project NBP

De examencommissie is in de pilotperiode twee keer bij elkaar geweest. De eerste keer (oktober 2013) is gesproken over taken en rollen van de Examencommissie en is het toetsbeleidsplan toegelicht. De tweede keer (september 2014) stonden we stil bij de stand van zaken examens op dat moment. Men adviseerde onder meer om de flexibele toetsen ad random samen te stellen. Verder heeft de Examencommissie een inhoudelijke discussie gevoerd over de praktijkopdrachten voor de module *HBO Identiteit en Identiteitsdocumenten II*.

6. Overzicht van de verbeter suggesties

Dit evaluatierapport bevat diverse verbeter suggesties. In dit hoofdstuk ordenen we deze verbeter suggesties, waarbij we uitgaan van de persoon/organisatie die de beschreven actie kan uitvoeren. De verbeter suggesties variëren van het nemen van een beslissing (geel) tot het uitvoeren van activiteiten zoals het doorvoeren van wijzigingen in bijvoorbeeld werkinstructies, examens of toetsvragen (groen). Voor alle verbeter suggesties geldt dat Stichting ICTU voorafgaand aan de uitvoering moet afwegen of ze de suggesties wil (laten) doorvoeren.

6.1 Verbetersuggesties ICTU/NVVB

1. Onderzoeken of de kwalificatiestructuur voor MBO-2 opgeheven kan worden en de twee modules ondergebracht kunnen worden in de MBO-4 kwalificatiestructuur of als losstaande startmodules die buiten de kwalificatiestructuur vallen (eventueel inclusief de BOBZ-module *Hoe werkt de gemeente* en een algemene module *Burgerzaken*).
2. Regelmatig onderhoud op toetstermen inplannen. Dit zou minimaal één keer per jaar moeten plaatsvinden en daarnaast waar nodig incidenteel, zoals bij de invoering van LO3.9 per 31 januari 2015. Bij elke wijziging moeten natuurlijk ook alle vastgestelde vragen en opdrachten (op het gebied van de wijziging) worden nagelopen.
3. Minder nadruk leggen op het onderscheid tussen PIV en IDM en in de diplomalijnen rekening houden met vrije keuzes in modules, die aansluiten bij de taken van een medewerker Burgerzaken binnen een grote gemeente (meer gericht op specialisaties) en binnen een kleine gemeente (meer gericht op allround inzetbaarheid).
4. Opleiders attenderen op de door cursisten gesignaleerde discrepantie tussen de inhoud van de opleiding en de toetstermen.
5. Nader onderzoeken wat de opleiders exact onduidelijk vinden aan de toetsplannen en waar mogelijk de toetsplannen uitbreiden en verduidelijken.
6. Duidelijkheid verstrekken over de verdere ontwikkeling van de kwalificatiestructuur en over de uitfasering van de huidige diplomalijnen BOBZ/BOP en BABZ/MLOBZ.
7. Deelnemers van tevoren beter informeren over doel, inhoud en belang van de cursussen.
8. Toekomstige vraagconstructeurs uitleggen hoe belangrijk het voor de kwaliteit van de examenvragen en de efficiëntie van het ontwikkeltraject is om de basiscursus *Vraagconstructie* te volgen indien men nog niet over de vereiste kennis en vaardigheden beschikt.
9. In de uitnodiging beter duidelijk maken dat deelname niet vrijblijvend is, bijvoorbeeld door deelnemers de uitnodiging te laten bevestigen en erop aandringen dat deelnemers zich tijdig afmelden als ze niet kunnen.

10. Bij de opdrachtverstrekking de toetstermen verdelen over zo min mogelijk verschillende vraagconstructeurs en de vraagconstructeurs geen eigen keuzen laten maken.
11. Het opstellen van de opdrachtverstrekkingen door Teelen Kennismanagement laten uitvoeren.
12. In de opdrachtverstrekking het belang van het ontwikkelen op het juiste taxonomie-niveau benadrukken (inclusief korte toelichting op de niveaus met voorbeelden).
13. Een overzicht maken van bij de toetstermen passende documenten, procedures of beschikbare leerstof en achtergrondinformatie en dit ter beschikking stellen aan de vraagconstructeurs.
14. De vraagconstructeurs verder professionaliseren in het ontwikkelen van andere vraagvormen. De verdiepingscursus is hiervoor een zeer geschikte mogelijkheid.
15. Per situatie afwegen of parallelle vragen het best door een vraagconstructeur of door een toetsdeskundige kunnen worden ontwikkeld.
16. Bepalen of het zinvol is om bij bepaalde toetstermen (op toepassingsniveau) vragen met videofragmenten te laten ontwikkelen.
17. Onderzoeken wat de beste manier is om de praktijkopdrachten aan te bieden aan de kandidaten en het proces van beoordelen te optimaliseren.
18. Bij de opdrachtverstrekking (nog) beter nagaan of de vraagconstructeur voldoende inhoudelijke kennis van de toetstermen heeft.
19. Een of twee keer per jaar de vulgraad van de itembank met toetsvragen controleren en waar nodig nieuwe (parallele) toetsvragen ontwikkelen.
20. Bij grootschalige actualisering van de aanpassingen zo kort mogelijk van tevoren in de toetsvragen en/of praktijkopdrachten doorvoeren en in de planning van het samenstellen van toetsen hiermee rekening houden.
21. De tijdsduur van *MBO-2 Documenten/Verzoeken en Migratie* verlagen naar 100 minuten.
22. De verlaagde tijdsduur van *HBO Identiteit en Identiteitsdocumenten II* heroverwegen: 120 minuten lijkt passender dan 90 minuten.
23. Heroverwegen van de cesuur van 70% voor de praktijkopdracht of altijd voor het bekendmaken van de resultaten beoordelen of er aanleiding is tot aanpassingen.
24. Overwegen of de lange praktijkopdrachten op de examenlocatie kunnen worden vervangen door meer kortere opdrachten.
25. Opnieuw overwegen of het toegestaan is dat een beoordelaar de uitwerkingen op de door hemzelf ontwikkelde praktijkopdracht beoordeelt en in een breder kader: bepalen of een systeem van functiescheiding ontwikkeld moet worden.
26. De oefentoetsen op de website (nog) beter onder de aandacht brengen van de kandidaten.
27. In de informatie (nog) duidelijker vermelden welke toetsen open vragen bevatten.

28. Eenduidig en duidelijk communiceren over het beschikbaar zijn van hulpmiddelen tijdens het examen en over het niet toegestaan zijn van eigen kladpapier.
29. Duidelijker informeren over de verstrekking van de uitslag, toelichten waarom de vragen niet achteraf meegenomen mogen worden (geheimhouding) en de mogelijkheid tot inzage benoemen.
30. De leesbaarheid van de toetsplannen verbeteren.
31. De opleiders informeren over de inhoud van de examens, bijvoorbeeld door hen de mogelijkheid te bieden een examen af te leggen of een informatiebijeenkomst over examenvragen te beleggen, waarin ook aandacht kan worden besteed aan (concept)toetstermen en het ontwikkelen van oefenvragen.
32. Opleiders er op wijzen dat het ongewenst is kandidaten te vragen om de inhoud van vragen door te geven.
33. In een procedure duidelijk beschrijven wie waarvoor verantwoordelijk is bij het verwerken van bezwaarschriften en de verwerking ervan volgens deze procedure afhandelen.
34. De werkwijze voor het afhandelen van niet officieel ingediende klachten en opmerkingen opnemen in het toetsbeleidsplan.
35. Actualiseren van het toetsbeleid, bijvoorbeeld ten aanzien van het inleveren van praktijkopdrachten, de inhoud en vaststelling van de toetsplannen, de vulgraad van de itembank, inschrijving praktijkopdrachten en beschikbaarheid oefenvragen.
36. Weglaten van de signaleringsnormen voor de cesuur in het toetsbeleid, omdat zij weinig nut hebben gezien de kleine aantallen kandidaten en de flexibele afnameperiode.
37. In de toetsplannen van de modules met meer onderdelen vermelden dat het eindcijfer is gebaseerd op het gewogen gemiddelde en dat dus niet altijd alle onderdelen met een voldoende hoeven te worden afgerond.
38. Kandidaten nog meer wijzen op de informatie die op de NVVB-site staat, om vragen te voorkomen waarop het antwoord nu al te vinden is.

6.2 Verbetersuggesties Examenadviesburo

39. De mogelijkheid onderzoeken om het cijfer vanuit TestVision Online in EMO over te nemen, zodat bij de cesuur rekening kan worden gehouden met de kansscore en er geen vaste maximale score meer noodzakelijk is en zodat ook bij een wijziging het cijfer uit TestVision Online kan worden overgenomen en niet (handmatig) hoeft te worden herberekend.
40. De opgetreden technische problemen nader onderzoeken en waar mogelijk verhelpen.
41. Het beschikbaar zijn van de hulpmiddelen nader onderzoeken/testen, in het bijzonder op locaties waar de links niet of een beperkt aantal keer geopend kunnen worden.

42. Vanaf TestVision Online versie 6 werken met het handmatig toewijzen van beoordelaars in TestVision Online, zodat elke beoordelaar alleen zijn eigen kandidaten kan zien. Daarbij kan ook voor anoniem beoordelen worden gekozen.
43. Vanaf TestVision Online versie 6 herbeoordelingen laten uitvoeren voor kandidaten van wie de uitslag minder dan 10% afwijkt van de cesuur.

6.3 Verbetersuggesties inhoudskundige

44. Toetstermen beoordelen op duidelijkheid, op (gedeeltelijke) overlap en op de mogelijkheid om gecombineerd te worden met inhoudelijk zeer gerelateerde andere toetstermen, zodat bepaalde onderwerpen minder of juist meer aan bod komen. Hierbij kan ook worden gekeken of het totale aantal vragen per toets kan worden verlaagd.
45. Beoordelen of meer modules moeten worden gesplitst in twee delen (zonder en met hulpmiddelen) en dit vervolgens realiseren in TestVision Online.
46. De (te) abstracte of complexe toetstermen voorzien van een uitgebreidere toelichting of enkele voorbeeldvragen.
47. Voor alle modules nogmaals controleren of voldoende en de juiste hulpmiddelen worden beschikbaar gesteld.

6.4 Verbetersuggesties toetsdeskundige

48. In de basis cursus Vraagconstructie aandacht besteden aan het verschil tussen de niveaus door voorbeelden van toetstermen over hetzelfde onderwerp op verschillende niveaus te bespreken.
49. In de basis cursus Vraagconstructie meer oefenen met het ontwikkelen van vragen op het juiste taxonomie niveau en het maken van vragen op RV- én B-niveau bij één toetsterm.
50. Voorafgaand aan de opdrachtverstrekking een actueel overzicht maken met de benodigde toetsvragen per (cluster van) toetsterm(en).
51. De werkinstructies actualiseren en uitbreiden met de genoemde informatie.
52. In de werkinstructie de informatie over de naamgeving van vragen en de iteminstellingen (vaste volgorde alternatieven, taxonomie niveau, wel of geen gedeeltelijke scoring) meer benadrukken/uitlichten.
53. Bij het ontwikkelen van parallelle of extra vragen bij een toetsterm de oorspronkelijke vragen als PDF-document bij de opdracht verstrekken.
54. De open vragen die ervoor geschikt zijn omzetten naar meervoudige invulvragen.
55. Het MS Word sjabloon actualiseren en opsplitsen in twee varianten: één voor de praktijkopdrachten op examenlocaties en één voor praktijkopdrachten die thuis of op de werkplek worden gemaakt.

56. Indien een constructeur of vaststeller onverhoopt niet aanwezig kan zijn bij een vaststellingsvergadering: afspraken dat Skype wordt ingezet, waarbij het beeldscherm van de toetsdeskundige wordt gedeeld, zodat de persoon op afstand mee kan kijken tijdens het bewerken van de toetsvragen of praktijkopdrachten.
57. In de werkinstructie opnemen dat vraagconstructeurs bij vragen waarbij het juiste antwoord niet voor elke inhoudsdeskundige evident is, in een notitie bij de vraag de onderbouwing van het goede antwoord kort moeten vastleggen.
58. Bij het aanpassen van een individuele toetsvraag of praktijkopdracht in de notities vastleggen waarom een aanpassing is doorgevoerd, zodat dit op een later moment snel terug te vinden is.
59. Indien mogelijk ook de inhoudelijk vaststeller voorafgaand aan de vergadering zijn commentaar schriftelijk laten geven.
60. Controleren of er binnen de vraagclusters waaruit vragen worden geselecteerd geen (te) grote verschillen in de vraagmoeilijkheid bestaan.
61. Als de vulgraad van de itembank dit toelaat, de vragen met een negatieve rit-waarde afschrijven. Indien dit niet het geval is, is het verstandig te onderzoeken of de vragen kunnen worden aangepast.
62. Doorvoeren van de achteraf bepaalde aanpassingen in de cesuur en criteria in de betreffende praktijkopdrachten voor toekomstige gebruik.
63. Een instructie voor beoordelaars opstellen met algemene beoordelingsregels en een voorbeeld van een op de juiste manier ingevuld beoordelingsformulier.
64. Een procedure ontwikkelen voor het aanvullen van de beoordelingsformulieren bij praktijkopdrachten die op examenlocaties worden gemaakt op basis van opmerkingen van beoordelaars.
65. In de instructie voor de praktijkopdrachten die thuis of op de werkplek worden gemaakt, vermelden welke bestandsformaten geüpload kunnen worden.
66. Een instructie voor afnameleiders op de afnamelocaties ontwikkelen met tips over het oplossen van problemen bij het beantwoorden van sleepvragen, het openen van hulpmiddelen en andere vaker voorkomende probleemsituaties (uit te voeren in overleg met Examenadviesburo).
67. Het gebruik van links met PDF-documenten en afbeeldingen zo veel mogelijk beperken. Vanaf TestVision Online versie 6 is het mogelijk voor kandidaten om afbeeldingen direct in de vraag te vergroten. Een aantal afbeeldingen kan daardoor alsnog direct aan de vraag worden toegevoegd.
68. In de instructie voor de praktijkopdrachten op de examenlocatie een indicatie geven voor de gewenste lengte van de uitwerking en de woordenteller bij de praktijkopdrachten aanzetten.
69. Formuleren van toetstermen en beoordelingscriteria voor schrijfvaardigheid.

70. In de instructie voor de praktijkopdrachten op de examenlocatie vermelden dat kandidaten geen teksten uit hulpmiddelen mogen knippen.
71. Bij de praktijkopdrachten die thuis of op de werkplek mogen worden gemaakt de instructiezin over het plegen van fraude uitbreiden met de informatie dat het beschikbaar stellen van de eigen uitwerking aan andere kandidaten ook als fraude wordt beschouwd.
72. De oefentoets voorzien van de genoemde 'probleemvragen' zodat de kandidaten beter zicht krijgen op het waarom van goede antwoorden.
73. Enkele voorbeelden van open vragen inclusief modelantwoord in de vorm van een PDF-document op de website van NVVB plaatsen.

6.5 Verbetersuggesties inhoudsdeskundige/toetsdeskundige

74. Verwerken van de resterende wijzigingen in het verzameldocument met toetstermen.
75. Vooraf beter nagaan of het mogelijk is om bij (set van) toetstermen geschikte praktijkopdrachten en eventuele parallelle versies van praktijkopdrachten te ontwikkelen.
76. Bij het beoordelen van een datum in een open vraag duidelijk vermelden of er wel of niet mag worden afgeweken van het format uit het antwoordmodel (bijvoorbeeld: 7 november 2013 mag ook worden ingevuld als 7-11-2013 of dit is juist niet toegestaan).
77. De vraagstelling van de open vragen in *HBO Identiteit en Identiteitsbewijzen II* aanscherpen op basis van de door kandidaten gegeven antwoorden. Dit kan het best worden uitgevoerd nadat TestVision Online 6 beschikbaar is, omdat in die versie de antwoorden makkelijker kunnen worden afgedrukt in combinatie met het beoordelingsvoorschrift. Een mogelijke verbetering is een paar correcte antwoorden als voorbeeld in de stam opnemen (bij vragen waar het aantal gevraagde antwoorden veel groter is dan het aantal mogelijke antwoorden).