



# Handboek Selectie

## Hoger Onderwijs



Sanne Schreurs, Sebastiaan Steenman,  
Ada Kool, Karen Stegers-Jager, Kitty Cleutjens  
en het OnderzoekersNetwerk Selectie

Dit handboek is mogelijk gemaakt door het Utrecht Stimuleringsfonds Onderwijs van het Centre for Academic Teaching and Learning van de Universiteit Utrecht.



**Universiteit  
Utrecht**

Een bijdrage vanuit de Expertgroep Toegankelijkheid Hoger Onderwijs (ETHO) via het Nationaal Regieorgaan Onderwijsonderzoek (NRO) heeft het mogelijk gemaakt de praktijk van het hbo beter in te bedden in dit handboek.

© 2023 Sanne Schreurs, Sebastiaan Steenman, Ada Kool, Karen Stegers-Jager en Kitty Cleutjens

Ontwerp: Taluut, Utrecht  
ISBN: 978-90-9037223-5



Leden van het **OnderzoekersNetwerk Selectie** die betrokken waren bij de totstandkoming van het handboek:

Timon de Boer – Universiteit Utrecht  
Kitty Cleutjens – Universiteit Maastricht  
Marc van Ekert – Universiteit Maastricht  
Eline Dubois – Leids Universitair Medisch Centrum  
Suzanne Fikrat-Wevers – Erasmus MC  
Simone Franssen – Universiteit Maastricht  
Renske van Gestel – Universiteit Utrecht  
Laura Koenders – Universiteit Utrecht  
Ada Kool – Universiteit Utrecht  
Anastasia Kurysheva – UMC Utrecht  
Claudy Oomen – Universiteit Utrecht  
Sanne Schreurs – Universiteit Utrecht & Universiteit Maastricht  
Sebastiaan Steenman – Universiteit Utrecht  
Karen Stegers-Jager – Erasmus MC  
Alicia Walkowiak – Universiteit Maastricht  
Anouk Wouters – Amsterdam UMC locatie VUmc

De student-assistenten van het project Select-UU waren onmisbaar voor (onder andere) het tot stand komen van dit handboek. Alle auteurs zijn hen veel dank verschuldigd.

Gijsbert van Elven  
Evelien van der Geest  
Janneke Jansink  
Janne Koekkoek





# Inhoudsopgave

1. Inleiding	7
<b>Deel 1: Kernoverwegingen en opzet van de procedure</b>	<b>15</b>
2. Alignment	17
3. Analyseren	25
4. Design (ontwerpen)	35
5. Develop (ontwikkelen)	43
6. Implementeren & Evalueren	49
<b>Deel 2: Instrumenten</b>	<b>59</b>
7. Koppeling criteria en instrumenten	61
8. Vooropleidingscijfers	67
9. Kennistoetsen	73
10. Capaciteitstoetsen	79
11. Persoonlijkheidsvragenlijsten	85
12. Motivatiebrieven	89
13. CV	95
14. Aanbevelingsbrieven	103
15. Situational Judgement Tests	111
16. Interviews	119
17. Psychomotorisch	125
18. Loting	129
19. Afsluiting	135
<i>Literatuur</i>	139
<i>Bijlagen</i>	148
<i>Verklarende woordenlijst</i>	161
<i>Over de auteurs</i>	163



# 1 Inleiding

Een goede selectieprocedure is van groot belang voor eerlijke toegang tot het hoger onderwijs. Het vormgeven van een goede selectieprocedure is echter geen gemakkelijke taak. Het vereist een goed inzicht in de doelen van de selectie en de eindtermen van de opleiding, (de kennis, vaardigheden en attitudes waarover studenten aan het einde van de opleiding moeten beschikken) als input om de selectie**criteria** te bepalen. Daarnaast vraagt het inzicht in de selectie-**instrumenten** die je kunt gebruiken om over die criteria in een selectie**procedure** informatie te verzamelen, waarbij je ook nog rekening moet houden met de legitimiteit van de procedure, de capaciteit van het onderwijs en de staf, de druk die kandidaten ervaren, risico's op ongewenste bias, et cetera. Concrete handvatten over de afwegingen die bij al deze vragen horen zijn dan ook van groot belang. Het doel van dit handboek is om elke stap in het vormgeven van een selectieprocedure te voorzien van handvatten, op basis van het onderzoek en de ervaringen op het gebied van selectie zoals die nu bekend zijn. Het boek richt zich daarbij primair op de mensen die bij het ontwerpen van de selectieprocedure betrokken zijn op opleidings- en instellingsniveau. Daarbij richt het boek zich op selectie in de breedte van het hoger onderwijs, zowel bachelor als master, en zowel hbo als wo. Elementen van het handboek zullen mogelijk ook breder voor selectie in het onderwijs bruikbaar zijn.

In dit handboek kiezen we er zeer bewust voor om niet voor te schrijven hoe een selectieprocedure eruit behoort te zien. Om twee redenen zou dat problematisch zijn. Allereerst zou dat ontkennen dat de context van een opleiding een grote rol speelt bij de inrichting van een selectieprocedure. De kennis, vaardigheden en attitudes die nodig zijn, zijn sterk afhankelijk van de inhoud en vorm van de opleiding. Een geneeskundeopleiding vraagt iets anders van studenten dan een opleiding bij een University College en een bachelor psychologie vraagt iets anders dan een master klinische psychologie. Ook het didactische model dat gebruikt wordt binnen de opleiding kan bepaalde vaardigheden vragen van studenten. Daarnaast verschilt natuurlijk

ook eventuele voorkennis die verondersteld wordt. Het is dan ook logisch dat de meest geschikte studenten voor de ene opleiding, niet direct het meest geschikt zijn voor de andere opleiding. Een goede selectieprocedure zal dat weerspiegelen. Studenten kunnen op verschillende manieren goed bij een opleiding passen, wat ook kan zorgen voor een diverse populatie studenten.

Een tweede reden waarom we geen eenduidige procedure voorschrijven voor een goede selectie is dat een deel van de afwegingen in het vormgeven van een selectieprocedure ook inherent normatieve componenten bevat waarvan het niet aan deskundigen of onderzoekers is om daarin keuzes voor te schrijven. Bijvoorbeeld de optie om loting als onderdeel van selectie toe te voegen. Dit vraagt om een afweging over wat een eerlijke manier is voor het verdelen van schaarse plekken in het hoger onderwijs. Wat eerlijk of rechtvaardig is komt niet uit empirisch onderzoek, hoewel een analyse van het debat laat zien dat er veel verschuivingen zijn in welke argumenten worden gebruikt (Steenman, 2018). We kunnen en zullen wel inzicht geven in afwegingen die daarbij een rol kunnen spelen. Verder vraagt de invulling van de procedure om afwegingen over hoe belangrijk bepaalde vaardigheden van kandidaten zijn, bijvoorbeeld in de weging van verschillende criteria. Ook die normatieve afwegingen kunnen we in dit handboek niet eenduidig maken, hoewel we wel meer inzicht geven in zowel relevante afwegingen als de gevolgen van keuzes. Veel van de keuzes vragen ook aandacht voor het onbedoeld benadelen van groepen kandidaten (bias, in het debat vaak ‘kansengelijkheid’ genoemd).

Dit handboek neemt ook verder expliciet geen normatieve positie in wat betreft het debat over de vraag of loting of selectie voor toegang tot het hoger onderwijs beter is. Het boek komt voort uit de vaststelling dat selectie nu gebruikt wordt voor programma’s met een beperkte instroom. Dit vraagt om een zorgvuldige en evidence-informed inrichting van de selectieprocedure. Het handboek is erop gericht om opleidingen daarbij te helpen. Daarnaast moeten opleidingen in bepaalde gevallen een keuze maken tussen loting en selectie of een combinatie van beide. Loting is daarom in dit handboek beschreven als mogelijk instrument. Ook hiervoor proberen we enkele relevante afwegingen in beeld te brengen zonder aan die afwegingen gewicht mee te geven.

Het handboek komt in eerste instantie voort uit het project Select-UU: een impuls voor de selectiepraktijk waarin voor opleidingen van de Universiteit Utrecht die voor toelating selecteren de afgelopen jaren onderzoek is gedaan naar de gebruikte instrumenten. Dat project en een groot deel van dit handboek zijn gefinancierd vanuit het Utrecht Stimuleringsfonds Onderwijs van de Universiteit Utrecht. Naast de inzichten uit dat project zijn ook de (inter)nationale literatuur en ander praktijkonderzoek naar selectie in dit handboek verwerkt. Met steun van NRO is de praktijk van het hbo ook beter ingebed in het handboek.





Veel van het onderzoek, zeker ook in de Nederlandse context, is uitgevoerd bij opleidingen in het veld van de gezondheidszorg (in het bijzonder geneeskunde). Als vanzelfsprekend gebruiken we de inzichten uit dat onderzoek ook in dit handboek. Het is echter een handboek dat geschreven is voor *alle* opleidingen in het (hoger) onderwijs die met selectie bezig zijn of dit overwegen. We vertalen de inzichten uit onderzoek dan ook altijd naar een bredere (en indien nodig Nederlandse) context. In voorbeelden zal bij instrumenten waar nog geen brede ervaring mee is opgedaan wel vaker een voorbeeld van een geneeskundeopleiding staan.

De vraag naar kennis over selectie leeft breed, bij verschillende opleidingen, bachelors, masters en aan de verschillende universiteiten en hogescholen in Nederland. De kennisbasis van dit handboek is dan ook gelegd door de onderzoekers die in Nederland met onderzoek naar selectie bezig zijn en die in het OnderzoekersNetwerk Selectie van Universiteiten van Nederland (UNL) kennis uitwisselen. In een startconferentie in oktober 2021 is voor alle instrumenten besproken wat de huidige stand van de kennis is, en welke conclusies op basis daarvan te rechtvaardigen zijn voor selectieprocedures in Nederland. Op basis van die bespreking zijn de onderdelen van het handboek uitgewerkt onder regie van de drie coördinatoren van het OnderzoekersNetwerk en het Select-UU-projectteam. In maart 2023 is een concept van dit handboek wederom in een conferentie met selectieonderzoekers besproken en naar aanleiding van de input in die bespreking aangepast.

De inhoud van dit handboek is in beweging: door nieuw onderzoek, door nieuwe regelgeving en door nieuwe ervaringen. Naast een papieren versie is daarom een digitale versie van dit handboek beschikbaar op het platform Onderwijskennis.nl van het Nationaal Regieorgaan Onderwijsonderzoek (NRO), die periodiek een update zal krijgen. Daarnaast werken we samen met NRO en de Expertgroep Toegankelijkheid Hoger Onderwijs (ETHO) aan de mogelijkheid om selectieprocedures laagdrempelig te evalueren met een digitale tool.

## Opzet van het handboek

Het handboek bestaat uit twee delen. In het eerste deel van het handboek gebruiken we een procesmodel, het ADDIE-model (Analysis, Design, Development, Implementation, Evaluation; Molenda, 2003), zoals dat ook vaak bij curriculumherzieningen wordt toegepast. Dit model helpt bij het geven van duidelijkheid en structuur aan het proces, en is bijvoorbeeld ook de basis van een recent gepubliceerde gids voor vormgeving van selectie specifiek voor geneeskundeopleidingen (Cleland et al., 2023). We benaderen dit proces alsof het gaat om het opzetten van een nieuwe selectieprocedure, maar de inzichten en stappen zijn evengoed relevant voor het evalueren en herzien van bestaande selectieprocedures. Figuur 1 laat dat ook zien.

We lichten in deze inleiding het ADDIE-model vast kort toe. Het eerste deel starten we in hoofdstuk 2 met het introduceren van het idee van *alignment*, of afstemming, tussen de selectieprocedure en de opleiding. Alignment is in dit handboek steeds een centrale afweging in de keuzes over de selectieprocedure. In de daaropvolgende hoofdstukken in deel 1 staan steeds onderdelen van het ADDIE-model centraal, en daarmee het ontwikkelingsproces van een selectieprocedure (hoofdstuk 3 t/m 6).

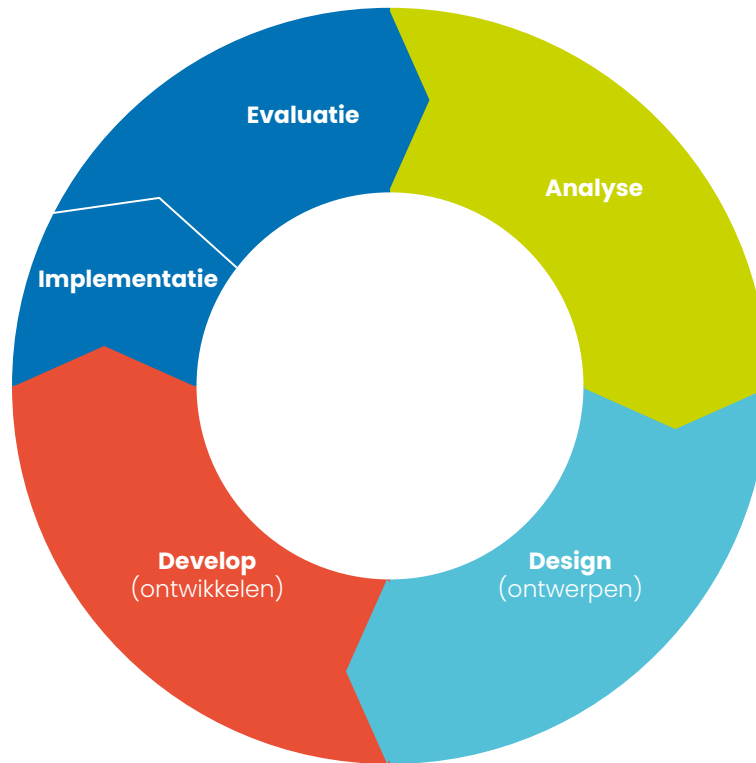
In het tweede deel van het handboek staan de verschillende mogelijke methoden of instrumenten die binnen een selectieprocedure gebruikt kunnen worden centraal: bijvoorbeeld het gebruik van resultaten van vooropleidingen of interviews. Daarbij is steeds aandacht voor wat het instrument is, enkele voorbeelden van het gebruik, belangrijke afwegingen voor het gebruiken van het instrument, hoe kandidaten tegen het instrument aankijken en wat we weten uit onderzoek. Daarnaast wordt telkens ingegaan op de vraag of bepaalde groepen kandidaten onbedoeld benadeeld worden door het (verkeerd) gebruik van het instrument (potentiële bias).

In het handboek hebben we vaak de keuze gemaakt om technische details en details over de praktische kant van de uitwerking weg te laten. Voor verdieping op elementen geeft de literatuurlijst mogelijke aanknopingspunten, maar we raden ook aan om bij andere opleidingen die bepaalde instrumenten hebben gebruikt te rade te gaan. De ervaring die zij hebben met de praktische uitvoering kan erg waardevol zijn. Ook expertise binnen de instelling kan helpen bij de uitwerking, bijvoorbeeld bij het ontwikkelen van toetsen. Daarnaast zijn de leden van het OnderzoekersNetwerk Selectie graag bereid om over verdiepende vragen mee te denken. Je kunt daarvoor contact opnemen met de coördinatoren van het netwerk: Kitty Cleutjens, Karen Stegers-Jager en Sebastiaan Steenman.

## Het ADDIE-model

ADDIE is een acroniem dat verwijst naar de belangrijkste fasen in een ontwikkel- of herzieningsproces: **A**nalyse, **D**esign, **O**ntwikkeling (**D**evelopment), **I**mplementatie en **E**valuatie. We bespreken de fasen kort een voor een, en gaan in de hoofdstukken 3 t/m 6 uitgebreider in op de verschillende onderdelen. Deze elementen geven structuur aan de rest van de bespreking, hoewel sommige stappen in de rest van het handboek meer aandacht zullen krijgen dan andere. Een samenvatting van de stappen voor het (her)ontwikkelen van de selectie is te vinden in figuur 1. Per stap is er ook een specifieke uitwerking opgenomen.





Figuur 1 • Het ADDIE-model

### Analyse

Het analyseren van de beginsituatie is de eerste stap voor het opzetten van een selectieprocedure. Hierbij overweeg je vooral wat het doel is van de selectieprocedure die ontwikkeld gaat worden: waarom wil je een selectieprocedure gaan opzetten, wat moet die toevoegen of veranderen aan de huidige situatie en wanneer ben je tevreden over de procedure? Of wat moet een aanpassing in de selectieprocedure opleveren?

Belangrijk in deze fase is dus om te onderzoeken:

- Waar is behoefte aan (waarom wordt selectie ingezet? Is het beperken van de instroom inderdaad passend bij de situatie?);
- Wat zijn de doelen van de selectie: wat wil je met je selectie bereiken en wanneer ben je er tevreden mee?
- Welke eisen zijn er aan datgene wat ontwikkeld gaat worden? Wat mag er (niet) vanuit wet- en regelgeving? Wat zijn de eisen vanuit de universiteit en de faculteit/opleiding die aan de procedure gesteld worden? Wat zijn de doelen en wat is de visie van het instituut? Hoe verhouden die elementen zich tot elkaar?
- Wat weten/kunnen de deelnemers aan de selectie al, wat mag je verwachten van de deelnemers aan je procedure (op welk niveau kun je de kandidaten testen, bijvoorbeeld uitgaan van het eindniveau van de middelbare school; Ab Latif & Mat Nor, 2020)?

**Design (ontwerpen)**

Na de analyse komt de eerste D van ADDIE: Design. In deze ontwerpfasen gebruik je de informatie die in de analyse is verzameld om een blauwdruk van de selectieprocedure te maken. Deze blauwdruk wordt gebruikt tijdens de ontwikkeling, en is bedoeld om ervoor te zorgen dat alle elementen van de selectie meten wat ze moeten meten. Daarmee helpt de blauwdruk om de vooraf opgestelde doelen te bereiken (Reinbold, 2013).

Het gaat daarbij om vragen als:

- Wat moet er in de selectieprocedure gemeten worden en waarom?
- Moet iedereen in dezelfde mate aan alle eisen voldoen, of is onderling compenseren ook goed? Hoeveel variatie tussen de vaardigheden van studenten is wenselijk of juist niet? Wat is de absolute ondergrens in vaardigheden, als die er is? En hoe verhoudt die ondergrens zich tot de toelaatbaarheidseisen voor de opleiding?
- Hoe bereid je kandidaten voor op de daadwerkelijke selectie, en welke voorbereiding verwacht je van kandidaten?
- En daarna: Welke instrumenten ga je voorleggen aan de kandidaten? Met welke (combinatie van) instrumenten kun je je doelen bereiken?

**Develop (ontwikkelen)**

Als het voorbereidende onderzoek- en denkwerk is voltooid, kom je bij de ontwikkelfase (development), de tweede D in het ADDIE-model. Dit is, zoals de naam al doet vermoeden, de fase waarin de selectie in zijn geheel wordt ontwikkeld nadat de opzet is vastgesteld. Er is nadrukkelijk overlap tussen de ontwikkeling en de ontwerp- en analysefasen. Het testen van de verschillende onderdelen (fasen en instrumenten) van de selectieprocedure hoort ook bij deze fase.

**Implementatie**

De implementatie behelst de daadwerkelijke uitvoering van de selectie. Hierbij observeer je ook vooral wat er gebeurt en hoe deelnemers aan het proces reageren (Ab Latif & Mat Nor, 2020). Dit is namelijk een belangrijke input voor de evaluatiefase.

**Evaluatie**

Eigenlijk is de evaluatiefase verweven met elke stap van het proces. In de laatste stap van de procedure vul je dat aan met een evaluatie van het hele proces en de uitkomsten van de selectieprocedure. Dit is input voor het verbeteren van de procedure voor de volgende keer dat er geselecteerd moet worden. Voor het herzien van de procedure zijn de bovenstaande stappen weer nuttig.





Alignment **2**

Analyseren **3**

Design (ontwerpen) **4**

Develop (ontwikkelen) **5**

Implementeren & Evalueren **6**

# Deel 1

## Kernoverwegingen en opzet van de procedure







## 2 Alignment

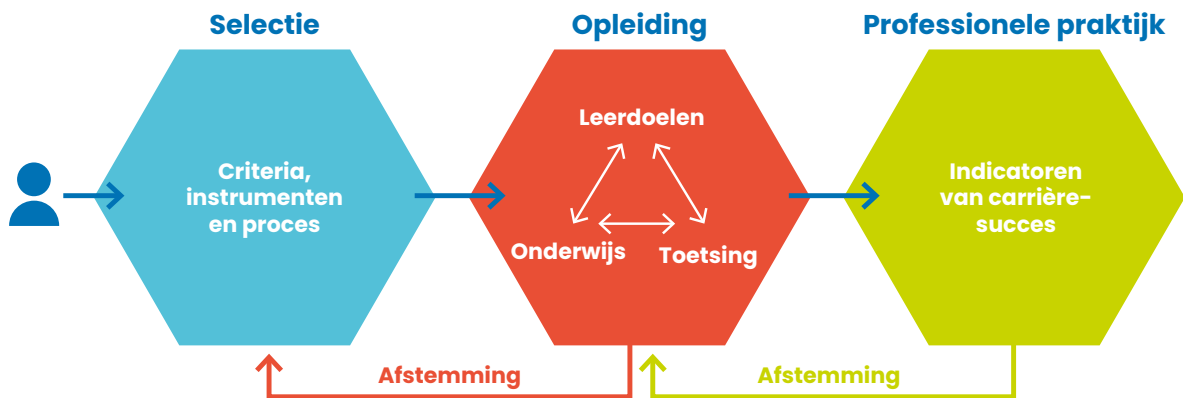
### Wat is alignment van selectie?

In dit handboek gebruiken we *alignment* als het centrale uitgangspunt voor het vormgeven van een selectieprocedure. Alignment draait om de afstemming tussen de selectieprocedure en de opleiding waarvoor geselecteerd wordt. Niet alle opleidingen die selecteren zijn op zoek naar dezelfde studenten; elke opleiding vraagt een specifieke combinatie van kennis, vaardigheden en attitudes. De eigenschappen van kandidaten die in een selectieprocedure centraal zouden moeten staan, zijn dus afhankelijk van de vorm en inhoud van de opleiding.

De term alignment van selectie is afgeleid van het idee van *alignment* binnen een opleiding zoals Biggs (1996) dat uitwerkt. Binnen een opleiding gaat het om *alignment* tussen leerdoelen, onderwijsleeractiviteiten en toetsing. *Alignment* van selectie is breder dan dat. Het gaat om afstemming van wat er in de selectie wordt gedaan met de vorm en inhoud van de opleiding. Dat wil zeggen dat de selectie (de criteria, de instrumenten en het proces) moet aansluiten op de leerdoelen, onderwijsleeractiviteiten en toetsing binnen een curriculum. In figuur 2 geven we dat aan de linkerkant aan. Verder leidt een opleiding ook op voor de professionele praktijk. Bij een goede alignment zal de opleiding een goede voorbereiding zijn op de professionele praktijk. Dit is te zien aan de rechterkant van figuur 2.

### Alignment tussen selectie en opleiding

Om een goede alignment van selectie te garanderen, is het van belang om eerst de vorm en inhoud van een opleiding in kaart te brengen. De vraag waarvoor je precies selecteert, is minstens zo belangrijk als hoe je selecteert. Daarbij hoort wat de leerdoelen van de opleiding zijn, welke onderwijsleeractiviteiten er in de opleiding worden gebruikt en hoe en wat er wordt



Figuur 2 • Alignment

getoetst. In een situatie waarin deze elementen op elkaar zijn afgestemd, spreken we over alignment binnen de opleiding. De selectieprocedure kan zich dan richten op de vaardigheden die belangrijk zijn voor de leerdoelen, onderwijsleeractiviteiten en toetsing van de opleiding. Dat is niet alleen een belangrijk uitgangspunt voor een selectieprocedure waarin de juiste elementen centraal staan, de selectieprocedure geeft in dat geval ook extra informatie over de vorm en inhoud van de opleiding aan kandidaten. Voor sommige opleidingen, waarbij kandidaten soms een verkeerd beeld hebben, kan die ‘matchingcomponent’ van een selectieprocedure die goed is afgestemd extra belangrijk zijn.

Een probleem met alignment binnen een opleiding of van selectie kan op verschillende manieren naar voren komen. Een opleiding kan bijvoorbeeld de vaardigheid kritisch redeneren erg belangrijk vinden en dit daarom onderdeel maken van de selectieprocedure. Echter, alle toetsen in de opleiding hebben vooral een sterke focus op kennis en begrip. In dat geval is de selectie niet goed afgestemd met wat er in de opleiding gevraagd wordt en zal de selectie ook niet voorspellend zijn voor de prestaties in de opleiding: er worden immers andere vaardigheden gevraagd in de selectie dan in de opleiding. Een vervolgstap kan zijn om de selectieprocedure aan te passen zodat die de focus op kennis en begrip weerspiegelt en daarmee past bij de opleiding. Aan de andere kant kan deze situatie ook leiden tot een aanpassing binnen de opleiding zelf waardoor kritisch redeneren terugkomt in de leerdoelen, onderwijsleeractiviteiten en toetsing. De selectie meet dan vaardigheden die ook in de opleiding belangrijk zijn. Op deze manier kan het nadenken over selectie, de doelen ervan en de *alignment* met de opleiding niet alleen tot een kwalitatief goede en voorspellende selectie leiden, maar is het ook een moment om op de inhoud en vorm van het onderwijs te reflecteren.

## Alignment tussen opleiding en professionele praktijk

Ook voor de professionele praktijk waartoe een opleiding opleidt, kunnen bepaalde eigenschappen in het bijzonder belangrijk zijn. Bij veel opleidingen waarvoor een duidelijke professionele praktijk af te bakenen is, spelen de kennis, vaardigheden en attitudes die voor die praktijk belangrijk zijn soms ook een rol in de selectie. Wanneer deze aspecten in de selectieprocedure worden meegenomen, zonder dat dit aansluit bij de vorm en inhoud van de opleiding, maakt dat het evalueren van de gebruikte selectie-instrumenten erg ingewikkeld. Een alternatieve manier om hier vorm aan te geven is door de afstemming tussen selectie en professionele praktijk in twee stappen te zien. De eerste stap is dat leerdoelen, onderwijsleeractiviteiten en toetsing binnen de opleiding zijn afgestemd op de professionele praktijk, de tweede stap is dat de selectie is afgestemd op de opleiding. Dat is het uitgangspunt zoals we dat in figuur 2 gebruiken.

## Alignment tussen selectie en opleiding: focus op vaardigheden

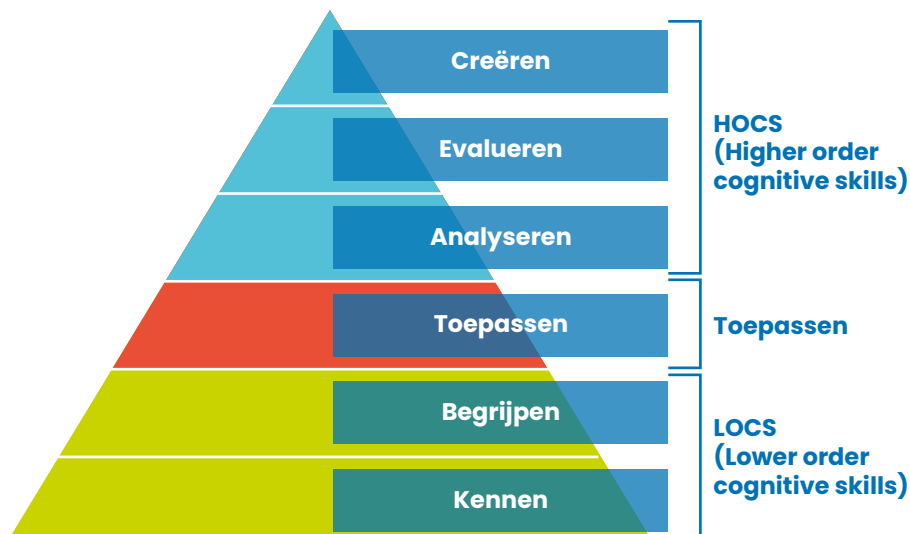
Als we het hebben over alignment van selectie, kan dat betrekking hebben op kennis, vaardigheden of attitudes (samen competenties). Bij kennis gaat *alignment* vooral om de voorkennis die van studenten verwacht wordt. Met name bij masteropleidingen is dat een belangrijk criterium voor toelaatbaarheid. Voor bacheloropleiding is dat doorgaans een minder duidelijk onderdeel van het toelatingsproces, omdat voorkennis doorgaans vervat is in de eisen aan de vooropleiding (mbo, havo of vwo) al dan niet aangevuld met specifieke eisen aan het vakkenpakket. Bij dat laatste speelt *alignment* natuurlijk wel een rol: zijn de eisen aan het vakkenpakket passend bij wat er in de opleiding van studenten verwacht wordt qua voorkennis? Omdat *alignment* van kennis meestal gaat over of iemand toelaatbaar is en minder over selectie, vormt dit onderdeel van het toelatingsproces in het handboek verder geen centraal element.

In dit handboek ligt de focus voornamelijk op de vaardigheden die studenten nodig hebben om succesvol te kunnen zijn binnen een opleiding. We gebruiken de term vaardigheden daarbij veel breder dan vaak gebeurt. Vaak wordt de term vaardigheden in het hoger onderwijs gebruikt voor professionele vaardigheden als presenteren of samenwerken. Wij gebruiken de term ook voor cognitieve vaardigheden. Dat type vaardigheid speelt in selectieprocedures vaak een belangrijke rol.

## Drie domeinen van vaardigheden binnen een selectieprocedure

In grote lijnen houden we het onderscheid tussen drie domeinen van vaardigheden aan zoals Bloom (1956) dat maakt: het cognitieve domein, het affectieve domein en het psychomotorische domein. Hieronder gaan we kort in op de verschillende domeinen en de categorisering van vaardigheden die binnen de domeinen horen.

Binnen het cognitieve domein zijn verschillende vaardigheden te onderscheiden. Elk van die vaardigheden vraagt iets anders van een persoon, hoewel ze ook met elkaar te maken hebben. Kennis kunnen toepassen vergt een andere cognitieve vaardigheid dan kennis onthouden. Natuurlijk heeft kennis kunnen onthouden en begrijpen wel verband met kennis kunnen toepassen. In het onderscheiden van de verschillende vaardigheden in het cognitieve domein baseren we ons op de indeling van Anderson en Krathwohl (2006) die is gebaseerd op het werk van Bloom (1956). Anderson en Krathwohl (2006) maken onderscheid tussen zes niveaus van complexiteit in cognitieve vaardigheden: kennen, begrijpen, toepassen, analyseren, evalueren en creëren. In dit handboek hanteren we een iets grovere indeling dan Anderson en Krathwohl doen, omdat dit voor het ontwikkelen van selectieprocedures voldoende handvatten biedt. In het handboek maken we onderscheid tussen *lagere orde cognitieve vaardigheden* (of LOCS, van Lower Order Cognitive Skills), waar kennen en begrijpen onder vallen; *toepassen* (dat in veel opleidingen een belangrijke rol speelt); en *hogere orde cognitieve vaardigheden* (of HOCS van Higher Order Cognitive Skills), waaronder analyseren, evalueren en creëren vallen. Zie ook figuur 3.



Figuur 3 • Vaardigheden in het cognitieve domein

Het affectieve domein beschrijft vaardigheden op het gebied van emotionele reacties van mensen en hun vermogen om pijn of vreugde van andere mensen te voelen. Affectieve doelen zijn gericht op het bewustzijn en de groei van houding, emoties en gevoelens. De volgende vijf niveaus worden onderscheiden: ontvangen, reageren, waarderen, organiseren en karakteriseren. In het psychomotorische domein gaat het over het kunnen verrichten van motorische bewegingen en handelingen.

Het uitgangspunt voor elke vaardigheid, of deze nu valt onder het cognitieve, affectieve of psychomotorische domein, is dat moet worden afgewogen wat het gewenste startniveau van kandidaten moet zijn bij de selectieprocedure. Van kandidaten kan niet verwacht worden dat zij al voldoen aan de einddoelen van de opleiding; dan zou de opleiding immers niet meer nodig zijn. Juist tijdens de opleiding is er ruimte om vaardigheden verder te ontwikkelen. Voor sommige vaardigheden is het echter wel van belang dat studenten al een bepaald startniveau hebben. Bijvoorbeeld als de vaardigheid vanaf de begin van de opleiding belangrijk is of doordat het hebben van een bepaald startniveau helpt bij het verder opbouwen van de vaardigheid en daardoor leidt tot een hoger niveau aan het eind van de opleiding. Als de vaardigheid meegenomen wordt in de selectie moet goed worden nagedacht over het niveau dat verwacht kan worden van kandidaten. Hier gaan we in [hoofdstuk 3](#) verder op in.

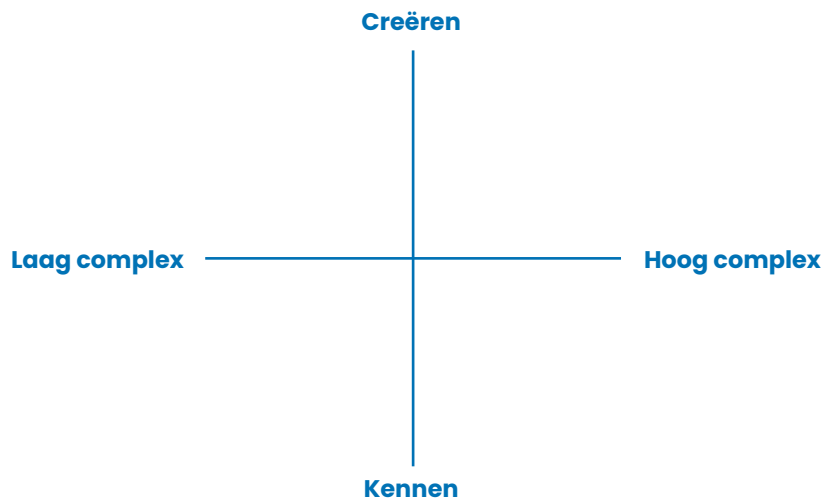
Het meenemen van alle relevante vaardigheden, of in elk geval de belangrijkste, in de selectie, maakt het mogelijk dat kandidaten met verschillende eigenschappen een plek kunnen krijgen in de opleiding. Niet alle kandidaten bezitten alle vaardigheden in dezelfde mate; sommigen zijn bijvoorbeeld minder goed in het consciëntieus aanleren van kennis, maar juist wel goed in analyseren en kritisch evalueren van kennis. Als allebei die vaardigheden in de opleiding belangrijk zijn, is het logisch dat beide terugkomen in de selectie. Een kandidaat die op een van die vaardigheden hoog scoort, maakt dus goede kans om beter uit de selectieprocedure te komen dan anderen. Juist door een divers palet aan relevante vaardigheden mee te nemen in de selectie, maakt een diverse groep kandidaten kans op een plek in de opleiding.

### Complexiteit van vaardigheden: aansluiting in selectie

De rangorde van Bloom gaat ervan uit dat sommige vaardigheden in zichzelf complexer zijn dan andere (analyseren bijvoorbeeld is ingewikkelder dan kennen). Daarnaast kun je binnen elke vaardigheid verschillende niveaus van complexiteit hebben: je kunt kennis opdoen van een paar simpele begrippen, of van complexe systemen. Op dezelfde manier kun je twee simpele definities met elkaar vergelijken en daar de verschillen uithalen (analyseren) of

bijvoorbeeld in een scriptie complexe theoretische stromingen met elkaar laten ‘praten’. De manier waarop opleidingen complex zijn, verschilt dan ook van elkaar. Veel hbo-opleidingen zijn vooral complex in het toepassen van kennis in praktijksituaties, waarbij bijvoorbeeld een academische opleiding meer hogere orde cognitieve vaardigheden vraagt en minder complexe toepassing in de praktijk. Figuur 4 geeft die verschillende dimensies van complexiteit in het cognitieve domein aan.

Binnen elke vaardigheid heb je dus te maken met een schaal van oplopende complexiteit. Dit is essentieel om bij je selectie over na te denken: je kunt de vaardigheid die je belangrijk vindt al in de selectie meenemen, maar op een lager niveau van complexiteit dan in de opleiding zelf. Een voorbeeld: stel je vindt het heel belangrijk dat je kandidaten kunnen evalueren. Je kunt dan relatief simpele kennis die kandidaten aan het begin van het eerste jaar op zouden doen (bijvoorbeeld laagdrempelige theorieën in sociale psychologie, zoals het omstandereffect) aan de kandidaten meegeven, en ze vragen hier vergelijkingen tussen te maken en een keuze te onderbouwen. Zo kun je relatief onafhankelijk van de inhoud wel de cognitieve vaardigheid meten.



*Figuur 4 • Complexiteit in het cognitieve domein*

Voor de selectieprocedure is het zaak om de vaardigheden te identificeren die in de opleiding belangrijk zijn en waarvan van kandidaten wordt verwacht dat ze op een hoog niveau van complexiteit eindigen (bijvoorbeeld toepassen of analyseren). Vervolgens gebruik je selectie-instrumenten die deze vaardigheden toetsen, maar op een lagere complexiteit (dus op dezelfde hoogte in figuur 4, maar meer naar links). Het doel is om kandidaten te selecteren die in potentie een hoog eindniveau zouden kunnen behalen aan het eind van de studie. Uitgangspunt daarbij is dat hoe beter je je selectie kunt laten lijken op

hoe de vaardigheid terugkomt in je opleiding, hoe beter je die voorspelling kunt maken (in het verleden behaalde resultaten bieden misschien geen garantie, maar wel de beste voorspelling van de toekomst). Dat is de reden dat onderzoek vaak laat zien dat toetsen die lijken op wat er in het eerste jaar van een opleiding gebeurt doorgaans goede voorspellers zijn voor de studieresultaten in jaar 1: goed sprinten in de training voorspelt goed sprinten tijdens een wedstrijd, goed gemaakte kennistoetsen op de middelbare school voorspellen goed gemaakte kennistoetsen in de opleiding.

Een nadeel aan het meenemen van cognitieve vaardigheden in je selectie is het feit dat niet alle kandidaten gelijke kansen hebben gehad om vaardigheden te ontwikkelen. Het volgen van dezelfde vooropleiding betekent niet automatisch dat omstandigheden waarbinnen prestaties geleverd zijn en vaardigheden ontwikkeld zijn voor iedereen ook gelijkwaardig zijn. Zo heeft een leerling die extra (ingekochte) hulp en bijles heeft gekregen meer mogelijkheden gehad om cognitieve vaardigheden te ontwikkelen dan een leerling die naast de opleiding noodgedwongen veel uren moest werken en daardoor minder tijd aan school kon besteden.

Sternberg (2012) laat zien dat als je andere vaardigheden dan kennis en begrip op een goede en gepaste manier meeneemt in je selectie, je meer diverse kandidaten de mogelijkheid geeft om uit te blinken in je selectieprocedure. Juist vaardigheden die in het voortgezet onderwijs minder dominant zijn, kunnen hieraan bijdragen. Hij heeft het over ‘praktische intelligentie’ en ‘wijsheid’ – in onze terminologie zou dat gaan in de richting van toepassen, HOCS en vaardigheden in het affectieve domein.

In het vervolg van dit eerste deel van dit handboek werken we uit wat dit idee van *alignment* betekent voor de verschillende fasen binnen ADDIE en voor het ontwikkelen en evalueren van een selectieprocedure. In het tweede deel van het handboek bepaalt het principe van *alignment* de hoofdstukindeling. We bespreken daar de instrumenten die zich het meest lenen voor verschillende typen vaardigheden.





## 3 Analyseren

Het analyseren van de beginsituatie (waarom wil je selecteren en wat is daarin dan belangrijk) is de eerste stap als je een selectieprocedure gaat opzetten, om zo alignment tussen selectie en opleiding te kunnen waarborgen. Hierbij ga je na wat de doelen van de selectie moeten zijn, wanneer je tevreden bent, en of een selectie er inderdaad voor zorgt dat die doelen behaald worden. Zoals eerder beschreven: de analysefase vraagt vooral uitzoekwerk, achtergrondonderzoek en informatieverzameling (Reinbold, 2013).

Belangrijk om te analyseren rondom selectie zijn de onderwerpen die ook in de cirkelfiguur genoemd worden:

- Is selectie echt nodig? Wat wil je precies bereiken en is selectie daar wel de beste aanpak voor?
- Waar is behoefte aan? (waarom wordt selectie ingezet: wil je instroom beperken of focussen op de beste studenten?);
- Wat zijn de einddoelen van de selectie: wat wil je met je selectie bereiken en wanneer ben je er tevreden mee? Bijvoorbeeld uitval verminderen, prestaties verbeteren, of een betere aansluiting bij het werkveld.
- Welke eisen zijn er aan de te ontwikkelen selectieprocedure (vanuit wet- en regelgeving)? Wat mag er (niet)?
- Wat zijn de eisen vanuit de universiteit en de faculteit/opleiding die aan de procedure gesteld worden? Wat zijn de doelen en wat is de visie van het instituut? Hoe verhouden die zich tot elkaar?
- Wat weten/kunnen de deelnemers aan de selectie al, wat mag je verwachten van deelnemers aan je procedure (op welk niveau kun je de kandidaten testen, bijvoorbeeld uitgaan van het eindniveau van de middelbare school; Ab Latif & Mat Nor, 2020)?



Figuur 5 • Analyseren

### Selectie?

Wat is de grondslag van je selectieprocedure? Is er een beperkt aantal plaatsen, of selecteer je op geschiktheid?

### Doelen van de selectieprocedure

Denk na over wat je belangrijk vindt om te meten in en te bereiken met de selectieprocedure. Hierbij hoort bijvoorbeeld ook de vraag welke vormen van diversiteit van belang zijn.

### Alignment

Wat zijn de vorm en inhoud van de opleiding zelf en wat betekent dat voor wat je verwacht van studenten, en daarom ook moet meten in de selectieprocedure?

### Wat mag en kan?

Wat is de relevante juridische context voor de procedure? Hoe ziet de verantwoording eruit? Hoe verhoud je je tot het beleid van de instelling? En wat is de financiële en organisatie context?

### Wat kun je vragen?

Wat kun je inhoudelijk en qua tijdsinvestering van kandidaten verwachten?

## Waarom selectie?

Er zijn verschillende vormen van selectie die gebruikt worden in Nederland en die een andere grondslag hebben. Voor bacheloropleidingen is de *numerus fixus* de meest voorkomende. Bij een *numerus fixus* wordt er een maximum aan het aantal studenten gesteld. Meestal is dit om de kwaliteit van de opleiding te kunnen waarborgen, maar vragen vanuit de arbeidsmarkt kunnen ook een rol spelen. Voorbeelden van *fixus*opleidingen zijn International Business, Mondzorgkunde, Geneeskunde en Psychologie. De reden voor selectie is in dit geval dus niet de kwaliteit of de passendheid van kandidaten, maar de noodzaak vanwege het beperkte aantal plekken ten opzichte van het aantal geïnteresseerde kandidaten.

Deze opleidingen willen natuurlijk wel dat studenten die binnenkomen zo 'goed' mogelijk passen (we gaan later verder in op hoe je daarover zou

kunnen nadenken). Tegelijkertijd accepteren deze opleidingen ook dat ze geschikte kandidaten afwijzen. De reden dat deze opleidingen selecteren ligt immers in de beperking van de instroom en niet in de kwaliteit van de kandidaten. Een belangrijk streven van deze opleidingen is dan ook dat ze, naast zo goed mogelijk selecteren, de fixus willen vullen.

Daarnaast zijn er opleidingen die *'hard' selecteren*. Dat betekent dat zij selecteren zonder dat er een vastgesteld maximum is. Zij wijzen dus kandidaten af die niet geschikt zijn voor de opleiding. Dit speelt vooral een rol bij masteropleidingen. Daarnaast is dit formeel het geval bij opleidingen met het Bijzonder Kenmerk Kleinschalig en Intensief onderwijs (BKKI), zoals de University Colleges en enkele specifieke hbo- en wo-opleidingen. In de praktijk heeft de kleinschaligheid van de opleiding tot gevolg dat ook daar meestal alleen de meest geschikt geachte kandidaten worden geselecteerd. Ook opleidingen met andere specifieke kenmerken, die bepaalde vaardigheden van kandidaten nodig hebben (zoals kunstopleidingen of de opleiding tot luchtverkeersleider) kunnen op die kenmerken selecteren.

Een opleiding mag in het Nederlandse stelsel dus nooit 'zomaar' besluiten om te selecteren. De selectie hangt altijd samen met de kenmerken van de opleiding of de arbeidsmarkt. De capaciteiten voor praktijkonderwijs, stageplaatsen, maar ook het onderwijsmodel zullen doorgaans een rol spelen. Het instellen van een numerus fixus is op dit moment nog een besluit dat het College van Bestuur van een instelling zelf kan nemen. Op het moment van schrijven zijn er echter serieuze bewegingen om dat in wetgeving aan te passen. Selecteren op basis van het Bijzonder Kenmerk Kleinschalig en Intensief onderwijs (BKKI) vraagt een accreditatieproces door de Nederlands-Vlaamse Accreditatieorganisatie (NVAO) en een besluit van de minister. Voor masteropleidingen is er in de huidige wetgeving veel ruimte met betrekking tot de wijze van selecteren, en is 'harde' selectie standaard toegestaan (natuurlijk kan binnen een instelling ander beleid gelden). Op nationaal niveau is tussen universiteiten wel afgesproken dat alle studenten met een Nederlands bachelordiploma toegang hebben tot ten minste één masteropleiding aan een Nederlandse universiteit.

Er is een wezenlijk verschil tussen selecteren voor een opleiding die een beperkt aantal plekken heeft en selecteren voor een opleiding die dat niet heeft. Bij het selecteren voor een beperkt aantal plekken is het uitgangspunt, zoals we hierboven al aangaven, dat niet alle geschikte studenten een plek kunnen krijgen. Het streven is dan om de voor die opleiding best passende kandidaten een plek te geven, waarbij je moet kunnen onderbouwen waarom de ene kandidaat beter geschikt is dan de andere. Bij het selecteren voor een opleiding die geen beperking van het aantal plekken heeft (harde selectie), richt het selecteren zich op de geschiktheid van de kandidaat. Het uitgangspunt is dan dat je moet onderbouwen waarom afgewezen kandidaten

de opleiding niet succesvol door kunnen komen. Dat is een verstrekkende uitspraak over de vaardigheden van een kandidaat.

### Waar is behoefte aan?

Of het primaire doel van de selectie nu is om de instroom te beperken of niet, de wens dat de juiste student op de juiste plek komt zal bij alle opleidingen die selecteren de eerste prioriteit zijn. Wat echter maakt dat een student op de goede plek zit, verschilt van opleiding tot opleiding. Het kan gaan om inhoudelijke interesse, vaardigheden of voorkennis. Het ligt sterk voor de hand hierover breed in de opleiding input op te halen, zeker ook van degenen die nauw betrokken zijn bij de huidige studenten (zoals tutores en werkgroepdocenten in het eerste jaar, maar ook de studieadviseur en natuurlijk de studenten zelf).

### Doel van de selectieprocedure

Het in kaart brengen van de behoeften binnen de opleiding geeft belangrijke input voor het formuleren van de doelstelling voor de selectieprocedure. Vanuit het instellings- of opleidingsbeleid liggen aanvullende overwegingen voor de hand. Dit kan bijvoorbeeld zitten in de prioritering van de behoeften, maar ook in de vorm van doelstellingen om specifieke groepen geen structureel nadeel te laten ondervinden van selectieprocedures, om de werklust van kandidaten en de organisatie te beperken, of om uitval binnen de opleiding te verminderen. Dit zijn allemaal elementen die veel invloed kunnen hebben op de opzet van de selectieprocedure. Het is dus belangrijk om die aan de start van het proces helder te formuleren en te zorgen dat daarvoor op de juiste plekken draagvlak is. Schrijf de doelen en overwegingen ook vooral uit: in de rest van het ontwerpproces is dat een belangrijk document om op terug te vallen.

In specifieke situaties kan het daarnaast belangrijk zijn om in de selectieprocedure voorbij de opleiding te kijken naar wat er verder in de toekomst van kandidaten verwacht wordt. Bij geneeskundeopleidingen is (afhankelijk van het curriculum) voor de bachelor vaak een procedure gefocust op kennis en begrip goed passend, maar richting de klinische master neemt het belang van toepassen en affectieve of (inter)persoonlijke vaardigheden sterk toe (Patterson, Knight et al., 2016; Schreurs, Cleutjens, Cleland & oude Egbrink, 2020; Schreurs, Cleutjens, Muijtjens, Cleland & oude Egbrink, 2018). Daarbij komt dat een grote meerderheid van de bachelorstudenten geneeskunde direct doorstroomt naar de master aan dezelfde instelling. Daarmee wordt het een belangrijke overweging of binnen de selectie voor de bachelor geneeskunde de vaardigheden die in de master belangrijk worden niet al

een rol zouden moeten spelen (Kuncel, 2011). Die latere competenties zijn daarnaast meer in lijn met wat de arbeidsmarkt van afgestudeerden vraagt.

### **Diversiteit**

Vanuit het perspectief van diversiteit is het vooral belangrijk om zorgvuldig af te wegen of een selectieprocedure systematisch kandidaten met bepaalde achtergronden benadeelt. Selectie maakt natuurlijk altijd onderscheid tussen kandidaten, maar dat onderscheid zou gebaseerd moeten zijn op de inhoudelijke criteria en de doelen van de selectieprocedure en niet op andere eigenschappen van kandidaten (Zwick, 2017). Het kan bijvoorbeeld gaan om onderscheid op basis van relevante cognitieve vaardigheden, maar niet op basis van de sociaal-economische status van het gezin van kandidaten.

In de discussie rondom loting en selectie en het uitgangspunt van kansengelijkheid of mogelijkheidengelijkheid speelt de angst een rol dat bepaalde groepen door zelfselectie zich al minder vaak zullen aanmelden voor selecterende opleidingen. Er zijn op dit moment echter geen duidelijke aanwijzingen voor structurele zelfselectie op basis van het instellen van een selectieprocedure alleen. Voor de meeste achtergrondkenmerken van kandidaten die de onderwijsinspectie als ‘risicogroep’ omschrijft geldt niet dat deze kandidaten zich minder snel aanmelden voor een opleiding die selecteert (Inspectie van het Onderwijs, 2022). De keuze voor selectie lijkt de diversiteit in de studentpopulatie dan ook niet aanvullend in de weg te hoeven staan, maar dit vraagt wel om een zorgvuldig en op basis van wetenschappelijke kennis ingevulde procedure. In dit handboek staan we daarom in het tweede deel bij alle instrumenten expliciet stil bij de potentiële bias en geven we handvatten om die te voorkomen.

Als het streven naar een diverse studentpopulatie belangrijk wordt gevonden, dan lijkt het bij selecteren voor de hand te liggen om te zorgen voor een combinatie van selectie-instrumenten die een breed palet aan vaardigheden die belangrijk zijn voor de opleiding meten, en daarbij in elke stap alert te zijn op bias. Dit betekent dat studenten met verschillende vaardigheden goed kunnen presteren op verschillende onderdelen van de selectieprocedure en zo toegang tot de opleiding kunnen krijgen. Tijdens de ontwikkeling en evaluatie van de betreffende instrumenten is aandacht voor verschillen in culturele achtergrond, de nadelen die specifieke groepen kandidaten kunnen hebben wanneer een instrument wordt gebruikt en differentiële voorspelende waarde daarbij van groot belang.

De Erasmus Universiteit in Rotterdam geeft kandidaten die zich willen voorbereiden op de decentrale selectieprocedure voor geneeskunde de mogelijkheid om een gratis coachingsdag bij te wonen. Deze coachingsdag is in het leven geroepen omdat kandidaten zich willen voorbereiden op de decentrale selectieprocedure en niet iedere student de prijzige trainingen van particuliere aanbieders kan of wil betalen.

De Erasmus Universiteit Rotterdam wil met de coachingsdag iedereen een gelijke kans op voorbereiding gunnen en onderstreept hiermee het belang van een diverse instroom. Uit onderzoek blijkt dat deelname aan de coachingsdag specifiek goed uitpakt voor de selectie-uitkomsten van mannelijke kandidaten en kandidaten met een niet-westerse migratieachtergrond en daarmee dus mogelijk kan bijdragen aan diversiteit (Fikrat-Wevers, Stegers-Jager, Afonso et al., 2023).

## Alignment

Naast het analyseren van de bredere doelen van selectie, is de belangrijkste stap in deze fase het analyseren van waar je selectie bij moet aansluiten: de opleiding zelf. In het vorige hoofdstuk gingen we al in op het belang van de afstemming of *alignment* tussen selectie en opleiding waarin de geselecteerde terecht komt (Steenman, 2018). De kern van wat er van studenten gevraagd wordt verandert vaak als studenten verder in de opleiding komen. In de verschillende fasen van de opleiding nemen andere complexe vaardigheden de overhand, en in de *alignment* van de selectie is het dan ook belangrijk om te expliciteren waarmee je de selectie in lijn wilt brengen (Schreurs, 2020).

De aanname is dat de opleiding aansluit bij vervolgoopleidingen en het beroepenveld. In dit handboek zal dus de aansluiting van de selectie bij de opleiding zelf centraal staan. In specifieke situaties kan het voor de hand liggen om toch verder te kijken (zoals we hiervoor bespraken voor de opleiding geneeskunde). In die situaties zou een zogenaamde ‘job analysis’ behulpzaam kunnen zijn (Patterson, Ferguson & Thomas, 2008).

Voor wat belangrijk is aan vaardigheden binnen een opleiding ligt het voor de hand om te kijken naar de leerdoelen van de opleiding zoals die in de Onderwijs- en Examenregeling (OER) staan vermeld, en naar de leerdoelen van de afzonderlijke cursussen. Ook kijken naar wat er in toetsing binnen de opleidingen eigenlijk echt van studenten gevraagd wordt is vaak erg behulpzaam. Wordt van studenten vooral gevraagd dat ze kennis en begrip opdoen en reproduceren? Moeten ze kennis toepassen of juist snel kritisch redeneren en zelf kennis genereren? Zijn affectieve of psychomotorische vaardigheden (bewuste, vanuit de hersenen gestuurde beweging, denk bijvoorbeeld aan oog-handcoördinatie) van belang binnen de opleiding? In deze fase is het belangrijk om dat goed in kaart te brengen, zodat je in de selectieprocedure

de vaardigheden die het meest belangrijk zijn in de opleiding op een simpeler manier een plek kunt geven (zie ook [hoofdstuk 2](#)).

## Wat mag en kan?

### Juridische context

Wat er wel en niet mag in een selectieprocedure wordt natuurlijk allereerst bepaald door de Nederlandse wetgeving. Die wetgeving laat op dit moment relatief veel toe. Er staat dat selectie plaatsvindt op basis van kwalitatieve criteria, en dat in de selectie ten minste twee soorten kwalitatieve criteria moeten worden gebruikt (artikel 7.53, tweede lid Wet op het hoger onderwijs en wetenschappelijk onderzoek (WHW) en 7.30b WHW voor de master). Kwalitatieve criteria is een lastige term, die eraan refereert dat de criteria moeten gaan om kwaliteiten van de kandidaat. De memorie van toelichting bij dit artikel geeft wel aan dat niet beide criteria gebaseerd mogen zijn op cijfers uit het voortgezet onderwijs.

De memorie van toelichting geeft als voorbeeld dat een van de criteria cognitief gericht zou kunnen zijn, en gebaseerd zou kunnen zijn op de cijfers in het voortgezet onderwijs, en dat het andere criterium dan niet-cognitief zou kunnen zijn. Nog los van de vraag wat niet-cognitief precies betekent, is dit geen verplichting die de wet oplegt. Beide criteria mogen ook betrekking hebben op cognitieve vaardigheden.

Inmiddels is het wetsvoorstel om de mogelijkheden van opleidingen met een numerus fixus (dus niet voor masteropleidingen of voor opleidingen met een BKKI of een ander speciaal kenmerk om op te selecteren) uit te breiden met de optie om ongewogen te loten, of om twee kwalitatieve criteria en loting te combineren, aangenomen door de Tweede Kamer. Dat kan zowel betekenen dat een deel van de studenten wordt geselecteerd en een deel ongewogen wordt geloot, of dat er gewogen wordt geloot waarbij de weging wordt bepaald door twee kwalitatieve criteria. Hiermee zullen de mogelijkheden voor de manieren waarop selectie vormgegeven kan worden veel breder worden. In [hoofdstuk 18](#) werken we uit wat het huidige wetsvoorstel zou betekenen als het in werking zou treden.

Andere overwegingen binnen de juridische context zijn bijvoorbeeld het omgaan met functiebeperkingen. Het is verplicht om rekening te houden met functiebeperkingen, en een algemene vuistregel die gebruikt wordt is dat kandidaten bij de selectie dezelfde faciliteiten krijgen die ook tijdens de opleidingen geboden worden, hoewel opleidingen hierin mogelijk ook andere afwegingen kunnen maken, bijvoorbeeld omdat de omstandigheden anders zijn dan tijdens de opleiding.

Een laatste belangrijke voorwaarde is dat er een specifieke verantwoordelijkheid is in het faciliteren van de selectieprocedure voor studenten die wonen in de Caraïbische delen van Nederland en van het Koninkrijk. Een aantal universiteiten lost dit bijvoorbeeld op door selectiedagen/-toetsen op locatie aan te bieden, of door voor deze groep online opties te geven.

Naast deze eisen aan de manier waarop de selectie vorm wordt gegeven, zijn er ook vereisten aan de procedure voor het vaststellen van de selectieprocedure en de inspraak van medezeggenschap daarbij. Het vaststellen van de wijze van selecteren gebeurt voor bacheloropleidingen anders dan voor masteropleidingen. Voor masteropleidingen moet de selectieprocedure opgenomen worden in de Onderwijs- en Examenregeling (OER) van de opleiding. De opleidingscommissie heeft adviesrecht over de kwalitatieve toelatingscriteria en de faculteitsraad over de selectieprocedure. Voor bacheloropleidingen moet het College van Bestuur van de instelling de selectieregeling vaststellen, waarbij de Universiteitsraad adviesrecht heeft. Bespreek tijdig binnen je instelling hoe medezeggenschap in dat proces is geregeld.

### **Verantwoording**

Denk binnen de formele context ook na over de verantwoording van de gehanteerde selectieprocedure. De onderwijsinspectie speelt in die verantwoording een formele rol, en voor BKKI-opleidingen ook de visitatiepanels bij het verkrijgen van het kenmerk. Het belang van een goed te verantwoorden selectie lijkt ook in het politieke debat een steeds grotere rol te spelen, en de verplichtingen op dat vlak voor opleidingen zouden goed kunnen toenemen. Daarnaast is het natuurlijk belangrijk om naar stakeholders (studenten en kandidaten, maar ook docenten, het werkveld en de samenleving) goed te kunnen verantwoorden wat je doet.

Daarbij kan evalueren een belangrijke rol spelen. Bij het ontwerpen van een selectieprocedure is het daarmee belangrijk om vast na te denken over hoe je straks kunt nagaan of de procedure goed werkt. Het coalitieakkoord spreekt over de volgende kwaliteitscriteria die daarbij waarschijnlijk een belangrijke rol kunnen spelen, en die sterk aansluiten bij het idee van *alignment* zoals dat in dit handboek centraal staat: “de selectieprocedure past bij de inhoud van de opleiding, [is] effectief en [borgt] gelijke kansen.”

### **Instituut (universiteit/hogeschool, faculteit, opleiding)**

Veel instellingen hebben eigen selectiebeleid. Vaak verschilt dat voor bachelor- of masteropleidingen en voor de verschillende grondslagen om te selecteren (vanwege een numerus fixus, omdat kandidaten specifieke vaardigheden moeten hebben voor een master, kunstopleiding of BKKI-opleiding, etc.). Ook is het van belang dat selectie onderdeel is van het geheel van de visie op het onderwijs binnen de instelling, faculteit en opleiding. Zorg dus dat de selectie aansluit bij de didactische visie en eventueel ook de



missie van het instituut. Als probleemgestuurd onderwijs bijvoorbeeld een centraal kenmerk van een programma is, dan heeft dat ook gevolgen voor de selectie. Vaardigheden rond communicatie en samenwerking zijn dan bijvoorbeeld belangrijk en zouden een rol in de selectieprocedure moeten spelen. In andere situaties is zelfstandig kunnen werken misschien juist wel van groter belang.

### **Financiële en organisatiecontext**

Een andere overweging zijn de kosten van goed selecteren. De tijd die en het geld dat een opleiding bereid is te investeren in de selectie vormen belangrijke randvoorwaarden: zijn er überhaupt werknemers die deze taak op zich kunnen nemen? Het ligt voor de hand hierbij zowel de kosten als de opbrengsten van een procedure mee te nemen. Als een selectieprocedure uitval en herkansingen voorkomt, kan dat financieel positief uitpakken voor een opleiding. In het onderzoek van Schreurs et al. (2018) komt bijvoorbeeld naar voren dat in dat geval ook een ‘dure’ selectieprocedure uiteindelijk financieel aantrekkelijker kan zijn dan een ‘goedkope’ loting. Daarbij speelt ook een rol hoe selectief een opleiding is: de verhouding tussen het aantal plaatsen en het aantal kandidaten. Wanneer een opleiding uiteindelijk vrijwel alle kandidaten toelaat ligt een grote investering in een selectieprocedure minder voor de hand.

### **Wat kun je vragen?**

Wat je van kandidaten kunt vragen heeft niet alleen een inhoudelijke component, maar gaat ook over de tijdsinvestering (en financiële investering) die je van kandidaten kunt vragen). De tijdsinvestering van kandidaten is een belangrijke afweging voor een selectieprocedure. Kandidaten zitten misschien in hun eindexamenjaar, of zijn bezig met het afronden van hun bacheloropleiding tijdens een selectieprocedure. Er moet een goede balans zijn tussen wat je van kandidaten vraagt (inhoudelijk, maar vooral ook in tijdsinvestering en op welke termijn) en wat dat in het geheel van het selectieproces oplevert. Zo weten we uit onderzoek dat het vragen om een inspanning tijdens de selectie, de selectie voorspellend kan maken (het zogenaamde participatie-effect; Schripsema et al., 2017), maar dat is ook wat een selectie stressvol kan maken voor de kandidaten. Dit kan tegen elkaar opwegen, maar denk er goed over na: is deze investering van de kandidaten de toename in de voorspellende waarde echt waard?

Besteed naast de tijdsinvestering ook aandacht aan wat je aan financiële bijdrage van kandidaten vraagt. Een directe financiële bijdrage voor kandidaten met een Nederlandse vooropleiding kan doorgaans niet, maar reiskosten kunnen in de praktijk wel kosten voor kandidaten betekenen. In afwegingen over zowel de tijdsinvestering als de financiële investering ligt het voor de

hand om specifiek na te gaan of dit voor bepaalde groepen kandidaten extra barrières opwerpt en wat dat betekent voor de diversiteit van de samenstelling van de studentenpopulatie.

Zoals eerder beschreven speelt bij wat je kunt vragen van kandidaten ook mee wat je op het moment van selecteren qua niveau van de kandidaten mag verwachten. Een goed beeld van wat kandidaten al (moeten) kunnen op het moment dat zij meedoen aan de selectie is daarvoor belangrijk. Daarbij is zicht op wat er in vooropleidingen gebeurt belangrijk (hoewel dit vaak ingewikkeld is, zeker voor opleidingen met een internationale instroom). Die voorkennis en de vaardigheden die studenten al ontwikkeld kunnen hebben zijn belangrijk voor het niveau waarop je de instrumenten ontwikkelt, maar ook voor wat je ermee kunt meten. Als je bijvoorbeeld wilt zien hoe kandidaten nieuwe informatie tot zich kunnen nemen, is het logisch om dat niet te doen met informatie die in hun vooropleiding al is behandeld.

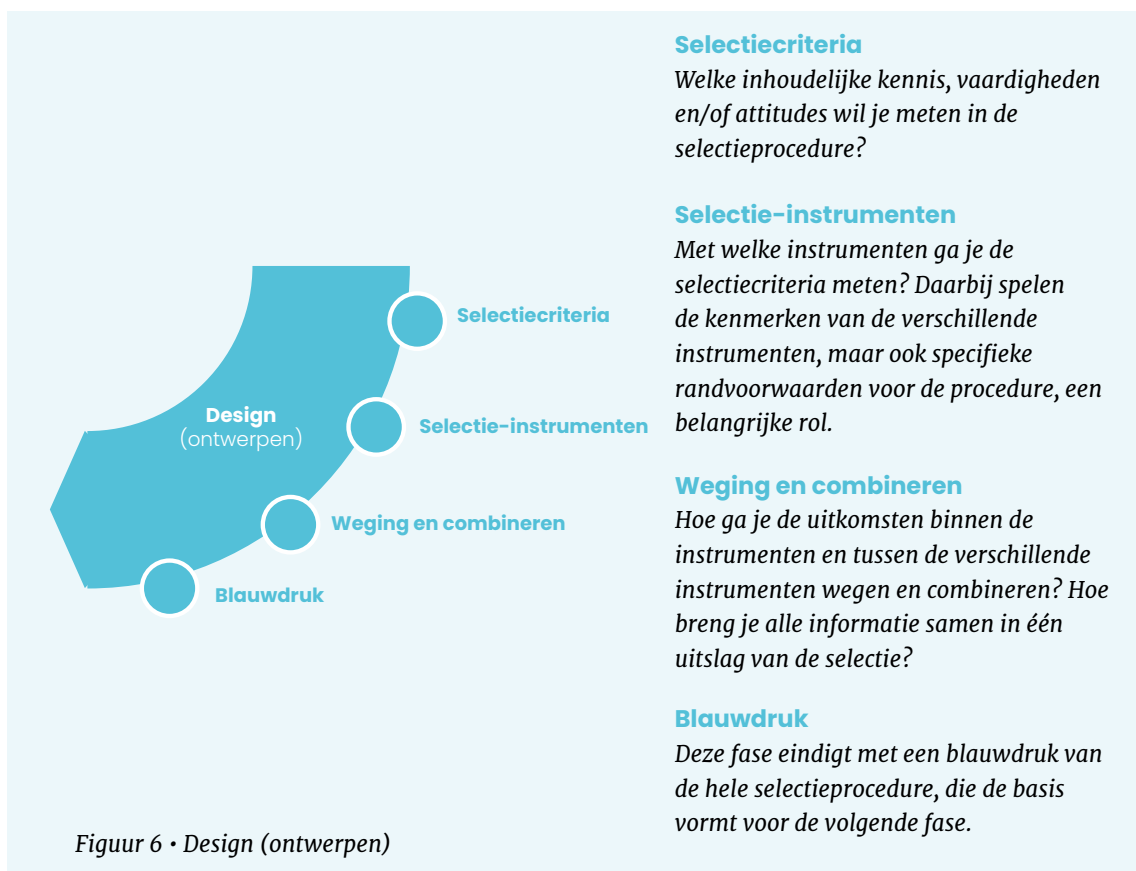
Daarnaast is een relevante, maar vaak over het hoofd geziene, vraag: wat kan selectie meegeven aan de kandidaat (zodat het meedoen aan de selectie voor de deelnemers in zichzelf ook nuttig is)? Een selectieprocedure is voor een kandidaat ook een mogelijkheid om een beter beeld te krijgen van wat ze van de studie kunnen verwachten: welke inhoud, op wat voor manier moeten ze het leren aanpakken (traditioneel, team-based learning, project-based, etc.), met wat voor studenten en docenten krijgen ze te maken en wat zijn de toekomstperspectieven? In het ontwerpen van een selectieprocedure zou deze kant ook nadrukkelijk een rol kunnen spelen. Daarnaast zou de uitkomst van de procedure de kandidaat ook informatie kunnen geven over eigen ontwikkelvragen, zoals in Nijmegen is onderzocht (de Visser, Laan, Engbers, Cohen-Schotanus & Fluit, 2018). De Nijmeegse onderzoekers concludeerden: “Selectie voor geneeskunde kan worden beschouwd als een leerervaring, en zowel studenten als geneeskundeopleidingen kunnen hiervan profiteren”.

Het Comeniusproject van Karen Stegers-Jager sluit hierbij aan. Zij doet onderzoek naar het integreren van een selectietoets die kandidaten een duidelijk beeld geeft van het beroep waarvoor ze meedoen aan de selectie (in dit geval geneeskunde). Hierbij omschrijft ze: “Onze uitdaging als instelling is om studenten aan te trekken en te selecteren die qua interesses en capaciteiten passen bij dit nieuwe curriculum. Want hoewel geneeskundestudenten vaak al vanaf hun vierde weten dat ze dokter willen worden, voldoet de opleiding niet altijd aan de verwachtingen. In dit project ga ik met een team van experts, waaronder studenten, een situationele beoordelingstest ontwikkelen, waarin aspirant-studenten kennis maken met realistische situaties uit de opleiding en het toekomstig werk als arts.” Deze test, bestaande uit videofragmenten, geeft aspirant-studenten een realistisch beeld van het leven van een geneeskundestudent en een arts en helpt hen zo beter geïnformeerd te beslissen om zich al dan niet aan te melden voor de opleiding in Rotterdam.



## 4 Design (ontwerpen)

In dit hoofdstuk gaan we in op het daadwerkelijk ontwerpen van je selectieprocedure. Hoe bepaal je welke kennis, vaardigheden en/of attitudes van de kandidaten je in de selectie wilt beoordelen? Met andere woorden, wat zijn de criteria van je selectie? En op welke manier kun je die criteria het beste meten (welke instrumenten gebruik je)? Het doel van de ontwerpfase is dat je komt tot een uitgewerkt plan, een blauwdruk van de selectieprocedure.



Figuur 6 • Design (ontwerpen)

## Selectiecriteria: Wat meet je?

Zoals in [hoofdstuk 2](#) beschreven, vormen de leerdoelen, de inhoud en vorm van je onderwijs en de toetsing het uitgangspunt voor de selectieprocedure. In de analysefase is het formuleren welke inhoudelijke kennis, vaardigheden en/of attitudes studenten nodig hebben om de opleiding goed te kunnen doorlopen een belangrijk onderdeel. Deze kennis, vaardigheden en/of attitudes die studenten nodig hebben in de opleiding zijn (onderdeel van) je selectiecriteria. De analyse- en designfase lopen zo in elkaar over.

In deze fase is het belangrijk om stil te staan bij het verschil tussen selectiecriteria en toelaatbaarheidseisen. Voor een masteropleiding waarin wordt voortgebouwd op statistische kennis, is een logische vereiste dat studenten bepaalde basiskennis hebben van statistiek. In dit handboek beschouwen we dit voorbeeld, in lijn met de UNL-afspraken, als een eis voor toelaatbaarheid en niet als selectie criterium. Afwegingen die passen bij selectiecriteria zijn bijvoorbeeld: Als er van studenten verwacht wordt dat ze analytisch kunnen denken, dan moeten we de vaardigheden die hieraan gekoppeld zijn meenemen in de selectieprocedure. Of: in de opleiding zitten veel schrijfopdrachten, dus studenten moeten over voldoende schrijfvaardigheid beschikken. Bij opleidingen met een internationale populatie zou je bijvoorbeeld ook de attitudes van studenten ten opzichte van internationale perspectieven mee kunnen nemen. Uiteraard zal er bij veel opleidingen sprake zijn van een combinatie van kennis, vaardigheden en attitudes.

Zoals we al in [hoofdstuk 2](#) bespraken is het belangrijk om stil te staan bij de mate van complexiteit waarop je vaardigheden in de selectieprocedure toetst. Dat is als vanzelfsprekend een andere dan tijdens de opleiding zelf. Bij deze stap hoort dus ook het terugredeneren van het eindniveau van de opleiding naar een niveau dat haalbaar en betekenisvol is voor kandidaten. Vragen die daarbij behulpzaam kunnen zijn, zijn bijvoorbeeld:

- Wat moeten de studenten kunnen binnen de opleiding?
- In hoeverre wordt er tijdens de opleiding tijd besteed aan het aanleren van deze aspecten? Hoe meer tijd je er tijdens de opleiding in steekt, hoe meer studenten zullen leren, hoe lager het niveau mogelijk kan zijn aan het begin. Dit is uiteraard afhankelijk van wat het eindniveau moet zijn. Aan de andere kant horen vaardigheden die in de opleiding belangrijk zijn mogelijk juist ook in de selectieprocedure een rol te spelen.
- Welk eindniveau wil je dat studenten bereiken aan het eind van de opleiding en helpt het daarbij als studenten een hoger startniveau hebben?
- Hoeveel kun je eigenlijk verwachten van de kandidaten, overschat je het niveau niet?

In de designfase neem je alle informatie uit de analysefase mee en kijk je er praktischer naar. Je hebt een overzicht van de kennis, vaardigheden en



attitudes die belangrijk zijn in de opleiding, en bepaalt wat daarvan echt in een selectieprocedure thuishoort en wat praktisch haalbaar is. Daarbij houdt je ook rekening met de randvoorwaarden die in de analysefase naar voren zijn gekomen. Deze stap mondt uit in een blauwdruk voor de selectie, die je vervolgens op instrumentniveau verder vormgeeft in de ontwikkelfase.

### Randvoorwaarden: waar houd je rekening mee?

Een eerste stap is om in kaart te brengen welke mogelijkheden je hebt. De hoeveelheid tijd en geld bepalen in grote mate welke opties er zijn. Denk bijvoorbeeld aan het aantal toetsen dat ontwikkeld (of gekocht) kan worden, de beschikbaarheid van de locatie voor het toetsen van kandidaten, de tijd die beschikbaar is voor het ontwikkelen en uitvoeren van de procedure, et cetera.

De doelgroep voor de selectie kan ook deels bepalen hoe de selectie eruit komt te zien. Voor een internationale opleiding kan het bijvoorbeeld onmogelijk zijn om de kandidaten op locatie op gesprek te laten komen of een toets te laten maken, waardoor er nagedacht moet worden over instrumenten die ook online afgenomen kunnen worden. Aan de andere kant kan een van de doelen zijn om kandidaten kennis te laten maken met de campus, waardoor een deel van je procedure juist op locatie moet zijn.

Verder hebben de principes en waarden die je meeneemt in je selectie ook invloed op de aanpak. Als het behouden van diversiteit bijvoorbeeld een belangrijk doel is van je selectie, zal hier binnen de instrumenten en in de procedure aandacht aan besteed moeten worden. Het is bijvoorbeeld belangrijk om na te gaan in hoeverre instrumenten gevoelig zijn voor betaalde coaching. Zowel voor gelijke kansen als voor de validiteit van de selectieprocedure is het een goed idee om een procedure te ontwikkelen waarbij dergelijke voorbereiding zo min mogelijk voordeel oplevert. Zo voorkom je dat je eigenlijk financiële armslag (of bereidheid om te betalen) meet, in plaats van kennis of vaardigheden van kandidaten.

Bij het vaststellen van de inhoud en het niveau van de selectiecriteria, en later ook bij het maken van de toetsing/opdracht, is het essentieel dat er mensen betrokken zijn die vanuit verschillende perspectieven meedenken. Dat kan geborgd worden door het oprichten van een selectiecommissie waarin verschillende expertises vertegenwoordigd zijn, of door het betrekken van collega's die regelmatig even gepolst kunnen worden voor specifieke onderwerpen. Denk hierbij ook aan de diversiteit in de groep collega's die je betreft (bijvoorbeeld in leeftijd, expertise, achtergrond, et cetera). Expertisegebieden die daarbij waardevol zijn, kunnen zijn:

- *Inhoudelijke expertise*: het liefst vanuit de breedte van de inhoudelijke onderwerpen die terugkomen in de opleiding. Denk hierbij na over welke

gebieden goed vertegenwoordigd moeten worden, bijvoorbeeld op basis van afstudeerrichtingen van studenten of juist richtingen waarin meer afgestudeerden nodig zijn.

- *Curriculum-expertise*: dit kunnen docenten en/of coördinatoren zijn die een grondige kennis hebben van de opbouw van het curriculum, zo veel mogelijk inhoudelijk de vakken kennen, bekend zijn met eventuele leerlijnen et cetera.
- *Onderwijskundige expertise*: om mee te denken over de validiteit en betrouwbaarheid van de toetsen, bijvoorbeeld of de vragen in een toets niet te veel inhoudelijke of contextuele expertise vragen die de kandidaten nog niet opgedaan kunnen hebben.
- *Ervaringsexpertise*: het studentperspectief is waardevol om inzichten vanuit de studie, studenten en het curriculum in te brengen (zij kunnen vaak goed aangeven wat voor huidige studenten lastig is, wat er in het curriculum gebeurt dat niet aansluit op de selectie, etc.). Ook kunnen zij zich inleven in de kandidaten.
- *Juridische expertise*: de regelgeving rond selectie biedt veel mogelijkheden, maar er zijn ook beperkingen (zie ook het vorige hoofdstuk). Daarnaast leveren selectiebesluiten regelmatig bezwaarprocedures op. Het kan daarvoor waardevol zijn om juridische expertise in te winnen om je ervan te verzekeren dat alles wat je doet binnen de kaders van de regelgeving blijft.
- *Evaluatie-expertise*: deze kan de selectieprocedure evalueren en kan daarmee ook verbeteringspunten inbrengen op basis van de evaluatie.
- *Organisatie-expertise*: is als laatste belangrijk om al deze expertise samen te kunnen brengen, en de ontwikkeling en implementatie binnen de opleiding goed vorm te geven.

## Selectie-Instrumenten: hoe toets je?

Na het in kaart brengen van de selectiecriteria en de randvoorwaarden, is de volgende stap om uit te zoeken welke selectie-instrumenten het meest geschikt zijn om de criteria te kunnen meten. Er zijn veel mogelijke instrumenten beschikbaar voor selectie. In deel 2 van dit handboek worden de eigenschappen van veelgebruikte selectie-instrumenten uitgebreid uitgewerkt, inclusief bij welke criteria de instrumenten aansluiten. Het ligt voor de hand om zeker voor belangrijke criteria een combinatie van instrumenten te gebruiken.

Een goede afstemming tussen instrumenten en criteria is dus van belang. Voor een eerste overzicht van instrumenten die mogelijk passen bij de geformuleerde criteria, zie het schema in [hoofdstuk 7](#). In dit schema wordt ook naar de hoofdstukken verwezen waarin de instrumenten besproken worden.

Bij het oriënteren op geschikte instrumenten kan het ook goed zijn te overwegen dat het mogelijk is om in meerdere verschillende ‘rondes’ te selecteren, bijvoorbeeld eerst op vooropleidingscijfers en een referentie en pas in de tweede ronde een ‘Multiple Mini Interview’. Zo kun je in de tweede ronde een instrument gebruiken dat (logistiek) niet haalbaar is met alle kandidaten. Afwegingen om gebruik te maken van meerdere rondes kunnen bijvoorbeeld het aantal aanmeldingen zijn, hoeveel kandidaten er op locatie ontvangen kunnen worden, hoeveel opleidingsplaatsen er zijn, hoeveel tijd er beschikbaar is, et cetera. Let daarbij wel op de voorwaarde dat kandidaten op minimaal twee criteria moeten worden beoordeeld. Een ander belangrijk aandachtspunt is dat het effect van een mogelijke bias in een instrument (bepaalde groepen kandidaten die onbedoeld benadeeld worden) groter is naarmate er minder instrumenten tegelijk gecombineerd worden. Bovendien ontnem je kandidaten de kans om een criterium waar ze minder op scoren te compenseren met een ander criterium. Hierdoor bestaat de kans dat je na de eerste ronde een minder diverse groep kandidaten overhoudt dan als je alle instrumenten tegelijk combineert. Bij het selecteren in verschillende rondes is het dus extra belangrijk om kritisch te kijken naar welke criteria je in de eerste ronde meeneemt en welke instrumenten je daarvoor gebruikt.

## Weging en combineren

Na het bepalen van de criteria (inhoud en niveau), en met welke instrumenten (en eventueel in hoeveel rondes) die criteria gemeten gaan worden, is de volgende stap om na te denken over de weging van de verschillende instrumenten. Wat daarbij belangrijk is, is de vraag of alle criteria even belangrijk zijn en of alle instrumenten de criteria even goed kunnen meten. Daarnaast kunnen er andere redenen zijn om een bepaald instrument meer te laten meewegen (bijvoorbeeld omdat het instrument veel lijkt op wat er in de opleiding aan bod komt). Ook hierbij liggen overwegingen rond *alignment* voor de hand als onderbouwing van de keuze.

Er zijn eindeloos veel opties in het combineren van criteria en instrumenten, het gebruik van combinaties van instrumenten en de weging binnen en tussen instrumenten. Zo hebben Lievens, Sackett en De Corte (2021) onderzoek gedaan naar een wegingsmethodiek (een soort AI) om voorspellende waarde te balanceren met diversiteit. Zij vonden dat iedere uitkomst en prioriteit die je stelt een andere weging met zich meebrengt. De ‘beste’ optie voor een specifieke opleiding komt altijd voort uit een combinatie van wat we uit onderzoek weten over bepaalde instrumenten (zoals dat ook in het tweede deel van het handboek wordt afgewogen) en de afweging van de opleiding over wat belangrijk is. Meijer, Neumann, Hemker en Niessen (2019) schetsen in grote lijnen twee mogelijkheden voor het bepalen van de weging: een weging op basis van de input van experts over welke informatie belangrijk

en betrouwbaar is, of een gelijke weging van alle scores. Ze concluderen dat het verschil in uitkomst tussen die twee opties vaak niet erg groot is. Een belangrijke afweging die ze meegeven is dat het (zwaar) meewegen van onbetrouwbare instrumenten (zoals een open, ongestructureerd interview) het meeste kwaad kan doen in een weging.

De vraag die hier bij hoort is of de precieze weging van ieder instrument wel vastgesteld moet worden. Een alternatief zou kunnen zijn dat een beoordeelaar alle informatie weegt en tot een ‘holistisch’ selectiebesluit komt. Dit is een vorm die vooral bij ‘harde’ selectie zou passen, en minder bij het maken van een ranking (zoals dat voor numerus-fixusopleidingen nodig is). Deze holistische selectie op basis van meerdere criteria is de tegenhanger van ‘mechanisch selecteren’, waarbij rekenregels zijn vastgesteld hoe scores gewogen moeten worden. Holistisch selecteren voelt voor beoordelaars vaak rechtvaardiger, maar het gevaar dat het bias introduceert is groot (bijvoorbeeld dat kandidaten waarin de beoordelaar zichzelf herkent beter scoren). Meijer et al. (2019) concluderen dat de keuze voor (in elk geval volledig) holistisch selecteren niet verdedigbaar is.

## Blauwdruk

De uitkomsten van de eerste twee stappen (analyse en design) vormen samen een blauwdruk voor je selectieprocedure. Hierin staat welke criteria er in de selectie beoordeeld worden, welke instrumenten er ingezet worden en in welke verhoudingen de verschillende onderdelen meegewogen worden. Het is eigenlijk een soort uitgebreide toetsmatrijs. In de blauwdruk wordt ook verwacht dat de verschillende criteria zijn uitgewerkt, waarbij wordt aangetoond dat (1) de verschillende instrumenten inderdaad passen bij de criteria, (2) de instrumenten het hele spectrum van het criterium afdekken en (3) geen criteria onbevestigd blijven of juist overvraagd worden. Deze blauwdruk kan ook het resultaat zijn van een iteratief proces, waarbij criteria soms kunnen worden aangepast aan wat praktisch haalbaar is. De blauwdruk vormt vervolgens de basis voor je selectieprocedure.

Aan de hand van de blauwdruk kunnen de instrumenten verder uitgewerkt worden. Dat kan bestaan uit het ontwikkelen van vragen voor een toets, of de vragen die je referenten (wanneer er gebruik gemaakt wordt van aanbevelingsbrieven) wil voorleggen. Dit verder uitwerken van instrumenten is de start van de develop-/ontwikkelfase. Het blijft belangrijk om te toetsen of deze uitwerking past bij de blauwdruk, inclusief het belang van de verschillende criteria die je daarin hebt bepaald.

Het doel is dat er uiteindelijk voldoende onderscheid gemaakt kan worden tussen de verschillende kandidaten, zodat de procedure voldoende infor-



matie oplevert voor het maken van een ranking of een selectiebesluit. Het is daarom van belang om ervoor te zorgen dat verschillende (onderdelen van de) instrumenten voldoende verschillende informatie opleveren. Bij het ontwerp van een instrument kan daar op verschillende manieren op worden gelet. Bijvoorbeeld door een manier van vragen stellen of scores die een groter verschil tussen kandidaten oplevert, een plan van wat je doet in het geval dat kandidaten op dezelfde score (en dus rang) terechtkomen, of een uitwerking van wat je doet als kandidaten op de grens tussen toelaten en afwijzen uitkomen.

## Vorbereiding kandidaten

Een laatste stap in de ontwerpfase, die ook al eerder in het handboek kort aan de orde kwam, is de voorbereiding die je van kandidaten wilt vragen. Voorbereiding kan betrekking hebben op toetsen die je afneemt, waarbij het doel is om te kijken of ze de stof voldoende geleerd en/of begrepen hebben, maar kan ook in andere instrumenten zitten. De afweging hoeveel tijd je van kandidaten vraagt ten opzichte van de informatie die het je oplevert is belangrijk in het ontwerpproces. De aansluiting tussen de rol die voorbereiding speelt en de criteria die je hanteert kan daarbij ook meewegen. Het kan zijn dat je planvaardigheden wilt toetsen en daarom elementen daarvan inbrengt in de procedure. Weeg wel ook daar steeds af wat dat voor (verschillende groepen) kandidaten betekent.

Aan de andere kant is het belangrijk om af te wegen of hulp bij voorbereiding bepaalde kandidaten een voordeel kan opleveren. Een procedure die sterk steunt op een kennistoets die studenten moeten voorbereiden is bijvoorbeeld meer vatbaar voor commerciële ondersteuning. Ook kan hulp vanuit de omgeving voor sommige voorbereiding een factor zijn die in de afweging speelt, bijvoorbeeld als het gaat om een uitwerking van de motivatie. Een alternatieve vorm is om vanuit de opleiding ondersteuning aan te bieden aan alle kandidaten om op die manier ongelijke mogelijkheden voor het organiseren van ondersteuning te ondervangen. Dat kan bijvoorbeeld in de vorm van voorbereidingsdagen, zoals dat voor de opleiding geneeskunde aan de EUR gebeurt.

Als je voorbereidingsdagen wilt aanbieden, denk dan goed na over wat je tijdens die dagen wilt behandelen. Zorg daarbij voor een zorgvuldige en tijdige planning die aansluit bij de doelgroep. De ervaring leert dat kandidaten het vaak vooral waarderen als ze een korte uitleg en zo veel mogelijk daadwerkelijke oefening met de toetsvorm en antwoordformats krijgen. Als je toets een bijzondere vorm heeft die kandidaten nog niet kennen, is een goede voorbereiding extra belangrijk en kan een voorbereidingsdag waardevol zijn.



## 5 Develop (ontwikkelen)

Als het voorbereidende onderzoeks- en denkwerk is afgerond, begint de ontwikkelfase van de selectieprocedure. In deze fase worden de selectie-instrumenten en alle bijbehorende informatie die in de ontwerpfase is uitgedacht ontwikkeld. Denk hierbij aan materialen zoals toetsen, beoordelingsrubrics of leermaterialen voor kandidaten. Een centrale afweging daarbij is of de instrumenten, zoals toetsen, door de opleiding zelf ontwikkeld gaan worden, of dat gebruik wordt gemaakt van elders ontwikkelde instrumenten. De focus in dit hoofdstuk ligt op de stappen voor het zelf ontwikkelen van een instrument; een marktverkenning van extern aangeboden toetsen is geen onderdeel van dit handboek. We gaan wel kort in op het proces van het kiezen van een extern ontwikkelde toets. Aan die optie zitten enkele belangrijke nadelen. Zo is de aansluiting op de inhoud en vorm van het eigen curriculum natuurlijk meestal niet optimaal.



## Zelf ontwerpen?

Voor een aantal onderwerpen (kennis) en vaardigheden bestaan er toetsen die je elders kunt aanschaffen. Een voorbeeld daarvan is de BMAT, de Bio-Medical Admission Test (die overigens binnenkort ophoudt te bestaan), of CASPer, een Situational Judgement Test (SJT) die door een Canadees bedrijf wordt aangeboden. De voordelen van het gebruik van een bestaand instrument liggen voor de hand: veel minder tijdsinvestering voor het ontwikkelen en vaak ook voor het beoordelen, de evaluatie en kwaliteitszorg liggen ergens anders, et cetera. Aan de andere kant zijn er ook belangrijke nadelen: veel elementen zijn context- en cultuurgebonden, en vertalen zich niet goed naar de context van het Nederlandse hoger onderwijs, laat staan de context van een specifieke opleiding. Die context speelt in elk geval vaak een belangrijke rol als het bijvoorbeeld gaat om (inter)persoonlijke vaardigheden (Stegers-Jager, 2018). Ook zijn veel toetsen in het Engels en moet er voor Nederlandstalige opleidingen eventueel vertaald worden. Bovendien kunnen de kosten voor het afnemen van een test een nadeel zijn. Het belangrijkste nadeel is echter dat er geen mogelijkheden zijn voor *alignment* met de specifieke criteria die in de eerdere fasen zijn uitgewerkt en dat kandidaten geen kennis kunnen maken met wat ze in jouw opleiding kunnen verwachten.

Aan de andere kant brengt het zelf ontwikkelen van instrumenten natuurlijk ook de nodige uitdagingen met zich mee. Het is niet altijd makkelijk om een goede toets voor de gewenste vaardigheden te ontwikkelen. In het tweede deel van het handboek gaan we uitgebreider in op de moeilijkheden die horen bij het ontwikkelen van verschillende instrumenten. In dit hoofdstuk doorlopen we een aantal algemene punten waarvan het raadzaam is dat ze meegenomen worden tijdens de fase waarin selectiemateriaal ontwikkeld wordt.

## Blauwdruk, vragen & items

Het ontwikkelen van de daadwerkelijke instrumenten vergt veel tijd en aandacht. De blauwdruk blijft een centrale rol spelen: die bepaalt de criteria en inhoud, en bij toetsen bijvoorbeeld ook de vorm van de toets en antwoordformats. Sommige stappen van de ontwikkelfase doorloop je jaarlijks, bijvoorbeeld toetsen die elk jaar opnieuw ontwikkeld moeten worden. Bij andere instrumenten stel je vooral bij naar aanleiding van de evaluatie, maar hoeft je niet jaarlijks helemaal opnieuw te ontwikkelen.

Voor het ontwikkelen van een goede, valide en betrouwbare toets, verwijzen we graag naar het standaardwerk van Van Berkel, Bax en Joosten-Ten Brinke (2017). Voor toetsen in een selectieprocedure is dat misschien nog belangrijker dan voor toetsen binnen de opleiding. Er hangt voor studenten (en de

opleiding) veel af van de uitkomst van de toets. En waar studenten binnen een opleiding doorgaans vaak getoetst worden, is het aantal toetsmomenten in een selectieprocedure beperkt.

### **Inhoud van het instrument**

De ontwikkeling van de inhoud van het instrument hangt natuurlijk sterk samen met welke vaardigheid centraal staat. Vaak wordt daarbij de ‘tegenstelling’ tussen cognitieve en niet-cognitieve of affectieve vaardigheden benadrukt (Schreurs, 2020). De scheiding tussen die twee domeinen is echter nooit zo scherp. Cognitie speelt vrijwel altijd een rol, ook als het gaat om vaardigheden in de andere domeinen (Van der Linden, 2018). Kandidaten moeten immers nadenken over wat er van hen gevraagd wordt. In veel praktijksituaties speelt een combinatie van cognitieve en affectieve vaardigheden ook een rol. Voor de opleiding geneeskunde is dat bijvoorbeeld het geval als studenten in een klinische setting kennis moeten gebruiken in interactie met patiënten. Zeker naarmate de opleiding vordert, wordt die combinatie belangrijk (Patterson et al., 2018; Stegers-Jager, 2018). Het is dan ook niet erg, en mogelijk zelfs krachtig, om dit soort combinaties – versimpeld – in de selectieprocedure een rol te laten spelen.

In deel twee van dit handboek is te vinden welke instrumenten passen bij welk domein (cognitief, affectief of psychomotorisch, voor meer uitleg over deze domeinen, zie [hoofdstuk 2](#)). In het tweede deel van het handboek gaan we ook dieper in op aandachtspunten voor instrumenten. Het is belangrijk om bij de inhoud van het instrument ook weer stil te staan bij het verschil tussen het niveau waarop de vaardigheid in de opleiding wordt gevraagd en het niveau waarop kandidaten al voor aanvang van de opleiding over die vaardigheid kunnen beschikken (zie ook daarvoor [hoofdstuk 2](#), figuur 4).

Een belangrijke vraag is de validiteit van je selectie: meet je wat je wilt meten? Zodra je instrument (ook) andere dingen meet, is dat een gevaar voor de kwaliteit van je selectieprocedure en kan het bias introduceren. Als je instrument bijvoorbeeld naast kennis ook meet of iemand tijdens de voorbereiding de mogelijkheid had om gebruik te maken van betaalde coaching, introduceer je bias én wordt de procedure minder valide. Het al dan niet kunnen betalen van hulp is immers geen onderdeel van je criteria, daarom wordt in de betreffende hoofdstukken in deel twee bij de instrumenten steeds benoemd of er iets bekend is over de gevoeligheid voor coaching. Afhankelijk van het instrument zijn er nog verschillende andere aandachtspunten, bijvoorbeeld hoeveel invloed lees- en taalvaardigheid heeft bij instrumenten die eigenlijk andere vaardigheden beogen te meten. In het tweede deel van het boek gaan we uitgebreider in op belangrijke overwegingen bij elk van de instrumenten.

### Antwoordformat van de vragen

Bij het ontwikkelen van een selectie-instrument is het belangrijk na te gaan welke vorm van informatie je van de kandidaten wilt ontvangen. Daarbij staat wat je wilt kunnen beoordelen centraal. Vaardigheden op het gebied van samenwerken laten zich waarschijnlijk niet goed vangen in een meerkeuzevraag. Bij dat soort vaardigheden zouden open vragen of het observeren van daadwerkelijk gedrag beter passen.

Ook bij het bepalen van het antwoordformat is de validiteit belangrijk om bias te voorkomen en de juiste vaardigheden te kunnen voorspellen. Daarnaast is het goed om na te gaan wat voor reactie je instrument oproept. Bij een toets of interview gaat dat bijvoorbeeld over het controleren of de vraagstelling eenduidig is. Over de validiteit en betrouwbaarheid van toetsing is veel bekend. In [bijlage 1](#) vatten we een aantal belangrijke inzichten samen.

Om bijvoorbeeld motivatie te testen, zou je kunnen vragen:  
Waarom ben je gemotiveerd voor de studie Toegepaste Psychologie?

- a) Ik wil mensen helpen
- b) Ik wil snappen hoe mensen werken
- c) Ik wil mezelf beter leren kennen
- d) Ik wil mensen kunnen manipuleren
- e) Ik kon geen andere studie vinden
- f) Ik vind deze selectie het makkelijkst

Door de antwoorden hier te geven, laat je kandidaten niet nadenken en geef je ze erg weinig ruimte. Je zou dit al kunnen verbeteren door ze meerdere opties te laten kiezen uit bovenstaande, of door voor iedere reden te vragen of die wel of niet op hen van toepassing is. Waarschijnlijk is het beter is om via een open vraag concreet te informeren naar de reden om voor de opleiding te kiezen, bijvoorbeeld:

Geef een specifiek voorbeeld uit de media of van iets dat je in je omgeving hebt meegemaakt dat je keuze voor de opleiding Toegepaste Psychologie heeft bevestigd. - Leg de link tussen dat voorbeeld en wat je later met de opleiding wilt kunnen. Geef je antwoord in maximaal 150 woorden.

Met je antwoordformat bepaal je ook wat je meet. De eerste optie laat vooral zien hoe goed kandidaten wenselijk antwoord kunnen geven, terwijl een open vraag een veel grotere kans geeft om daadwerkelijke informatie over de kandidaat te verkrijgen. Dit werkt ook zo met andere lastig te meten kenmerken of vaardigheden, denk bijvoorbeeld aan ethisch bewustzijn, empathie, etc.

Minstens net zo belangrijk is het uitwerken van de manier van beoordelen. Het ontwikkelen van een beoordelingsvorm waarbij kandidaten consistent worden beoordeeld, waarbij het niet uitmaakt welke beoordelaar een kandidaat ‘treft’, vraagt nadrukkelijke aandacht en evaluatie. Als er veel ruimte is voor de eigen interpretatie van een beoordelaar, is de kans op verschillen en mogelijke bias groter. Aan de andere kant zijn sommige vaardigheden lastig te vatten in één goed antwoord. Het kan dan helpen om meerdere mensen bij de beoordeling van elke kandidaat te betrekken. Zelfs als niet elk instrument door meerdere mensen beoordeeld wordt, kan het betrekken van een diverse groep beoordelaars bij de verschillende instrumenten potentiële bias al verkleinen.

## Testen

Tijdens en na het ontwikkelen van instrumenten (bijvoorbeeld toetsen of beoordelingsrubrics voor antwoorden tijdens interviews) en leer materiaal, is het raadzaam collega's, studenten of andere betrokkenen te laten meekijken. Dit kan zowel tussendoor, waardoor ze kunnen meedenken en input kunnen leveren, als aan het einde wanneer er een concept ligt. Vragen die je bijvoorbeeld zou kunnen stellen zijn:

- Zorgt de ontwikkelde procedure ervoor dat de criteria die beoordeeld moeten worden tijdens de selectie goed worden uitgevraagd?
- Is de manier van beoordelen vergelijkbaar met hoe deze criteria in de opleiding worden beoordeeld?
- Zijn eventuele toetsvragen eenduidig en op het juiste niveau geformuleerd?
- Is de aangeboden leerstof voor kandidaten beheersbaar en representatief voor wat er tijdens de opleiding gevraagd wordt?
- Zouden groepen kandidaten met bepaalde achtergrondkenmerken met (elementen van) dit instrument meer moeite hebben dan andere kandidaten?
- Is te verwachten dat de resultaten van het instrument makkelijk (efficiënt) te beoordelen zijn?

Een *pilot* kan goed werken om de procedure en instrumenten te testen. Het doel van een pilot is om de volledige selectieprocedure een keer te testen zonder consequenties voor de opleiding of de kandidaten. Hiermee kan bijvoorbeeld bepaald worden of ontwikkelde toetsen op het juiste niveau zijn en of de beschikbare tijd voor onderdelen in de selectie voldoende is. De blauwdruk en instrumenten zijn dan al door collega's of een aantal studenten kritisch bekeken. Een manier om een pilot uit te voeren kan zijn door een nieuw instrument toe te voegen aan een selectieprocedure, maar het nog niet te laten meetellen, of door het mee te nemen als onderdeel van een matchingprocedure voordat selectie wordt geïntroduceerd. Een alternatief is om een groep huidige studenten te vragen om de procedure of het instrument

uit te proberen. In het volgende hoofdstuk gaan we verder in op hoe je de resultaten van een pilot kunt evalueren en verwerken in aanpassingen van de procedure.

## Dataveiligheid en privacy

Een laatste overweging bij het ontwikkelen van de selectie-instrumenten zijn overwegingen die te maken hebben met het opslaan van de gegevens, antwoorden en scores van kandidaten. Zowel privacy als een goede bescherming van die gegevens is belangrijk. De AVG vormt daarbij een kader, dat ook consequenties heeft voor waar data kunnen worden opgeslagen. Bij het gebruik van bepaalde software kan dat beperkingen opleveren. Zorg dat je dit tijdig uitzoekt, denk na over wie welke toegang tot informatie nodig heeft (of niet nodig heeft) en over hoe je de evaluatie van de selectie wilt inrichten. Bij dat laatste moet bekeken worden of er een *informed consent* van kandidaten nodig is. Het kan erg helpen om hiervoor deskundig advies binnen de instelling in te winnen, zeker als je in de selectie gebruik wilt maken van software die verder niet gebruikt wordt.



## 6 Implementeren & Evalueren

Als de procedure is uitgedacht en ontwikkeld volgt de **implementatiefase**. Voor een handboek is deze fase minder interessant, omdat als het goed is de implementatie niet meer is dan de uitvoering van de selectieprocedure. Natuurlijk lopen er in de praktijk altijd dingen net anders dan gepland, maar de uitgangspunten die eerder zijn uitgedacht en uitgewerkt zouden voldoende handvatten moeten bieden om daarmee om te gaan. Wel is het belangrijk dat alles wat er gebeurt gedurende de uitvoering van de procedure goed bijgehouden wordt. Deze informatie is namelijk van groot belang in de **evaluatiefase**.

Bij bijhouden gedurende de procedure valt te denken aan:

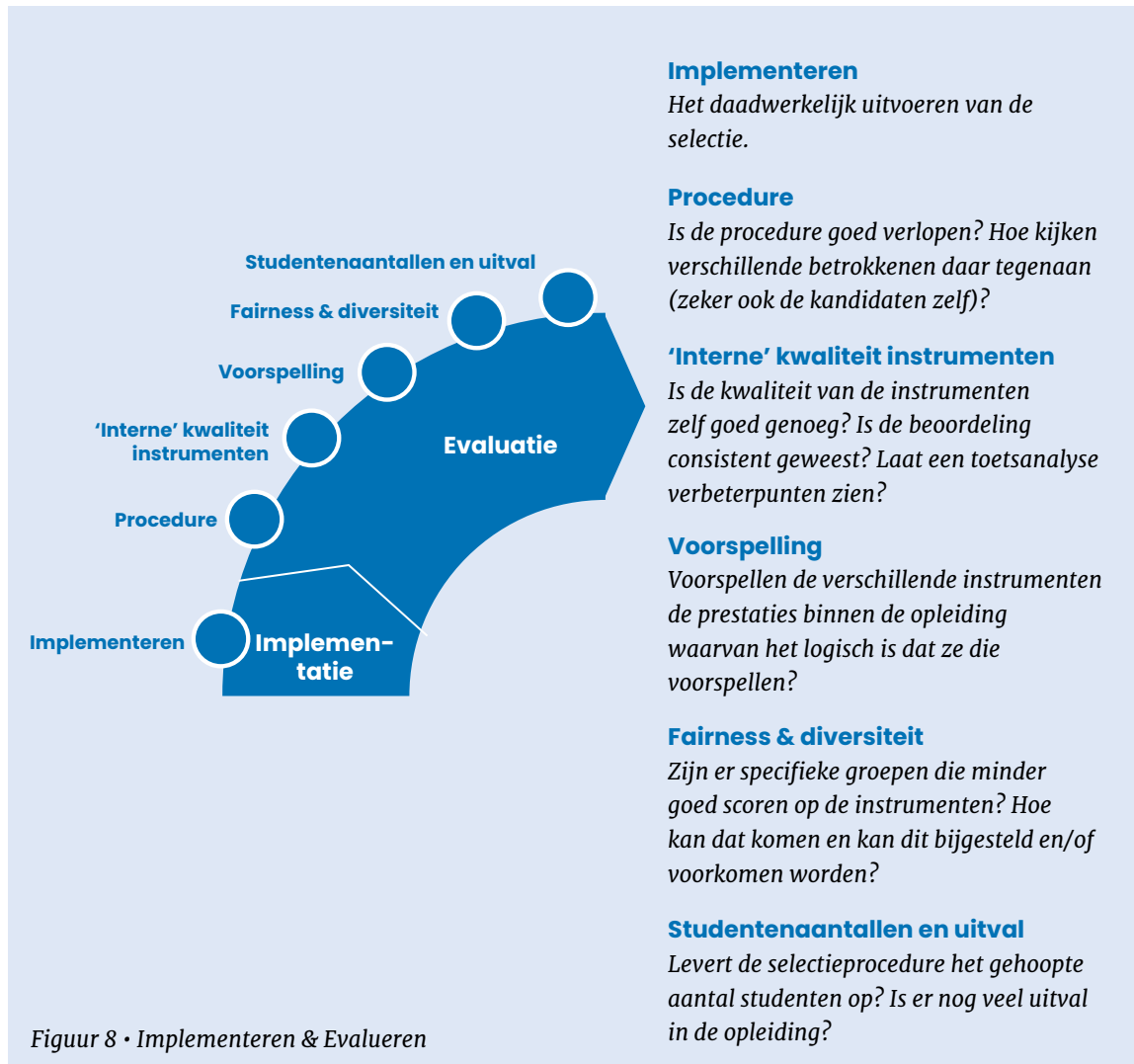
- Stellen kandidaten gedurende de procedure vragen, en zo ja, wat voor vragen? Lopen zij tegen problemen aan?
- Lopen collega's die (een deel van) de procedure uitvoeren tegen problemen aan?
- Zie je zelf ergens problemen ontstaan (bijvoorbeeld tijdens de aanmelding of vallen veel kandidaten af bij bepaalde stappen in de procedure, etc.)?
- Hoe werken regelingen voor specifieke groepen die bijvoorbeeld extra tijd krijgen (dyslexie, faalangst, beperkingen, studenten met specifieke achtergronden)? Blijken die groepen niet achtergesteld te worden?
- Hoeveel tijd kosten de verschillende stappen in de procedure voor zowel kandidaten als voor de opleiding?
- Hoeveel ondersteuning is er nodig om de procedure goed te laten verlopen?
- Eventuele andere observaties, problemen of merkwaardigheden die je opvallen.

Nadat de selectieprocedure is uitgevoerd, volgt de laatste stap: een uitgebreide **evaluatie** van het geheel.

Aan het einde van je evaluatie wil je vooral kunnen concluderen:

- Of je vooraf opgestelde doelen behaald zijn

- Of je instrumenten hebben gemeten wat ze hadden moeten meten
- Of de selectie ook positieve effecten heeft (zoals studieresultaten, kosteneffectiviteit of m.b.t. diversiteit) die opwegen tegen eventuele negatieve effecten
- Hoe de procedure en/of de instrumenten voor de volgende keer verder versterkt kunnen worden.



## Verschillende elementen in de evaluatie

Voor het evalueren van de selectie zijn er verschillende elementen te onderscheiden. Een uitgebreide evaluatie richt zich zowel op de procedure als geheel, als op de verschillende selectie-instrumenten die daar onderdeel van zijn. Voor de instrumenten kun je daarnaast kijken naar de 'interne' kwaliteit

van instrumenten (bijvoorbeeld of ze voor voldoende onderscheid zorgen tussen kandidaten, of dat de beoordeling consistent gebeurt), en in hoeverre ze de prestaties van de studenten binnen de opleiding kunnen voorspellen. Daarnaast kun je kijken welke toegevoegde waarde elk los instrument heeft op het totaal aan informatie dat je verzamelt over kandidaten. In dit hoofdstuk staan we kort stil bij deze verschillende elementen van het evalueren van de selectieprocedure, in bijlage 2 geven we hiervoor verdere handvatten. Een goede evaluatie vereist ook een zekere expertise, het kan waardevol zijn om die voor deze fase in te brengen als die niet binnen de opleiding zelf aanwezig is.

Vanuit de Expertgroep Toegankelijkheid Hoger Onderwijs wordt gewerkt aan een praktische toepassing die het mogelijk maakt om laagdrempelig zelf de evaluatie van een selectieprocedure uit te voeren. Die toepassing zal zich met name richten op wat de uitkomsten van de selectieprocedure zeggen over de prestaties van studenten in de opleiding, en over de toegevoegde waarde van verschillende instrumenten ten opzichte van elkaar. Deze mogelijkheden worden op [onderwijskennis.nl](http://onderwijskennis.nl) toegankelijk gemaakt zodra ze beschikbaar zijn.

## Procedure

Voor het evalueren van de procedure is de informatie die gedurende de procedure is verzameld van belang (zie de punten die aan het begin van dit hoofdstuk vermeld staan). Naast het bijhouden van informatie kan er ook gedacht worden aan het actief verzamelen van informatie, bijvoorbeeld door het afnemen van eventuele vragenlijsten onder kandidaten. Centraal staat de vraag hoe kandidaten en andere belanghebbenden de selectie hebben ervaren en of er in de procedure dingen verbeterd kunnen worden. Voor het afnemen van een vragenlijst onder kandidaten ligt het voor de hand om, zeker bij de eerste keer dat de procedure wordt uitgevoerd, vragen te stellen over onderdelen van de procedure waarover bij de ontwikkeling twijfels waren en om veel open ruimte te houden voor opmerkingen. Het advies is om de vragenlijst anoniem af te nemen voordat de uitslag van de selectie bekend is. Informatie over de gehele procedure kan ook verzameld worden onder andere betrokkenen.

Vragen die bij een evaluatie voor de hand liggen zijn:

- **Vorbereiding**
  - Was er voldoende tijd om de procedure voor te bereiden vanuit de opleiding?
  - Hadden kandidaten genoeg tijd om zich voor te bereiden? Hoeveel tijd heeft de voorbereiding gevraagd?
  - Eventueel: Was het studiemateriaal relevant en had het samenhang met de toets?

- **Selectie-instrumenten**
  - Was de inhoud duidelijk en relevant?
  - Hoe is de manier van afnemen ervaren?
  - Worden de instrumenten ervaren als legitiem?
- **Logistiek**
  - Waren er momenten dat kandidaten of anderen niet goed wisten wat er van hen gevraagd werd of wat ze moesten doen?
- **Persoonlijk**
  - In hoeverre hebben de kandidaten stress ervaren naar aanleiding van de procedure?
  - In hoeverre hebben kandidaten het gevoel dat ze van zichzelf hebben kunnen laten zien wat zij denken dat belangrijk is voor de studie om te weten?
  - In hoeverre voelen kandidaten zich gezien?
  - Hoe hebben kandidaten zich voorbereid op de procedure?
  - Mogelijke aanvullingen voor de selectie

### ‘Interne’ kwaliteit van de instrumenten

In de selectieprocedure worden instrumenten gebruikt om de vooraf gestelde criteria te meten. Een eerste stap in het evalueren van instrumenten is om de uitkomsten verder te analyseren. Zo kun je bekijken of instrumenten voldoende onderscheid maken tussen kandidaten. Als alle kandidaten dezelfde score behalen, dan voegt het instrument natuurlijk weinig toe. Daarnaast zijn verdere analyses ook mogelijk. Welke analyses nuttig zijn hangt af van het instrument. Bij toetsen kun je de moeilijkheid bepalen door te kijken naar hoeveel procent van de kandidaten de verschillende vragen goed heeft beantwoord (*p-waarde*) en onderzoeken in hoeverre een vraag goed onderscheid maakt tussen goede en minder goede kandidaten (*Rir en Rar*). Dit zijn onderdelen van een zogenaamde toets- en itemanalyse, die ook binnen veel opleidingen wordt gebruikt om de kwaliteit van toetsing te onderzoeken. Omdat toetsen een veelgebruikt instrument zijn in selectieprocedures, is uitgebreidere informatie over toetsanalyses opgenomen in [bijlage 1](#). Bij veel instellingen is daarnaast op dit vlak ook specialistische kennis beschikbaar.

Bij het beoordelen van bijvoorbeeld CV's en motivatiebrieven van of interviews met kandidaten is het aan te raden om te onderzoeken of twee beoordelaars consistent dezelfde producten beoordelen wanneer ze dat onafhankelijk van elkaar doen. Dat kun je berekenen door de zogenaamde interbeoordelaar-betrouwbaarheid uit te rekenen.

Deze stappen horen natuurlijk in een pilot thuis, maar ook na de selectie is het goed om dit soort evaluaties uit te voeren richting de volgende ronde. Bij grotere herzieningen in de procedure, of als er andere mensen betrokken

raken bij de procedure heeft dat natuurlijk in het bijzonder toegevoegde waarde.

## Voorspelling

Naast de interne kwaliteit is het van belang om te beoordelen of de instrumenten inderdaad de vooraf bedachte criteria meten. De manier om dat te onderzoeken is door te bepalen of kandidaten die hoger scoren op een bepaald instrument, het ook binnen de opleiding op dat criterium beter doen. Dus: kunnen de instrumenten toekomstige prestaties goed voorspellen? Er gaat echter wat tijd overheen voordat je dat goed kunt evalueren. Daarvoor moeten er immers gegevens beschikbaar zijn over de prestaties van kandidaten binnen de opleiding. Tijdens een pilot met huidige studenten kunnen soortgelijke analyses vaak wel al eerder uitgevoerd worden.

Voor het voorspellen van prestaties is het belangrijk om verschillende soorten studieresultaten binnen de opleiding uit elkaar te houden. Net als dat sommige instrumenten vooral bedoeld zijn voor het toetsen van lagere orde cognitieve vaardigheden, en andere voor het toepassen van kennis (zie [hoofdstuk 2](#)), is dat ook het geval voor de toetsen binnen de opleiding. Voor de evaluatie geldt dus wederom het belang van *alignment*. Als een selectie-instrument bedoeld is om de vaardigheid toepassen van kennis te voorspellen, dan moet het samenhangen met toetsresultaten binnen de opleiding die ook toepassen van kennis toetsen. Het is dus niet erg als zo'n instrument niet goed met andere toetsen in de opleiding samenhangt, deze meten immers andere vaardigheden. Omdat bepaalde vaardigheden misschien pas later in de opleiding belangrijk worden, ligt het voor de hand om niet alleen naar de resultaten in het eerste jaar te kijken. In sommige gevallen komen criteria uit de selectie voor de bachelor pas in de master terug, zoals (Schreurs, 2020) laat zien voor de opleiding Geneeskunde.

Het zal nooit zo zijn dat selectie-instrumenten alles kunnen voorspellen. Tijdens de opleiding gebeurt er van alles, zowel in motivatie en leren, als in het persoonlijk leven van studenten. Als een instrument niet samenhangt met de resultaten in de opleiding voor de vaardigheden die het beoogt te voorspellen, is de kans echter groot dat het ook bias introduceert. Dan is er een risico dat bepaalde groepen ten onrechte systematisch benadeeld worden.

Deze vorm van evaluatie is complexer voor opleidingen met een vorm van 'harde' selectie, waarbij bepaalde groepen kandidaten worden afgewezen omdat ze onvoldoende beschikken over de voor de opleiding benodigde competenties. In dat geval zijn er immers geen resultaatgegevens van afgewezen kandidaten. In sommige gevallen zal dat in een pilot wel te onderzoeken zijn. In andere gevallen kan het helpen om kandidaten waarover getwijfeld werd,

maar die toch zijn toegelaten, goed te registreren. Als die groep op de criteria waarover twijfel bestond inderdaad slechter presteert (of niet) dan de overige studenten, dan zegt dat iets over de voorspellende waarde van het instrument.

## Samenhang tussen instrumenten

Naast het evalueren van de voorspellende waarde van losse instrumenten, is het onderzoeken van de samenhang tussen verschillende instrumenten interessant. Op die manier kun je vaststellen of alle instrumenten nodig zijn en voldoende toevoegen aan het geheel, of dat de procedure ook met minder instrumenten de gewenste resultaten kan opleveren.

Allereerst kan een analyse van de samenhang tussen instrumenten aan het denken zetten over wat er daadwerkelijk gemeten wordt. Als instrumenten die eigenlijk andere vaardigheden zouden moeten meten sterk met elkaar samenhangen, dan kan het betekenen dat ze ook een andere, onderliggende vaardigheid meten. Denk bijvoorbeeld aan verschillende instrumenten waarin schrijfvaardigheid of nauwkeurigheid een belangrijke, onderliggende rol speelt. Soms is samenhang juist erg logisch. Dat kan het geval zijn bij instrumenten die vaardigheden in communicatie en in samenwerking meten. Het is immers niet mogelijk om samen te werken zonder communicatie. Samenhang is niet per se goed of slecht, maar geeft wel aanleiding om in het licht van de blauwdruk verder te denken over de rol van elk van de instrumenten.

Een sterke samenhang kan ook betekenen dat een van de instrumenten onvoldoende informatie toevoegt. Het kan daarbij bijvoorbeeld gaan om twee instrumenten die dezelfde vaardigheid beogen te meten, zoals vooropleidingscijfers en een kennistoets. Een regressiemodel kan de toegevoegde waarde van een extra instrument voor het voorspellen van dezelfde vaardigheid laten zien. De omvang van die toegevoegde waarde kan waardevolle informatie zijn voor de evaluatie.

## De uitkomsten

Waar de evaluatie tot nu toe vooral ging over het selectieproces, zijn ook de uitkomsten van de selectie van belang. Wanneer bepaalde instrumenten naast het meten van specifieke criteria, ook een bias hebben die zorgt dat groepen systematisch worden uitgesloten dan wel benadeeld, is het natuurlijk van bijzonder groot belang dat dit in de evaluatie snel naar boven komt (daarnaast zijn voorzorgsmaatregelen om dat te voorkomen natuurlijk nog belangrijker; daarop gaan we in het tweede deel van het handboek nader in).

Een aantal belangrijke afwegingen voor de evaluatie van de uitkomst behandelen we in dit laatste deel van dit hoofdstuk.

### **Fairness & Diversiteit**

In de ontwikkelfase hoort bij ieder instrument de vraag gesteld te worden of er een onterechte bias verwacht zou kunnen worden. Tijdens de evaluatie is het belangrijk om te toetsen of er onderdelen zijn waarop bepaalde groepen onbedoeld significant anders scoren. Hoewel het lastig is om dit goed te onderzoeken, vooral omdat veel achtergrondgegevens van kandidaten niet geregistreerd worden, ligt het zeker bij een nieuwe procedure voor de hand om hier aanvullend onderzoek naar te doen. Voor het verzamelen van gegevens van kandidaten is hun nadrukkelijke en geïnformeerde toestemming nodig. Daarnaast kunnen dit soort analyses vanuit onderzoeksethische overwegingen alleen uitgevoerd worden wanneer de uitkomsten niet herleidbaar zijn naar een of enkele individuen. Hiervoor moet de te onderzoeken groep groot genoeg zijn. Ook statistisch gezien zitten er beperkingen aan het vergelijken van kleine groepen. Mogelijke (achtergrond)kenmerken van kandidaten die voor dit type analyse van belang kunnen zijn:

- Genderidentiteit
- Migratieachtergrond
- Etnische identiteit
- Moedertaal
- Eerste generatie in het hoger onderwijs
- Opleiding en beroep van ouders/verzorgers
- Sociaaleconomische status ouders
- Dyslexie en overige functiebeperkingen
- Vooropleiding

### **Studentenaantallen en de vaardigheden van de studenten die binnenkomen**

Een belangrijke factor voor veel opleidingen, zeker bij 'harde' selectie, is wat de procedure doet met het aantal studenten dat uiteindelijk aan de opleiding begint. Een aantal overwegingen speelt hierbij in de evaluatie een mogelijke rol. Als er veel studenten worden geselecteerd die uiteindelijk toch besluiten niet met de opleiding te starten, kan het analyseren van de eigenschappen van die groep belangrijk zijn. Zijn er bijvoorbeeld bij bepaalde instrumenten al twijfels over hun match met de opleiding te vinden? Of moet het aantal toe te kennen plaatsen misschien omhoog (waar dat binnen de regelgeving een optie is)? Aan de andere kant kan een te strenge ondergrens voor toelating bij 'harde' selectie er ook voor zorgen dat een opleiding kleiner wordt.

Ook de ervaringen met bepaalde cohorten studenten die door middel van de selectieprocedure zijn toegelaten kunnen relevant zijn. Die signalen zijn vooral belangrijk om verder uit te zoeken, en niet direct voor feitelijk waar aan te nemen. Signalen kunnen bijvoorbeeld betrekking hebben op de houding of het welzijn van studenten, over de prestaties in onderdelen van

de opleiding, op groepsprocessen, et cetera. De potentiële bijdrage daaraan (positief of negatief) van de selectieprocedure kan onderwerp zijn van nadere evaluatie. Overigens kunnen zowel probleem als oplossing natuurlijk ook heel ergens anders liggen, en kan er sprake zijn van toevallige verschillen of dynamieken in een specifiek cohort.

### **Uitval**

Een doel van de selectieprocedure kan ook zijn om uitval tijdens de opleiding (door besluit van de student, of vanwege het bindend studieadvies) te minimaliseren. Het volledig voorkomen daarvan is een illusie, al was het maar omdat er ook factoren een rol kunnen spelen die nog niet bestonden op het moment van de selectie. Als een opleiding een beperkte instroom heeft, hadden andere kandidaten in het geval van uitval aan de opleiding kunnen beginnen. Bovendien is uitvallen ook voor de student zelf doorgaans geen prettige ervaring. Het analyseren van uitval is dan ook een waardevol onderdeel van de evaluatie. Zijn er signalen terug te vinden in de verschillende instrumenten die hierop hadden kunnen wijzen? Zegt het bijvoorbeeld iets over de weging van de instrumenten? Ook hierbij neemt de waarde van de analyse paradoxaal genoeg toe als de uitval groter is. Te veel focus op een kleine groep uitvallers kan ook een vertekend beeld geven, terwijl de oorzaak in omstandigheden kan liggen die buiten de selectie of zelfs helemaal buiten de opleiding vallen.









# Deel 2

## Instrumenten



Koppeling criteria en instrumenten	7
Vooropleidingscijfers	8
Kennistoetsen	9
Capaciteitstoetsen	10
Persoonlijkheidsvragenlijsten	11
Motivatiebrieven	12
CV	13
Aanbevelingsbrieven	14
Situational Judgement Tests	15
Interviews	16
Psychomotorisch	17
Loting	18
Afsluiting	19



## 7 Koppeling criteria en instrumenten

Het grootste deel van de literatuur over onderzoek naar selectie gaat over verschillende instrumenten die gebruikt kunnen worden in een selectieprocedure (Patterson et al., 2018). Het gaat dan vooral over de kwaliteiten van de instrumenten (betrouwbaarheid, voorspellende waarde) die gebruikt kunnen worden om af te wegen welke instrumenten je kiest (Patterson, Knight et al., 2016). Dit deel van het handboek gaat over de verschillende instrumenten die beschikbaar zijn, wat zij inhouden en met welke andere instrumenten zij zoal gecombineerd worden.

Dit deel is een aanvulling op de stappen uit het vorige deel. Het blijft van belang om eerst in beeld te hebben wat de procedure moet opleveren voordat de focus op selectie-instrumenten komt te liggen (Schreurs, Cleutjens, Collares, Cleland, & oude Egbrink, 2020). De belangrijkste vraag daarbij is welke vaardigheden het selectie-instrument moet meten en op welk niveau (Steenman, 2018), conform het idee van alignment (Biggs, 1996; zie ook [hoofdstuk 2](#)). We beginnen dit deel daarom met een overzicht van beschikbare instrumenten en welke type vaardigheden zij vooral kunnen meten. Ook het evalueren van de selectie-instrumenten en -procedure blijft van groot belang.

Eerst bespreken we per domein (cognitief, affectief en psychomotorisch) welke instrumenten vooral geschikt zijn en welke combinaties van instrumenten voor de hand liggen binnen dat domein. Meer detail per instrument wordt vervolgens in losse hoofdstukken besproken. In onderstaande tabel worden de domeinen, en voor het cognitieve domein ook de niveaus binnen het domein, afgezet tegen de instrumenten die we in dit handboek bespreken.

Instrumenten	Cognitief			Affectief	Psycho- motorisch
	LOCS	Toepassen	HOCS		
Vooropleidingscijfers (H8)	++	+	-	--	--
Kennistoetsen (H9)	++	--	--	--	--
Capaciteitstoetsen (H10)	-	++	++	+	--
Persoonlijkheidsvragenlijsten (H11)	-	-	--	+	--
Motivatiebrieven (H12)	-	-	-	+	--
CV (H13)	+	++	+	++	-
Aanbevelingsbrieven (H14)	+	++	++	++	-
Situational Judgement Tests (H15)	-	++	++	++	--
Interviews, ongestructureerd (H16)	--	--	--	--	--
Interviews, gestructureerd (H16)	-	++	+	+	--
Multiple Mini Interviews (H16)	-	++	++	++	-
Psychomotorisch (H17)	--	--	--	--	++
Loting (H18)	--	--	--	--	--

Let op: de bruikbaarheid van elk instrument is afhankelijk van de context en de implementatie!

++	Het instrument kan dit goed meten als je het goed en zorgvuldig inzet.
+	Het instrument kan gebruikt worden om dit te meten, mits je extra aandacht besteedt aan de manier waarop je het instrument inzet. Mogelijk zijn er betere alternatieven beschikbaar.
-	Het instrument is niet goed bruikbaar om dit te meten, maar in incidentele gevallen kan het zijn dat er geen betere of haalbare alternatieven zijn.
--	Het instrument is ongeschikt om dit te meten.

Figuur 9 • Alignment van instrumenten

## Cognitief

De instrumenten die in dit handboek beschreven worden voor het meten van het cognitieve domein zijn vooropleidingscijfers, kennistoetsen en capaciteitstoetsen. Hierbij zijn vooropleidingscijfers en kennistoetsen vooral geschikt voor het meten van LOCS. Capaciteitstoetsen zijn meestal gericht op het niveau van toepassen en HOCS. Verschillende instrumenten die in eerste instantie met name gericht lijken op het affectieve domein kunnen ook worden gebruikt voor vaardigheden in het cognitieve domein. Zo kunnen (gestructureerde) aanbevelingen iets zeggen over de eerdere prestaties van kandidaten in het cognitieve domein, en daarmee ook betekenisvol zijn voor het voorspellen van toekomstige prestaties in het cognitieve domein (zie ook figuur 9).

### LOCS

Als het gaat om het voorspellen van prestaties voor LOCS, laat onderzoek doorgaans zien dat vooropleidingscijfers het beste passen. Vooropleidingscijfers zijn betrouwbaar, voorspellend en simpel mee te nemen en vormen daarom ook vaak een belangrijk onderdeel van een selectieprocedure. Belangrijke kanttekening daarbij is echter de vergelijkbaarheid van vooropleidingscijfers uit verschillende onderwijssystemen (bijvoorbeeld in een vergelijking met kandidaten uit andere landen). Bij de toelating voor de master is ook voor bacheloropleidingen binnen Nederland de vergelijking vaak al complexer, maar zijn cijfers vaak goed te gebruiken in een selectieprocedure. Vooral als het gaat om resultaten in een aanpalende discipline en als vaardigheden overeenkomen is dat het geval (De Boer & Van Rijnsoever, 2022). Als vooropleidingscijfers minder betrouwbaar zijn ligt het gebruik van een kennistoets voor LOCS voor de hand. Ook in combinatie met vooropleidingscijfers kan een op het curriculum gebaseerde kennistoets bijdragen, mogelijk voornamelijk voor kandidaten met een lager vooropleidingscijfer (De Visser et al., 2017).

### Toepassen

Voor de vaardigheid *toepassen van kennis* kan mogelijk ook een deel van de vooropleidingscijfers gebruikt worden. Gemiddelde cijfers van het voortgezet onderwijs blijken doorgaans de vaardigheid *toepassen* echter veel minder goed te voorspellen (Steenman, Bakker & Van Tartwijk, 2016). Specifieke cijfers uit een bacheloropleiding zouden, afhankelijk van de toetsing, als een indicator voor toepassen gebruikt kunnen worden bij selectie voor een masteropleiding.

Kennistoetsen meten vooral het vermogen om kennis te begrijpen en te reproduceren en zijn daarmee niet geschikt als instrument om toepassingsvaardigheden te bepalen. Capaciteitstoetsen zijn beter geschikt. Capaciteitstoetsen zijn niet gericht op inhoudelijke kennis, maar zijn juist gericht op de vaardigheid, bijvoorbeeld het toepassen van kennis in een nieuwe context.

### HOCS

Van de toetsen die zich het meest direct lenen voor het cognitieve domein, is alleen de capaciteitstoets geschikt voor het meten van HOCS. Dat kan bijvoorbeeld in de vorm van een toets die kritisch denken meet. Verder laten HOCS zich ook meten met instrumenten die we onder het affectieve domein bespreken. Hiervoor kun je bijvoorbeeld een Multiple Mini Interview (MMI) gebruiken. In motivatiebrieven of aan referenten kan daarnaast gevraagd worden naar voorbeelden van het gebruik van HOCS door de kandidaten, of in een CV met extracurriculaire activiteiten kunnen kandidaten laten zien waar ze HOCS hebben gebruikt. Afhankelijk van de beschikbare informatie kunnen sommige vooropleidingscijfers vanuit de bacheloropleiding mogelijk gebruikt worden voor HOCS bij mastertoelating.

## Affectief

Het aantal instrumenten dat primair bedoeld is voor affectieve vaardigheden is groot. Het is een breed domein, waar veel verschillende vaardigheden onder verstaan kunnen worden. In dit boek behandelen we zes instrumenten, die alle ook op verschillende manieren gebruikt kunnen worden. Er is geen simpel antwoord op de vraag welk instrument in welke situatie het beste past, maar zoals altijd is een goede afstemming tussen wat in de opleiding gebruikt en getoetst wordt en de manier waarop het selectie-instrument wordt ingezet het beste uitgangspunt. Hieronder staan we kort stil bij de kern van de verschillende instrumenten; in de hoofdstukken per instrument werken we dat veel verder uit.

*Persoonlijkheidstests.* Elementen van persoonlijkheid kunnen goed gemeten worden. Het feit dat bij dit instrument doorgaans gebruik wordt gemaakt van zelfrapportage maakt het een lastiger instrument om in te zetten bij selectie: kandidaten kunnen zich heel makkelijk anders voordoen, op de manier die zij denken dat goed is voor hun toelatingskansen (Niessen, Meijer & Tendeiro, 2017b; Schripsema, Van Trigt, Van der Wal & Cohen-Schotanus, 2016). Daarbij is het vaak lastig te bepalen welke persoonlijkheid, en welke precieze scores op de verschillende vragen, het beste past bij een opleiding.

*Motivatiebrieven.* In de kern gaat het hierbij om tekst over de studiekeuze van kandidaten. Vaak worden motivatiebrieven gebruikt om na te gaan in hoeverre kandidaten passen bij de opleiding of de onderwijsinstelling. De vorm van de motivatiebrief die voor de selectie gevraagd wordt kan sterk verschillen. Het grootste verschil zit vaak in de hoeveelheid structuur die de kandidaten wordt meegegeven.

*Extracurriculaire activiteiten* worden binnen een selectieprocedure soms gevraagd om affectieve vaardigheden van kandidaten, zoals samenwerking, leidinggeven en communicatie, te beoordelen. Hier kan bijvoorbeeld naar gevraagd worden in de vorm van een CV. Ook in de vorm waarin dit instrument gebruikt wordt kan veel variatie zitten. Er zitten bijvoorbeeld veel keuzes in de manier waarop uit het CV een score komt. Extracurriculaire activiteiten worden vooral vaak gebruikt als maat voor motivatie, maar kunnen ook gebruikt worden om “ervaring met / vaardig zijn in” bepaalde competenties te beoordelen.

Ook voor *aanbevelingsbrieven* is de vorm van het instrument belangrijk: zonder structurering zijn aanbevelingsbrieven erg lastig van elkaar te onderscheiden en is de kans op bias groot. Maar met een goed uitgewerkte structurering (duidelijke vragen waar de *aanbeveler* antwoord op kan geven, liefst met vergelijkingen met andere leerlingen) lijken aanbevelingsbrieven



verschillende aspecten van de kandidaten te kunnen meten en daarmee ook bepaalde studieresultaten te kunnen voorspellen.

De *Situational Judgement Test (SJT)* of situatie-beoordelingstest is een brede categorie met veel afwegingen. SJT's zijn toetsen waarbij kandidaten situaties voorgelegd krijgen waarover vragen worden gesteld. Situaties kunnen zowel in de vorm van een stukje tekst voorgelegd worden als bijvoorbeeld als videofragment. Afhankelijk van het soort SJT dat gedaan wordt, kan gevraagd worden 'hoe zou je reageren', 'wat denk je dat de goede manier is om te reageren' of bijvoorbeeld 'hoe heb je op soortgelijke situaties in het verleden gereageerd'. Binnen een SJT kunnen ook uiteenlopende antwoordformats toegepast worden, zoals open vragen, meerkeuze-, orderings- of rankingvragen. De SJT kan op deze manieren (situatiebeschrijving, vraag, antwoordformat) volledig worden aangepast aan de context waarbinnen kandidaten moeten gaan functioneren. Een SJT leent zich dan ook om uiteenlopende capaciteiten te meten, denk bijvoorbeeld aan communicatie en samenwerking, maar ook ethisch bewustzijn, kritisch denken, en het inschatten van de waarden van kandidaten.

Ook bij *interviews* is de mate van structuur de belangrijkste onderscheidende factor tussen verschillende vormen. Ongestructureerde interviews zonder vooraf bedachte vragen voorspellen niets en zijn daarmee niet geschikt als selectie-instrument. *Semi-gestructureerde interviews* zijn vooral geschikt als aanvulling op andere instrumenten, zoals een aanbeveling of een SJT. *Gestructureerde interviews* kunnen daadwerkelijk kwaliteit en voorspellende waarde toevoegen aan je procedure. Een belangrijk voorbeeld van een gestructureerd interview is het MMI, het *Multiple Mini Interview*, waarbij kandidaten meerdere en qua onderwerp van elkaar verschillende korte gestructureerde interviews doorlopen.

## Psychomotorisch

Het laatste domein in de taxonomie van Bloom is het psychomotorische domein. Hierbij gaat het om fysieke beweging, coördinatie en in het algemeen het gebruik van motorische vaardigheden. Het ontwikkelen van deze vaardigheden vereist oefening en wordt gemeten in termen van snelheid, precisie, afstand, procedures of technieken in uitvoering (Bloom, 1956). Net als in de andere domeinen ligt de kern hier bij alignment tussen de selectie en de vaardigheden die in de opleiding belangrijk zijn. Als er duidelijke redenen zijn dat een goede basis van psychomotorische vaardigheden belangrijk is voor het presteren binnen de opleiding (zoals vaardigheden voor het bespelen van instrumenten voor het conservatorium, het beroep van tandarts of het beoefenen van sporten voor een sportopleiding), kan dat een goede reden zijn om dit als criterium in de selectieprocedure mee te nemen.

## De bespreking per instrument

Hieronder gaan we in aparte hoofdstukken in op alle instrumenten die hiervoor al kort aan de orde zijn gekomen. Al die hoofdstukken hebben een vergelijkbare structuur. We starten met een korte bespreking van *wat het instrument is*, meestal met ook een paar voorbeelden uit de praktijk. Die voorbeelden zijn niet bedoeld als ‘best practices’, maar om inzicht te geven in hoe het instrument er in de praktijk uit kan zien. In het volgende deel staan we stil bij *belangrijke afwegingen* bij de manier waarop het gebruikt kan worden. Onder ‘*Wat vinden kandidaten?*’ bespreken we wat we weten over hoe kandidaten het instrument ervaren. Dat doen we voornamelijk op basis van het brede onderzoek van Fikrat-Wevers et al. (2023), waarin kandidaten in een vragenlijst hebben aangegeven hoe ze tegenover een breed scala aan instrumenten staan. Dit onderzoek is uitgevoerd binnen een groep medische opleidingen in Nederland. In het onderzoek van Niessen, Meijer en Tendeiro (2017a) bij een psychologieopleiding zien we echter vergelijkbare resultaten. We eindigen elk hoofdstuk met een bespreking van *wat we weten uit onderzoek* en van *potentiële bias* voor elk instrument.

Bij die potentiële bias bespreken we zowel de onbedoelde factoren die onderdeel van de meting kunnen zijn, als de nadelen die specifieke groepen kandidaten kunnen hebben wanneer een instrument wordt gebruikt (wat vaak wordt aangeduid met *adverse impact*). De bias van een instrument kan bijvoorbeeld zijn dat in de beoordeling van een interview een belangrijke rol speelt of de geïnterviewde lijkt op de interviewer/beoordelaar (dat is niet de bedoeling, maar gebeurt toch). Adverse impact kan ontstaan omdat een bepaalde groep structureel minder goed scoort op een bepaalde vorm van het instrument. Bijvoorbeeld jongens die vaak een lagere score behalen op een persoonlijkheidstest voor nauwkeurigheid. Dat is geen bias in de meting, maar levert wel structureel nadeel voor een bepaalde groep op.

## 8 Vooropleidingscijfers

### Wat is het?

In dit eerste instrument is het uitgangspunt om gebruik te maken van informatie die al bekend is over eerdere prestaties. Welke cijfers beschikbaar zijn, verschilt natuurlijk per situatie. Bij selectie voor bacheloropleidingen gaat het vooral om cijfers uit het voortgezet onderwijs, voor masteropleidingen voornamelijk om cijfers uit de bacheloropleiding. In dit hoofdstuk gaan we in op het gebruik van cijfers van vooropleidingen voor selectie. De vakken die kandidaten hebben afgesloten (of nu volgen) worden ook vaak gebruikt om toelaatbaarheid te beoordelen. In dat geval ligt de nadruk op het beoordelen van voorkennis. In het handboek gaan we daar minder expliciet op in omdat toelaatbaarheid en selectie twee afzonderlijke fasen zijn in de toegang tot de opleiding.

Het uitgangspunt bij het gebruik van vooropleidingscijfers in de selectieprocedure is dat de prestaties die kandidaten eerder hebben geleverd ook iets zeggen over waartoe ze in de toekomst in staat zijn. Daarbij kan gebruik worden gemaakt van een algemeen gemiddelde, of van specifieke resultaten (bijvoorbeeld de bachelorscriptie bij de selectie voor een masteropleiding).

Cijfers uit het voortgezet onderwijs zijn vooral geschikt als instrument voor het meten van de lagere orde cognitieve vaardigheden, kennen en begrijpen (Steenman, Bakker & Van Tartwijk, 2016). Dat wijst erop dat toetsen uit het voortgezet onderwijs ook vooral die vaardigheden toetsen, dus dat daar een nadruk ligt op het toetsen van kennis en begrip. Hoewel de resultaten dus voorspellend zijn voor een uitsnede van alle vaardigheden die in de opleiding belangrijk zijn, zijn vaardigheden in het verwerven van kennis en begrip wel ook een belangrijk onderdeel van de meeste academische opleidingen.

Het principe achter deze redenering geldt ook in het algemeen voor cijfers van vooropleidingen. Wat er vooral in de vooropleiding getoetst wordt, is ook wat er met die cijfers het best voorspeld kan worden. Onderzoek van Vulperhorst, Lutz, De Kleijn en Van Tartwijk (2018) laat bijvoorbeeld zien dat het Internationaal Baccalaureaat (IB) als geheel andere vaardigheden meet dan vwo-cijfers. Voor vergelijkbare resultaten is een uitsnede van de kernvakken in het IB beter te vergelijken en dat is ook een betere voorspeller voor resultaten binnen de opleiding.

## Belangrijke afwegingen

De eerste belangrijke afweging is welke cijfers die potentieel beschikbaar zijn je wel of niet mee laat tellen in de selectie. Voor het gebruik van cijfers uit het voortgezet onderwijs gaat dat bijvoorbeeld over welke vakken je meeneemt. Dat kan het gemiddelde van alle vakken zijn, maar er zijn ook opleidingen die bijvoorbeeld alleen de bètavakken meetellen of alleen Nederlands, Engels en wiskunde. Onderzoek laat zien dat de voorspellende waarde van het gemiddelde van alle cijfers voor vrijwel alle opleidingen de grootste voorspellende waarde heeft voor resultaten binnen de opleiding. Een specifieke uitsnede van vakken blijkt maar zeer zelden sterker samen te hangen met prestaties binnen de opleiding, en als dat gebeurt blijkt het verschil klein te zijn en tussen cohorten niet consistent. De meest voor de hand liggende verklaring daarvoor lijkt te zijn dat er simpelweg meer meetmomenten zijn meegenomen als je meer vakken meeneemt.

Ook voor masteropleidingen lijkt doorgaans te gelden dat het gemiddelde van alle resultaten samen de beste voorspeller voor toekomstige prestaties is (De Boer, 2020). Daarbij geldt wel dat de voorspellende waarde van resultaten in de bachelor groter is als de inhoud van de bacheloropleiding meer overeenkomsten vertoont met de inhoud van de masteropleiding. Als het verschil in inhoudelijke focus tussen bachelor en master groter wordt, kan het daarmee waardevol zijn om ook andere instrumenten te gebruiken in de selectieprocedure.

Voor opleidingen waarbij de overgrote meerderheid van de kandidaten een vergelijkbare vooropleiding heeft (bijvoorbeeld havo of vwo), ligt het vrijwel altijd voor de hand om deze resultaten mee te nemen in de selectieprocedure omdat ze toekomstige prestaties voor LOCS goed voorspellen. Als de vooropleiding van kandidaten meer divers is, wordt dat ingewikkelder. Dat geldt bijvoorbeeld voor opleidingen met een grote internationale instroom met vooropleidingen uit zeer diverse onderwijssystemen. In die systemen worden andere vakken gegeven, die anders ingevuld worden, een ander niveau hebben, een andere tijdsinvestering vragen, et cetera. Dat maakt het vergelijken, en het brengen van al die resultaten op een enkele, betrouwbare schaal erg ingewikkeld.

Het Nuffic publiceert informatie die die vergelijking kan ondersteunen, maar ook met die informatie is een fijnmazige vergelijking vaak lastig. Een oplossing die sommige opleidingen gebruiken is om internationale cijfers niet om te rekenen naar de Nederlandse 1-10 schaal, maar om een grovere schaal te gebruiken, bijvoorbeeld een die van 1-4 loopt. Een andere optie zou kunnen zijn om andere instrumenten te gebruiken voor het beoordelen van lagere orde cognitieve vaardigheden in selectieprocedures, bijvoorbeeld een kennistoets.

Ook Nederlandstalige masteropleidingen kunnen te maken krijgen met een instroom van sterk uiteenlopende bacheloropleidingen. Omdat de vaardigheden die kandidaten nodig hebben gehad in hun bacheloropleiding dan waarschijnlijk ook zullen verschillen, zal de waarde van het gemiddelde cijfer dat kandidaten voor hun bacheloropleiding hebben behaald waarschijnlijk kleiner zijn dan wanneer iedereen een vergelijkbare vooropleiding heeft gevolgd. In die gevallen zou een optie kunnen zijn om alleen bepaalde typen resultaten (bijvoorbeeld voor kennistoetsen) mee te nemen. Dat vraagt natuurlijk wel relatief veel uitzoekwerk.

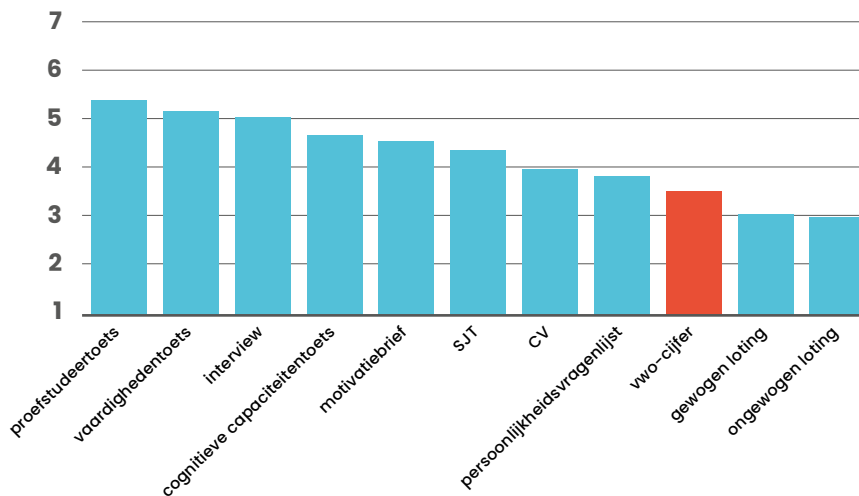
Zelfs als een grote meerderheid van de kandidaten een vergelijkbare vooropleiding heeft, is een belangrijke afweging wat je doet met kandidaten die een andere vooropleiding hebben. Verschillende opleidingen kiezen ervoor om die kandidaten voor dit instrument het gemiddelde van alle kandidaten toe te kennen, maar voor die kandidaten zou het ook een optie kunnen zijn om de resultaten die wel beschikbaar zijn om te zetten naar dezelfde schaal. Een afweging die hieraan raakt is of je bij kandidaten die een of meerdere tussenja(a)r(en) hebben genomen wel de eindresultaten meeneemt. Zo veel mogelijk vergelijkbare resultaten meenemen lijkt hier de voorkeur te hebben, maar het is praktisch niet altijd gemakkelijk om die te verkrijgen na het afronden van de vooropleiding.

Een andere belangrijke afweging is de peildatum. Daar zitten een aantal praktische beperkingen aan. De meeste opleidingen selecteren immers voordat de vooropleiding is afgerond. Bacheloropleidingen met een numerus fixus moeten de ranking bijvoorbeeld al bekendmaken vóór 15 april, geruime tijd voordat bij de meeste kandidaten het centraal schriftelijk eindexamen wordt afgenomen. Voor bacheloropleidingen geldt voor het gebruik van Nederlandse vooropleidingen dan ook vooral de vraag of het gemiddelde schoolexamencijfer op 15 januari wordt meegenomen, of het overgangsrapport tussen 5 en 6 vwo. In UNL-verband is voor academische opleidingen afgesproken om die laatste optie te hanteren, waarbij eventuele herkansingen die in de zomerperiode nog zijn afgelegd meegenomen worden. Ook wordt voor het doorgeven van deze cijfers gewerkt aan een uniform formulier voor alle academische opleidingen.

Een laatste belangrijke overweging is de druk die kandidaten kunnen ervaren tijdens hun vooropleiding doordat de cijfers daarvan worden meegenomen in de selectieprocedure. Dit speelt vooral voor opleidingen met veel kandidaten voor elke beschikbare plek, en waar kandidaten al vroeg een duidelijk beeld bij hebben. De meest voor de hand liggende voorbeelden zijn de opleidingen geneeskunde en diergeneeskunde. Aan de andere kant heeft dit instrument niet het nadeel van de druk op een enkel toetsmoment. Afwegingen rond de druk die kandidaten tijdens hun vooropleiding ervaren kunnen ook een rol spelen bij het bepalen welk peilmoment gebruikt wordt en eventueel zelfs welke vakken worden meegenomen.

### Wat vinden kandidaten?




Veel kandidaten zijn niet erg positief over vooropleidingscijfers als selectie-instrument. Onderzoek naar ervaringen van kandidaten voor bachelor selectieprocedures laat zien dat het meenemen van cijfers uit het voortgezet onderwijs maar nipt beter beoordeeld wordt dan (gewogen en ongewogen) loting.



Figuur 10 • Wat vinden kandidaten

Kandidaten geven onder andere aan dat ze vinden dat ze zich nog snel ontwikkelen en dat het dus oneerlijk is om cijfers uit het verleden mee te nemen. Bovendien wordt er niet meegenomen onder welke omstandigheden een cijfer tot stand is gekomen, bijvoorbeeld met behulp van bijles, naast heel veel sporten of werken, of op een middelbare school die zeer goed/minder goed bekendstaat.

In figuur 11 staat welke argumenten in het onderzoek van Fikrat-Wevers (2022) door kandidaten vooral worden gegeven.

		
Voorargumenten	Tegenargumenten	Voorwaarden
<ul style="list-style-type: none"> <li>• Relevante informatie</li> <li>• Geen momentopname</li> <li>• Je best doen loont</li> </ul>	<ul style="list-style-type: none"> <li>• Mist relevante informatie</li> <li>• Te generiek</li> <li>• Geen toegevoegde waarde</li> <li>• Informatie uit het verleden</li> <li>• Verhaal achter cijfer onbekend</li> <li>• Toegankelijkheid (bijles, geslacht, vooropleiding)</li> <li>• Grote verschillen tussen scholen</li> </ul>	<ul style="list-style-type: none"> <li>• In combinatie met andere methodes</li> <li>• Recente cijfers</li> <li>• Goede informatievoorziening</li> </ul>

Figuur 11 • Argumenten van kandidaten

### Wat weten we uit onderzoek?

De betrouwbaarheid van combinaties van eerdere cijfers is goed, en dit is ook een van de sterkste punten aan het gebruik van vooropleidingscijfers voor selectie. Een algemene toetsingsregel is dat hoe meer toetsen er meegenomen worden in een gemiddelde, hoe betrouwbaarder dit gemiddelde is: bij één toets meet je altijd ook specifieke toevalligheden die speelden bij die toets terwijl die bij meerdere toetsen steeds meer uitmiddelen. Het effect van één toets meer neemt wel af, maar blijft iets bijdragen. Voor selectie voor masters gaat dezelfde redenering op: het combineren van meer cursussen met meer toetsen geeft een betrouwbaarder beeld van het algemene studievermogen van de kandidaat.

Cijfers uit het voortgezet onderwijs zijn een van de beste voorspellers van latere prestaties in de bacheloropleiding, in het bijzonder voor de prestaties op kennis- en begripstoetsen in de opleiding. Hoe groot die voorspellende waarde is hangt af van welke cijfers je gebruikt en hoe en wat er in de opleiding zelf getoetst wordt. Bij een breed scala aan internationale vooropleidingen, met andere beoordelingssystemen, neemt de voorspellende waarde snel af. Voor een groep kandidaten die allemaal van het vwo of de

havo afkomen, zijn de havo- of vwo-cijfers betrouwbare metingen van een algehele vaardigheid.

Meerdere (internationale) studies hebben ook laten zien dat vooropleidingscijfers correleren met latere prestaties (Patterson, Knight et al., 2016). Richting het eind van de opleiding en de start van de master neemt de voorspellende waarde van de cijfers van de middelbare school af. Dat komt vooral omdat dan vaak de hogere niveaus binnen Bloom's taxonomie (toepassen en HOCS) steeds belangrijker worden en ook andere vaardigheden steeds meer aan bod komen (Schreurs, 2020). Als die andere vaardigheden belangrijk zijn in een opleiding ligt het (aanvullend) gebruik van andere instrumenten dan ook voor de hand.

Ondanks dat vooropleidingscijfers een sterke voorspeller zijn van toekomstige prestaties, verklaren ze altijd maar een deel van de toekomstige academische prestaties (Ferguson, James & Madeley, 2002; McManus, Woolf, Dacre, Paice & Dewberry, 2013). Er zijn enkele aanwijzingen dat de toegevoegde waarde van aanvullende instrumenten daarnaast groter kan zijn voor kandidaten met een lager gemiddelde in het voortgezet onderwijs (De Visser et al., 2017)

### Potentiële bias

Binnen Nederland zijn de meeste middelbare scholen relatief vergelijkbaar in kwaliteit en resultaten. Desalniettemin treden er ook bij het gebruik van cijfers uit het voortgezet onderwijs verschillen op, onder andere doordat bijles niet voor alle leerlingen op dezelfde manier toegankelijk is. Bias kan ook verder veroorzaakt worden door de context waarbinnen resultaten behaald worden, zoals hulp van ouders of taalachtergrond. Zo laten Fikrat-Wevers, Stegers-Jager en anderen (2023) zien dat studenten waarvan de ouders niet gestudeerd hebben significant lagere cijfers in het voortgezet onderwijs behalen. Ook zijn er significante verschillen tussen de gemiddelde cijfers van vrouwelijke en van mannelijke kandidaten, zeker voor het schoolexamen.





## 9 Kennistoetsen

### Wat is het?

Er zijn verschillende instrumenten die vaardigheden binnen het cognitieve domein meten. De kennistoets is een cognitieve toets die kennis en begrip (LOCS) meet over een bepaald onderwerp. Wat een kennistoets onderscheidt van bijvoorbeeld een capaciteitstoets is dat er bij een kennistoets meestal voorbereidende stof verstrekt wordt en dat het doel is te achterhalen welke kennis de kandidaat heeft vergaard in plaats van wat de kandidaat met de kennis doet (dit laatste past meer bij een capaciteitstoets). Een proefstudeertoets kan dan ook in een van of beide categorieën vallen, afhankelijk van wat er in die toets gevraagd wordt.

Een kennistoets kan een meer algemene, gestandaardiseerde toets zijn (bijvoorbeeld om de biologiekennis in kaart te brengen) of juist heel sterk op de opleiding gebaseerd zijn waarvoor geselecteerd wordt (bijvoorbeeld een tentamen dat gaat over stof uit de eerste periode van de opleiding). Daarnaast kan het gaan om kennis waarvan je verwacht dat kandidaten die al hebben, of kennis die studenten voor de toets moeten opdoen. Grofweg betekent dat dat er op basis van deze twee eigenschappen vier verschillende soorten kennistoetsen zijn:

- 1. Algemeen, geen voorbereiding nodig.** Hier gaat het om standaard toetsen die kennis meten die kandidaten in hun vooropleiding opgedaan zouden moeten hebben. Een voorbeeld hiervan is onderdeel 2 van de BioMedical Admissions Test (BMAT). Hierin zitten vragen over biologie, natuurkunde, scheikunde en wiskunde, zoals:

One healthy male human testis cell undergoes meiosis.  
How many chromosomes, in total, are present after meiosis is completed?  
A 23      B 44      C 46      D 92      E 184

- 2. Op het programma gericht, geen voorbereiding nodig.** Hierbij zou het bijvoorbeeld kunnen gaan om een tentamen of essay zoals dat ook in het curriculum van studenten gevraagd zou worden, over een onderwerp waar zij de kennis al over bezitten vanuit de vooropleiding. Zeker voor masteropleidingen is dit een aantrekkelijke optie om ook de voorkennis te controleren, maar ook voor geneeskunde zou je bijvoorbeeld naar specifieke voor de opleiding belangrijke biologievoorkennis kunnen vragen op een manier die in lijn is met de toetsen binnen de opleiding.
- 3. Algemeen, voorbereidingsstof gegeven.** Hierbij zou je een toets gebruiken om te onderzoeken of je kandidaten voldoende in staat zijn kennis te vergaren (je toetst de vaardigheid om kennis op te doen, zonder dat de context van de opleiding daarin belangrijk is). Hierbij zou je de kandidaten informatie kunnen geven over een niet direct aan de opleiding gerelateerd onderwerp, waarvan het onwaarschijnlijk is dat ze daarmee eerder in contact zijn gekomen. Op die manier kun je meten hoe kandidaten volledig nieuwe kennis verwerken.
- 4. Op het programma gericht, voorbereidingsstof gegeven.** Hier gaat het, net als bij optie 2, om een toets die je binnen het programma zelf zal moeten ontwikkelen. De kern is hier dat het om een toets gaat die goed aansluit op de vorm en inhoud van het daadwerkelijke curriculum. Vaak kijken opleidingen daarvoor naar onderwerpen die in het eerste jaar van de opleiding aan bod komen en schrijven ze kandidaten literatuur voor die daarin terugkomt of daar sterk op lijkt. Ook in de vorm ligt het voor de hand om aan te sluiten bij hoe kennis in de opleiding wordt getoetst, bijvoorbeeld in een meerkeuzetentamen, of in de context van een casus.

Kennistoetsen kunnen prestaties op vergelijkbare toetsen in de bacheloropleiding vaak goed voorspellen, en worden dan ook relatief vaak gebruikt. Hoe beter de aansluiting tussen de toets en de vorm en inhoud van de opleiding, hoe beter de voorspellende waarde doorgaans is. De kwaliteit van de toets speelt daarbij natuurlijk wel een grote rol. Daarnaast is een enkele toets gevoelig voor toevalligheden die op de dag van de toets gebeuren, en is het goed om stil te staan bij de verhouding tussen wat de kennistoets meet en de informatie die eventueel via vooropleidingscijfers te achterhalen is.

Hieronder staat een voorbeeld van hoe het gebruik van een kennistoets wordt uitgelegd op de [website](#) van de opleiding toegepaste psychologie aan de Hanzehogeschool in Groningen.

In deze selectieprocedure wordt zowel kennis als inzicht getoetst, gebaseerd op een greep uit de inhoud van de opleiding. Dit ziet er als volgt uit:



- het volgen van een hoorcollege en meerdere flitscolleges van twee representatieve cursussen
- het volgen van een werkcollege (ter voorbereiding maak je opdrachten uit het werkboek)
- afsluitende (online) toets bestaande uit drie onderdelen: Inleiding in de Psychologie, Onderzoek en Statistiek en een onderdeel over het beeld van de opleiding en toekomstige beroepsmogelijkheden.

Je kunt je voorbereiden op de toets door het volgen van de colleges, het bestuderen van de literatuur en beschikbare materialen op de digitale leeromgeving, je in te lezen in het vakgebied van Toegepaste Psychologie, het bestuderen van de website en het bezoeken van onze open dagen en meeloopdagen. Na (tijdige) aanmelding krijgen ze toegang tot de digitale leeromgeving.

## Belangrijke afwegingen

Als kandidaten vooraf leerstof wordt aangereikt ter voorbereiding is een belangrijke afweging hoeveel tijd je kandidaten geeft voor de voorbereiding en hoeveel leerstof je geeft. De selectie voor een bacheloropleiding vindt plaats op het moment dat veel kandidaten in hun eindexamenjaar zitten. Voor een master geldt dat veel kandidaten op dat moment bezig zijn hun bachelor af te ronden. Het moeten voorbereiden van leerstof voor een selectieprocedure komt bovenop de leerstof en opdrachten die ze voor hun huidige opleiding moeten doen in een voor veel kandidaten drukke en stressvolle periode. Hoewel het moeten voorbereiden van wat grotere hoeveelheden leerstof in kortere tijd wellicht dicht ligt bij wat er in de opleiding gevraagd wordt, waardoor het een beter beeld kan geven van de relevante vaardigheden van kandidaten, is het goed dit ook af te wegen tegen de mogelijk negatieve consequenties hiervan voor kandidaten.

De vaardigheden die je met een kennistoets meet, overlappen vaak veel met de vaardigheden die gereflecteerd worden in vooropleidingscijfers. Het verschil is dat vooropleidingscijfers vaak een gemiddelde zijn over meerdere toetsen die op verschillende momenten in de tijd zijn gemaakt. Een kennistoets is slechts één toetsmoment waardoor de uitslag meer beïnvloed wordt door externe omstandigheden, zoals stress of slecht slapen. Hiermee is een kennistoets een minder betrouwbaar selectie-instrument. Als het doel is om een beeld te krijgen van de vaardigheid van kandidaten in het verwerven en begrijpen van leerstof ligt de inzet van een kennistoets dus niet altijd voor de hand. Als er geen goed vergelijkbare vooropleidingscijfers beschikbaar zijn, wordt de afweging anders. Dat geldt bijvoorbeeld bij internationale instroom, bij instroom vanuit veel verschillende bacheloropleidingen of vanuit

mbo-opleidingen. Daarnaast kan een kennistoets die een afspiegeling is van de inhoud van de opleiding het beeld dat kandidaten van de opleiding hebben versterken, en kan de motivatie voor deze kennistoets anders zijn dan de motivatie tijdens de vooropleiding.

Zoals hierboven beschreven zijn er de meer algemene kennistoetsen of kennistoetsen die worden afgestemd op de opleiding. De aanschaf van een gestandaardiseerde kennistoets ligt in Nederland niet voor alle opleidingen voor de hand. Anders dan in sommige andere onderwijssystemen hebben kandidaten in het voortgezet onderwijs op de meeste vlakken vergelijkbare basiskennis meegekregen, zodat die basiskennis voor bacheloropleidingen met voornamelijk directe instroom uit het voortgezet onderwijs niet opnieuw hoeft te worden vastgesteld. Er wordt dan ook meestal gebruikgemaakt van een toets die wordt afgestemd op de opleiding.

Het ontwikkelen van een goede kennistoets vraagt tijd, maar je kunt de leerstof en inhoud van de toets precies afstemmen op wat er in de opleiding verwacht wordt. Het is belangrijk om tijdig met de ontwikkeling van de toets te starten. Een belangrijke afweging is daarbij bijvoorbeeld ook de taligheid in de vragen. Bij erg talige vragen gaat taalvaardigheid van kandidaten vanzelfsprekend een steeds grotere rol spelen in de resultaten, waardoor kandidaten die bijvoorbeeld thuis een andere taal spreken mogelijk in het nadeel zijn. Het is aan te bevelen om bij het ontwikkelen van de kennistoets in een vroeg stadium ook een toetsdeskundige te betrekken.

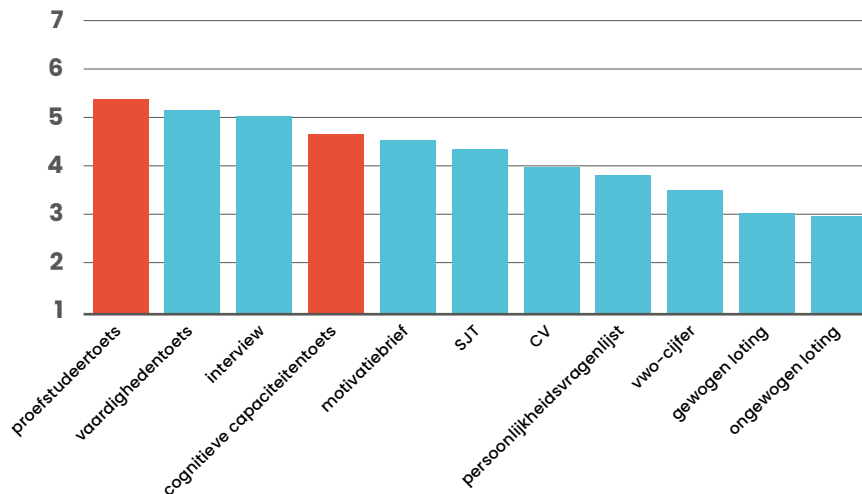
Een relatief nieuwe overweging rondom het afnemen van kennistoetsen (die ook opgaat voor een groot deel van de capaciteitstoetsen, maar ook SJT's) is of je de toets online of op locatie zou willen afnemen. Beide opties hebben voor- en nadelen. Online heeft voor de meer internationaal georiënteerde opleidingen als belangrijk voordeel dat kandidaten ook vanuit het buitenland zonder lang reizen kunnen deelnemen. De andere kant is dat online toetsen meer mogelijkheden tot fraude geven. Er wordt in die situatie dan ook doorgaans gebruikgemaakt van *proctoring*. Deze vorm van digitaal toezicht stelt wel veel eisen aan de apparatuur, internetverbinding en omgeving van de kandidaat. Toetsen op locatie heeft als voordeel dat je beter op fraude kunt controleren. In dat geval vraagt de logistiek op locatie wel meer aandacht.

## Wat vinden kandidaten?

In de internationale context lijken kandidaten geen bezwaren te hebben tegen het gebruik van kennistoetsen (Patterson, Knight et al., 2016), maar als ze direct vergeleken worden met andere toetsen geven kandidaten aan een voorkeur te hebben voor bijvoorbeeld capaciteitstoetsen of Situationele Beoordelingstoetsen (SJT's) (Koczwara et al., 2012). Als het gaat om de keuze



tussen gestandaardiseerde kennistoetsen of toetsen die meer op de opleiding gebaseerd zijn, hebben kandidaten de voorkeur voor meer opleidingsgebaseerde toetsen. Dit blijkt ook uit het onderzoek van Suzanne Fikrat-Wevers (2023), waarin kennistoetsen die een goede afspiegeling vormen van de opleiding voor kandidaten zelf de meest gewaardeerde instrumenten blijken te zijn. Gestandaardiseerde kennistoetsen komen op de vierde plaats uit.



Figuur 12 • Wat vinden kandidaten

## Wat weten we uit onderzoek?

Wat betreft voorspellende waarde blijkt dat kennistoetsen prestaties in de opleiding kunnen voorspellen. Ze voorspellen met name hoe kandidaten zullen presteren in kennisgebaseerde bacheloropleidingen of de op kennis gerichte elementen binnen een opleiding, maar ook hoe snel kandidaten door de opleiding heen gaan (Hissbach, Klusmann & Hampe, 2011). Ook voor masterselectie lijken kennistoetsen bruikbaar, en laten bij het gebruik voor postdoctorale selectie gunstige betrouwbaarheid en validiteit zien (Patterson & Zibarras, 2018, p. 5). In een directe vergelijking met capaciteitstoetsen en SJT's (Situational Judgement Tests) lijkt de voorspellende waarde van kennistoetsen soms wel onder te doen voor die van deze andere toetsen (Koczwara et al., 2012).

Eerder onderzoek heeft ook laten zien dat de inhoud van een kennistoets van belang is voor de voorspellende waarde ervan (Sackett, Walmsley, Koch, Beatty & Kuncel, 2016). In Nederland heeft een kennistoets gericht op de inhoud van het curriculum een betere voorspellende waarde laten zien voor resultaten in de opleiding dan een algemene kennistoets. De inhoudelijke kennistoets komt qua voorspellende waarde zelfs in de buurt van cijfers uit

het voortgezet onderwijs (Niessen, Meijer & Tendeiro, 2016). Dat is allemaal natuurlijk sterk afhankelijk van de kwaliteit van de toets. Het is dus belangrijk om de gebruikte toets steeds goed te evalueren.

### Potentiële bias

Een kennistoets is een instrument dat gevoelig is voor coaching. De prestatie van de kandidaat kan dus beïnvloed worden door het deelnemen aan een (betaalde) cursus. Dat kan ongelijkheid met zich meebrengen (zie ook [hoofdstuk 5](#)). Coaching voor de selectie van sommige opleidingen lijkt erg populair onder kandidaten, vooral wanneer het verschil tussen het aantal aanmeldingen en het aantal plaatsen toeneemt. Dit lijkt de uitkomsten van selectieprocedures minder voorspellend te maken (Griffin, Yeomans & Wilson, 2013) en geeft verschillende groepen kandidaten ongelijke toegang tot ondersteuning. Kennistoetsen lenen zich meer voor coaching dan toetsen op een hoger cognitief niveau of affectieve toetsen (Krathwohl, 2002; Patterson, Knight et al., 2016). Vanwege de ongelijke toegang tot betaalde coaching bieden verschillende opleidingen gratis coaching aan voor kandidaten die daar belangstelling voor hebben. Ook de tijd om de toets voor te bereiden is ongelijk verdeeld. Kandidaten die bijvoorbeeld veel tijd kwijt zijn aan mantelzorgtaken kunnen bijvoorbeeld mogelijk minder gemakkelijk tijd vinden voor de voorbereiding van de toets.

Uit het onderzoek van Fikrat-Wevers et al. (2023) blijkt dat kandidaten met een niet-westerse migratieachtergrond (zowel eerste als tweede generatie) een lagere score behaalden op proefstudeertoetsen dan kandidaten zonder migratieachtergrond. Mogelijk hebben zij minder kans om voor te bereiden. Het is ook mogelijk dat er een culturele bias in bijvoorbeeld de vraagstelling van de toets zelf zat. Ook kandidaten met een buitenlandse vooropleiding bleken in dit onderzoek lager te scoren op de proefstudeertoets. Mogelijke verklaringen daarbij zijn de taligheid van de toets of het feit dat de toets een andere vorm had dan waaraan kandidaten gewend waren.

# 10 Capaciteitstoetsen

## Wat is het?

Een capaciteitstoets is een instrument dat bedoeld is om te meten wat de mogelijkheden van kandidaten zijn om bepaalde vaardigheden te ontwikkelen (Patterson, Ferguson & Knight, 2010, p. 359). Capaciteitstoetsen kunnen worden ontworpen om een breed scala aan vaardigheden te meten. Het is wereldwijd een breed gebruikt instrument, ook bijvoorbeeld voor personeelsselectie.

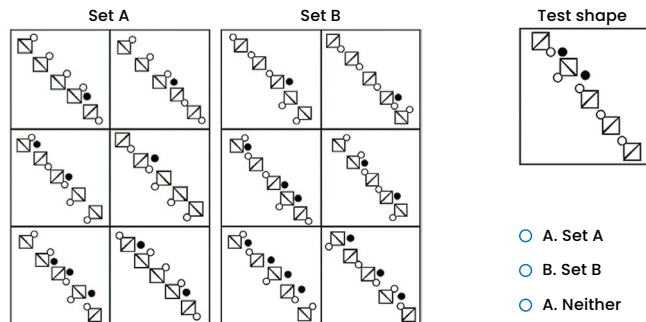
Een capaciteitstoets meet niet de kennis die studenten hebben, maar gaat over cognitieve of andere vaardigheden. Het kan dus bijvoorbeeld gaan over de vaardigheden die studenten hebben in het verwerken van informatie op een hoger cognitief niveau, in plaats van over de kennis zelf. Dat is dan ook het grote verschil met de kennistoets die in het vorige hoofdstuk centraal stond. Capaciteitstoetsen kunnen zich bijvoorbeeld ook richten op de vaardigheid om kritisch te redeneren of om informatie te gebruiken in een casussituatie. De vorm van het instrument kan een soort tentamen zijn, maar ook bijvoorbeeld een opdracht die kandidaten uitvoeren.

Hoewel bij het toetsen van die vaardigheden natuurlijk wordt uitgegaan van enige basiskennis, staat die in de toetsing bij dit instrument niet centraal. Er is meestal geen sprake van specifieke voorbereiding of leerstof voor de toets. Kennis waarvan niet verondersteld kan worden dat kandidaten die hebben, wordt dan ook meestal in de toets zelf gegeven. Hieronder worden enkele voorbeelden beschreven van hoe capaciteitstoetsen ingezet worden in selectieprocedures.

### Voorbeelden

Als onderdeel van medische selectieprocedures bij veel Britse, Australische en Nieuw-Zeelandse universiteiten wordt de UCAT afgenomen, de University Clinical Aptitude Test. Die test wordt bijvoorbeeld gebruikt om vaardigheden in abstract redeneren te meten. In het kader staat een voorbeeld van een vraag uit deze test.

Een voorbeeld van een (oefen)item uit de subtest ‘abstract redeneren’ van de UK Clinical Aptitude Test (UKCAT), waarbij kandidaten moeten aangeven bij welke set de ‘test shape’ behoort.



10; Uit Patterson & Zibarras (2018), p. 29.

Bij de selectie voor de opleidingen van de Academie Verloskunde Amsterdam / Groningen wordt ook een capaciteitstoets gebruikt. Kandidaten worden in die selectieprocedure gevraagd om feedback te geven aan een andere kandidaat. Daarnaast moeten ze reflecteren op de feedback die ze zelf ontvangen. Hierbij staat de vaardigheid van het geven van feedback (en de cognitieve processen die daarbij horen, zoals evalueren) centraal. Ook in de reflectie worden de reflectievaardigheden beoordeeld. Deze vaardigheden sluiten aan bij wat er binnen de opleiding van studenten gevraagd wordt:

**Werkgroep.** Net als in het eerste studiejaar van de opleiding gaan de kandidaten tijdens de tweede ronde van de selectie in werkgroepen aan de slag met een studietaak. Deze taak sluit aan bij het onderwerp uit het zelfstudiepakket dat voor de eerste ronde van de selectie is bestudeerd.

Tijdens het werken aan de studietaak beoordelen twee assessoren (beoordelaars) de gespreks-, samenwerkings-, en andere contactuele vaardigheden. Het materiaal voor deze taak wordt door de opleiding verstrekt. Het is niet toegestaan zelf materiaal mee te nemen.

**Peerassessment.** Na het afronden van de studietaak wordt aan de kandidaat gevraagd schriftelijk feedback te geven op het functioneren van één ander lid van de werkgroep. Met deze opdracht wordt gekeken hoe vaardig de kandidaat is in het geven van feedback aan een medekandidaat. Dit is iets wat ook in de opleiding regelmatig gebeurt.



**Schriftelijke reflectie.** Tot slot wordt aan de kandidaat gevraagd schriftelijk te reflecteren op competenties van het beroepsprofiel. Tijdens deze opdracht wordt gekeken naar zowel de reflectievaardigheden, de schriftelijke communicatievaardigheden en het toepassen van de reflectiemethodiek.

Ook in Rotterdam wordt voor de opleiding geneeskunde gebruikgemaakt van een capaciteitstoets. Het gaat hierbij om een openboektoets (om ervoor te zorgen dat voorkennis geen rol speelt in de prestaties op de toets) waarin de focus ligt op het begrijpen en toepassen van de stof. Op de website van de opleiding wordt deze toets als volgt beschreven:

Op de toetsdag worden voor de bacheloropleiding geneeskunde benodigde vaardigheden getoetst. Het gaat om studievoordigheden zoals het verzamelen en lezen van informatie, het onderscheiden van hoofd- en bijzaken, bestuderen van teksten, aantekeningen maken e.d. Academische vaardigheden zoals analyseren van artikelen, interpreteren van resultaten en numerieke vaardigheden worden tevens getoetst. Voor deze toetsen is, anders dan bovengenoemd, geen verdere voorbereiding nodig.

Voor masteropleidingen wordt internationaal veel gebruik gemaakt van de GMAT (Graduate Management Admission Test®). De GMAT meet verbaal (zinscorrectie, begrijpend lezen en kritisch redeneren), kwantitatief (probleemoplossing en dataverwerking), geïntegreerd redeneren (informatie-evaluatie) en analytisch schrijven (analyse van een probleem en argument). Het is een computergestuurde test met vier getimedede onderdelen.

## Belangrijke afwegingen

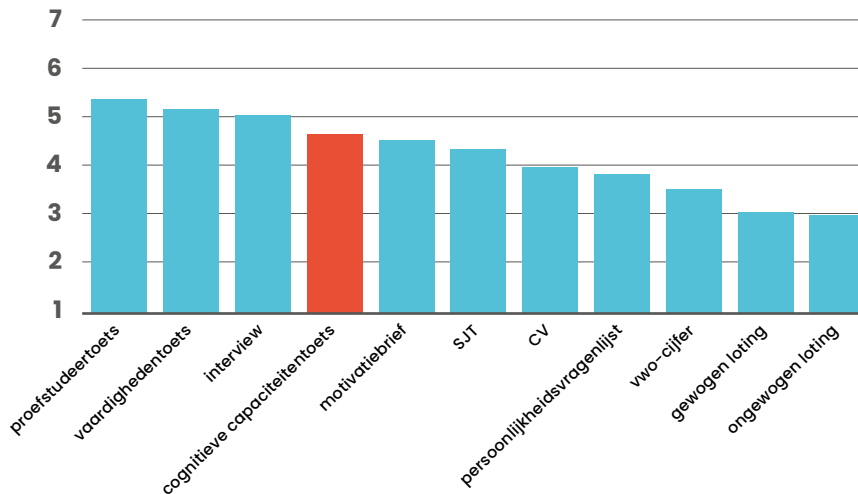
Capaciteitstoetsen kunnen heel breed zijn: je kunt dit instrument op verschillende manieren inzetten en veel verschillende capaciteiten meten. De resultaten van een kandidaat zullen vooral iets zeggen over hoe die kandidaat het in de opleiding gaat doen op het moment dat dezelfde soort vaardigheid gevraagd wordt als er in de capaciteitstoets gemeten is. Als er binnen de opleiding vooral wordt ingezet op het aanleren van kennis, ligt het dus niet voor de hand een capaciteitstoets op te nemen in de selectie, maar voldoet een kennistoets. Is het doel echter dat studenten zich gedurende de opleiding ontwikkelen tot bijvoorbeeld probleemoplossers die met nieuwe situaties om kunnen gaan, dan is het opnemen van een capaciteitstoets in je selectie een goede manier om de alignment tussen selectie en opleiding te waarborgen.

Een capaciteitstoets kan meer algemeen zijn of afgestemd worden op de opleiding. Zeker voor het meten van cognitieve vaardigheden zijn er veel opties voor het aanschaffen van een gestandaardiseerde, meer algemene toets (zoals de UCAT of GMAT uit de voorbeelden hierboven). Dit geeft veel gemak en bespaart een hoop tijd die anders nodig zou zijn om zelf een toets te ontwikkelen. Aan de andere kant is een meer algemene toets niet afgestemd op de opleiding. Het geeft de kandidaten dus geen informatie over de opleiding. Afhankelijk van wat je precies wilt meten en hoe de rest van de procedure eruitziet kan dit een nadeel zijn.

Bij het ontwikkelen van een capaciteitstoets kan het tijdsaspect een afweging zijn in de keuze voor het type vragen dat gesteld wordt. Het stellen van open vragen kan ertoe leiden dat er een beter beeld ontstaat van de kandidaat, maar zorgt er ook voor dat de beoordeling meer tijd kost en beter afgestemd moet worden in het geval van meerdere beoordelaars. Deze afwegingen kunnen per opleiding en selectieprocedure verschillen.

### Wat vinden kandidaten?

Uit het onderzoek van Fikrat-Wevers et al. (2023) blijkt dat kandidaten capaciteitstoetsen redelijk waarderen. Net als bij kennistoetsen geldt ook hier dat hoe meer de inhoud en vorm in lijn zijn met de inhoud en vorm van de opleiding, hoe hoger de toets gewaardeerd wordt door de kandidaten.



Figuur 13 • Wat vinden kandidaten

Ook uit internationaal onderzoek blijkt dat deelnemers deze toetsen als nuttig ervaren en dat de toetsen hiermee een belangrijk onderdeel van de selectieprocedure kunnen vormen (Patterson, Knight et al., 2016).

## Wat weten we uit onderzoek?

Het ligt voor de hand, maar het bewijs voor de betrouwbaarheid, validiteit en voorspellende waarde van capaciteitstoetsen hangt af van de specifieke invulling daarvan en van hoe gedegen het instrument is ontwikkeld. Uit onderzoek blijkt dan ook dat er wisselend bewijs is over de betrouwbaarheid en, met name, de voorspellende waarde (Patterson & Zibarras, 2018, p. 5). Dat vraagt daarmee ook iets van de evaluatie van het individuele instrument in de context van de opleiding. Het is wel mogelijk om een betrouwbare capaciteitstoets te ontwikkelen; zie ook het eerste deel van het handboek over het ontwikkelen en evalueren van instrumenten.

Onderzoek laat zien dat goede capaciteitstoetsen binnen selectieprocedures prestaties in de opleiding kunnen voorspellen. Als capaciteitstoetsen in combinatie met vooropleidingscijfers worden gebruikt in een selectieprocedure, blijkt dat capaciteitstoetsen toegevoegde waarde kunnen hebben, hoewel de extra toegevoegde voorspellende waarde vaak niet heel groot is. Door het toevoegen van capaciteitstoetsen naast vooropleidingscijfers kan een breder scala aan relevante vaardigheden worden meegenomen in de selectie, waardoor prestaties op specifieke vaardigheden binnen de opleiding mogelijk wel veel beter voorspeld kunnen worden. Capaciteitstoetsen zijn over het algemeen het meest geschikt om cognitieve prestaties te voorspellen, maar kunnen afhankelijk van de inhoud ook belangrijke informatie uit het affectieve domein opleveren.

Op het gebied van bestaande tests is het bewijs soms wisselend. Het is dus goed om eerst goed uit te zoeken wat voor kennis beschikbaar is voordat je van deze tests gebruikmaakt. Recent onderzoek zet bijvoorbeeld bij enkele aspecten van de GMAT vraagtekens. Onderzoek van Kass, Kim, LaMacchia en Bekisz (2020) laat zien dat de GMAT weinig toevoegt in de voorspelling van prestaties binnen een opleiding, en ander onderzoek geeft een indicatie dat het weglaten van de GMAT uit een selectie de diversiteit van de toegelaten kandidaten verhoogt (Hooker, Mallow, Downes & Baidwan, 2022).

## Potentiële bias

Uit onderzoek komen wisselende beelden over de eerlijkheid en vergelijkbaarheid van resultaten op de toetsen van groepen deelnemers met specifieke achtergronden (denk aan verschillen in leeftijd, geslacht, moedertaal, SES). Sommige onderzoeken laten weinig tot geen verschillen zien voor specifieke toetsen, andere onderzoeken laten zien dat er wel degelijk verschillen kunnen optreden. Het is dus belangrijk om bij het ontwerpen het potentieel voor bias expliciet mee te nemen, en de toetsen op dit vlak goed te evalueren. Een ‘kritisch redeneren’-toets kan bijvoorbeeld heel talig zijn, en daarmee bepaalde

groepen op achterstand stellen, hoewel het ook mogelijk is een instrument te ontwikkelen dat kritisch redeneren meet zonder dit nadeel.

In recent onderzoek waarin selecties voor een aantal medische opleidingen meegenomen worden, blijken capaciteitstoetsen gericht op het curriculum weinig bias op SES te vertonen. Kandidaten met een niet-westerse migratieachtergrond blijken echter wel lager te scoren (Fikrat-Wevers et al., 2023). Of dit voor een specifieke toets geldt is natuurlijk afhankelijk van de specifieke kenmerken van een toets.

Een voordeel van capaciteitstoetsen is dat ze weinig vatbaar zijn voor coaching. Daarvoor is het dan wel belangrijk om ieder jaar op zijn minst een aantal vragen te laten rouleren of details aan te passen. Er is wel onderzoek dat laat zien dat oefenen met dit soort toetsen de prestaties van kandidaten verbetert (Puddey, Mercer, Andrich & Styles, 2014). Desalniettemin lijkt het bij een goed ontwikkeld instrument nog steeds mogelijk om het vermogen om met vergelijkbare, hoewel nieuwe problemen om te gaan, vooral onder tijdsdruk, goed te meten (Patterson & Zibarras, 2018).

# 11 Persoonlijkheidsvragenlijsten

## Wat is het?

Persoonlijkheid is een ingewikkeld begrip. Hoewel we er allemaal een idee bij hebben is het lastig te definiëren. Persoonlijkheid is te verdelen in wat je van buitenaf kunt zien (gedrag) en wat je niet kunt zien (gevoelens, emoties, opvattingen) (Kohnstamm, 2009). De aanname achter persoonlijkheidsvragenlijsten is dat de uitkomsten een verband hebben met hoe iemand zich gedraagt, niet alleen in het persoonlijke leven, maar ook met betrekking tot opleiding en werk. Hiermee kun je (tot op zekere hoogte) voorspellen hoe iemand zich op een opleiding gedraagt en daarmee ook een deel van de prestaties en of iemand bij het type opleiding past.

Voor het meten van persoonlijkheidskenmerken binnen een selectie is een grote variatie aan persoonlijkheidsvragenlijsten beschikbaar (Woods & Anderson, 2016). Meestal wordt de kandidaat hierbij gevraagd zichzelf te scoren op een aantal stellingen of kenmerken. Een veelgebruikte persoonlijkheidsvragenlijst is de *Big Five*, waarin aan de hand van vijf algemene kenmerken iemands persoonlijkheid wordt beschreven (Kohnstamm, 2009; Meeus, 2019). Voor het meten daarvan wordt gebruikgemaakt van standaard vragenlijsten; in een klinische context wordt vaak gekozen tussen bijvoorbeeld de NEO-PI-R, de NEO-FFI (een verkorte versie van de NEO-PI-R) of de Nederlandse Persoonlijkheids Vragenlijst (NPV).

Het kiezen van een goede vragenlijst is vaak al ingewikkeld, en daarnaast zijn er naast de *Big Five* ook nog verschillende andere theorieën over persoonlijkheid. Zo is er ook een theorie die uitgaat van 6 persoonlijkheidsdimensies, HEXACO: ‘Honesty-humility’ (Integriteit), ‘Emotionality’ (Emotionaliteit), ‘eXtraversion’ (Extraversie), ‘Agreeableness’ (Verdraagzaamheid), ‘Conscientiousness’ (Consciëntieusheid) en ‘Openness to experience’ (Openheid voor ervaringen). Ook hier zijn vragenlijsten voor te vinden, evenals voor andere theorieën.

## Belangrijke afwegingen

De samenhang tussen persoonlijkheid en prestaties is complex. Wat in de ene context positief werkt, kan in de andere context juist tegenwerken. Een voorbeeld daarvan is dat het persoonlijkheidskenmerk ‘zorgvuldigheid’ een positief effect heeft op iemands prestaties in het (bachelor) onderwijs, terwijl uit onderzoek bij geneeskunde blijkt dat meer zorgvuldigheid niet direct positief is voor prestaties in de medische beroepspraktijk. Studenten met een hoge zorgvuldigheid presteerden in de kliniek soms structureel lager dan minder zorgvuldige studenten (Patterson, Knight et al., 2016). Een eerste belangrijke afweging is dus het zorgvuldig bepalen (en onderzoeken) welke persoonlijkheidskenmerken voor de opleiding belangrijk zijn.

Persoonlijkheidskenmerken hebben vaak twee uitersten. Zorgvuldigheid kan belangrijk zijn, maar te zorgvuldig zijn gaat over in rigiditeit, met negatieve gevolgen voor prestaties in een omgeving met weinig structuur en voorspelbaarheid. Als je overweegt persoonlijkheidsvragenlijsten mee te nemen in je selectie, is hoe je omgaat met extreem hoog of laag scorende kandidaten dan ook een belangrijke afweging.

Verder is een belangrijke overweging of een diversiteit aan persoonskenmerken niet meer waarde heeft voor de opleiding dan kandidaten met een specifieke eigenschap. Vooral als persoonlijkheidskenmerken veel gewicht krijgen in de selectiebeslissing kan dat een negatief effect hebben op de diversiteit in de studentenpopulatie (Patterson, Knight et al., 2016).

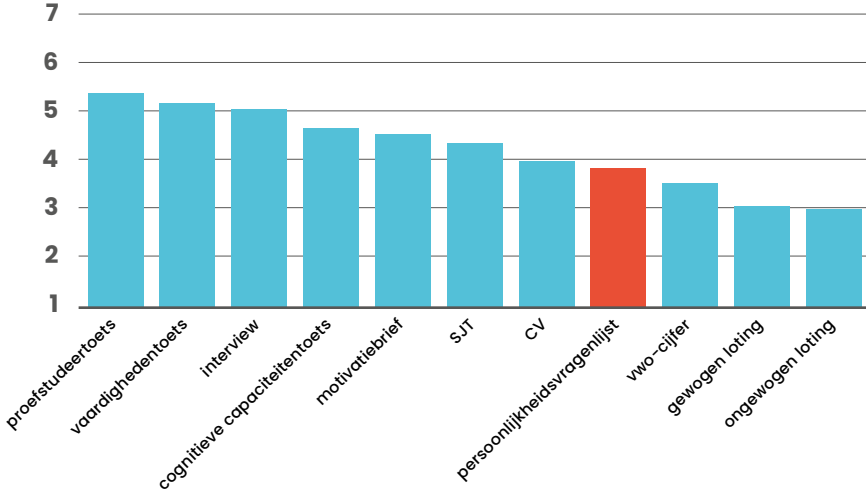
Als je een persoonlijkheidsvragenlijst wilt meenemen, is een belangrijke afweging welke vragenlijst je daarvoor wilt gebruiken. Patterson en Zibarras (2018) gaan verder in op waar je op moet letten bij de keuze voor een vragenlijst voor een selectieprocedure. Een belangrijke kanttekening is dat de meeste van deze vragenlijsten gevalideerd zijn voor de klinische praktijk. Dat is niet een-op-een hetzelfde als de context van een selectieprocedure. De inschatting van kandidaten wat de gewenste eigenschappen of antwoorden zijn kan in een situatie waarin toelating kan afhangen van de score soms belangrijker zijn dan welk antwoord ze bij zichzelf het beste vinden passen (Prideaux et al., 2011; Schripsema et al., 2016). Een manier om hiermee om te gaan kan zijn om ook een bekende van de kandidaat te vragen de lijst over de kandidaat in te vullen.

Een laatste belangrijke afweging is welke rol de persoonlijkheidsvragenlijsten in het geheel van de procedure spelen. Dat bepaalt ook in hoeverre sommige van de hierboven beschreven nadelen van het instrument van toepassing zijn. Een optie is bijvoorbeeld om de persoonlijkheidsvragenlijst te gebruiken als input voor een interview, zodat kandidaten hun antwoorden op de vragenlijst verder kunnen toelichten en hierop verder bevraagd kunnen

worden. Ook kan van kandidaten worden gevraagd om op basis van de uitkomsten van de test te reflecteren op situaties binnen de opleiding waarbij het persoonlijkheidskenmerk effect kan hebben.

**Wat vinden kandidaten?**

Persoonlijkheidsvragenlijsten worden door kandidaten niet sterk positief gewaardeerd als onderdeel van een selectieprocedure. Kandidaten begrijpen vaak niet wat het nut is van een persoonlijkheidstest voor een opleiding, en geven zelf ook aan dat de methode erg gevoelig is voor sociaal wenselijke antwoorden. Het onderzoek van Fikrat-Wevers et al. (2023) laat zien dat kandidaten persoonlijkheidsvragenlijsten maar net boven de beide vormen van loting en het gebruik van vwo-cijfers waarderen.



Figuur 14 • Wat vinden kandidaten

**Wat weten we uit onderzoek?**

De betrouwbaarheid en validiteit van persoonlijkheidsvragenlijsten is niet gunstig in de context van een selectieprocedure. Persoonlijkheidsvragenlijsten zijn ontwikkeld voor de klinische psychologie, waarbinnen ze een hoge betrouwbaarheid en validiteit hebben (e.g. Fossati, Borroni, Marchione & Maffei, 2011). Zodra deze vragenlijsten echter gebruikt worden in situaties waarbij er veel van de resultaten afhangt (high-stake situaties), waar selectie overduidelijk een voorbeeld van is, blijkt het simuleren van goed gedrag (sociaal wenselijk antwoorden) een groot probleem te worden (Patterson & Zibarras, 2018; Schripsema, 2017). Uit onderzoek blijkt dat er grote verschillen gevonden worden tussen persoonlijkheidsvragenlijsten die

tijdens en na de selectie worden ingevuld (Griffin & Wilson, 2012; Kuncel & Borneman, 2007). Verder is de kwaliteit van de persoonlijkheidsvragenlijst in grote mate afhankelijk van de specifieke vragenlijst die er gebruikt wordt.

Persoonlijkheid is een aspect dat veel onderzocht wordt en waarbij ook de relatie met prestaties in studie en werk meegenomen wordt. Meerdere studies laten zien dat persoonlijkheid een relatie vertoont met prestaties op school en op het werk. Zo worden iemands mate van zorgvuldigheid (Poropat, 2009) en integriteit (De Vries, Reinout, Born & Van den Berg, 2014) positief in verband gebracht met latere prestaties in de opleiding. Ook andere persoonlijkheidskenmerken van de *Big Five* kunnen correleren met verschillende prestaties in de opleiding (Patterson & Zibarras, 2018).

### Potentiële bias

Zoals eerder genoemd kan het (zwaar) selecteren op bepaalde persoonlijkheidskenmerken ervoor zorgen dat de diversiteit van verschillende typen studenten in de opleiding kleiner wordt (Patterson & Zibarras, 2018). Hier komt nog bij dat van bepaalde persoonlijkheidskenmerken bekend is dat specifieke groepen kandidaten gemiddeld structureel anders scoren dan anderen. Een voorbeeld hiervan is dat mannelijke kandidaten gemiddeld lager scoren op zorgvuldigheid dan vrouwelijke kandidaten. Dat kan ook voor andere groepen gelden.

Verder kunnen persoonlijkheidsvragenlijsten gevoelig zijn voor coaching. Kandidaten die zich anders presenteren dan ze in werkelijkheid zijn, kunnen hierdoor een hogere score behalen en zo voordeel ondervinden in de selectieprocedure, zeker als ze hulp krijgen van experts in dit type instrument. Dat kan eisen stellen aan de manier waarop persoonlijkheidsvragenlijsten in de procedure worden gebruikt en hoe daarover wordt gecommuniceerd.



## 12 Motivatiebrieven

### Wat is het?

Een motivatiebrief vatten we breed op als elke tekst waarin kandidaten uitleggen waarom ze graag aan een opleiding zouden willen studeren of waarom hun interesse of vaardigheden passen bij de vorm en inhoud van de opleiding. Dat hoeft dus niet per se over motivatie te gaan, en het hoeft ook niet per se in de vorm van een brief te gebeuren. Het belang van motivatie van studenten voor de opleiding die ze volgen blijkt wel duidelijk uit onderzoek naar motivatie (Deci & Ryan, 2012). Ook onderzoeken die bekijken wat maakt dat studenten goed functioneren, benadrukken de rol van motivatie (zie bijvoorbeeld: Liu, Bridgeman & Adler, 2012; Pinxten, Van Soom, Peeters, De Laet & Langie, 2019). Wat een motivatiebrief precies meet, hangt af van wat kandidaten gevraagd wordt, hoeveel structuur daarbij meegegeven wordt en hoe de brieven worden beoordeeld en gewogen.

Motivatiebrieven worden op veel verschillende manieren gebruikt en beoordeeld. Het is binnen Nederland een instrument dat relatief vaak wordt gebruikt in selectieprocedures. In de praktijk gaan motivatiebrieven vooral over de aansluiting tussen interesses en de inhoud en vorm van de opleiding. Het gaat dus om een specifieke invulling die dicht ligt bij het idee van matching of de studiekeuzecheck. Motivatie om bijvoorbeeld hard te werken is ook lastig te beoordelen, vooral omdat het erg veranderlijk is (De Boer & Van Rijnsoever, 2021). Een goed gesprek met een docent kan dit type motivatie bijvoorbeeld erg versterken. Interesse in de inhoud van de opleiding, en aansluiting bij de manier waarop kandidaten het liefste leren en de manier waarop het onderwijs bij de opleiding wordt aangeboden bieden veel meer handvatten. Soms worden motivatiebrieven ook gebruikt om achtergrondinformatie te verzamelen, bijvoorbeeld om te begrijpen waarom kandidaten langer over een vooropleiding hebben gedaan. Hieronder staat een aantal voorbeelden van de manier waarop motivatiebrieven worden gebruikt.

### Voorbeelden

Bij psychologie in Groningen wordt een motivatiebrief op de volgende manier gebruikt:

Om deel te kunnen nemen aan de selectieprocedure moet iedere kandidaat een motivatiebrief schrijven waarin je aangeeft waarom je Psychologie in Groningen zou willen studeren. Het schrijven van de brief is vooral bedoeld om ervoor te zorgen dat jouw keuze voor Psychologie weloverwogen gemaakt is. We beoordelen of je het schrijven van je motivatiebrief voldoende serieus hebt genomen, maw dat je daadwerkelijk beschrijft waarom je Psychologie wilt studeren. Op deze wijze gaan wij na of jouw aanmelding serieus is. Daarnaast biedt dit gelegenheid tot essentiële reflectie op jouw keuze. De specifieke inhoud van de motivatie zal niet meegewogen worden bij de selectie, hiervoor wordt de selectietoets gebruikt.

De motivatiebrief wordt dus niet meegewogen in de daadwerkelijke selectieprocedure, maar is een voorwaarde om deel te mogen nemen aan de selectie. De opleiding geeft aan dat matching hier vooral het doel is: kandidaten worden gevraagd zich nog eens goed in te lezen in de opleiding en na te gaan of deze echt bij ze past. Dit kan ervoor zorgen dat kandidaten die bij het inschrijven nog geen goede afweging hebben gemaakt vroeg in het proces uitvallen.

Je zou ook een motivatiebrief kunnen gebruiken waarin je net iets meer richting en structuur geeft, bijvoorbeeld door een aantal algemene vragen te laten beantwoorden. De opleiding Bestuurs- en Organisationswetenschap van de Universiteit Utrecht vraagt kandidaten bijvoorbeeld om een aantal vragen te beantwoorden, zoals:

- Waarom vind je B&O inhoudelijk interessant? (150 woorden)
- Waarom kies je ervoor om deze opleiding juist in Utrecht te volgen? (let op: deze vraag gaat over de opleiding, niet over waarom je de stad Utrecht leuk vindt) (100 woorden)
- Heb je een voorlichtingsdag en/of middagje studeren bij B&O bijgewoond? Zo nee, waarom niet? Zo ja, wat sprak je aan in de didactiek en/of inhoud? (100 woorden)
- Wat denk je dat het belangrijkste is dat je bij de opleiding B&O gaat leren dat je later in een baan kan gebruiken? (Licht toe) (100 woorden)

In dit voorbeeld wordt meer structuur geboden door specifieke vragen te stellen.

Je kunt ook de structuur die je meegeeft aan de motivatiebrief versterken zonder het in los te beantwoorden vragen te vervatten. Het volgende fictieve voorbeeld laat zien hoe dat vorm zou kunnen krijgen:



### Schrijf een motivatiebrief

Wil je de voltijdsopleiding Fysiotherapie volgen, dan moet jouw brief uiterlijk 22 januari in ons bezit zijn. Hieronder vind je de eisen waaraan jouw brief moet voldoen:

1. Je persoonlijke gegevens (volledige naam, emailadres, telefoonnummer, woonadres)
2. Uitleg van je motivatie voor de studie Fysiotherapie (maximaal 400 woorden)
  - a. Waarom wil je de studie Fysiotherapie volgen?
  - b. Hoe heb je informatie verzameld over de studie Fysiotherapie?
  - c. Welke andere opleidingen heb je bekeken, en waarom kies je toch voor Fysiotherapie?
3. Keuze voor Fysiotherapie in Groningen (maximaal 200 woorden)
  - a. Fysiotherapie kun je op verschillende hogescholen in Nederland studeren. Ze leiden je op tot dezelfde kwalificatie. De wijze waarop de verschillende hogescholen dat doen is echter verschillend. Waarom heb je gekozen voor de opleiding Fysiotherapie aan ons instituut?
4. Keuze voor het beroep 'Fysiotherapeut' (maximaal 200 woorden)
  - a. Waarom denk jij in de toekomst een goede Fysiotherapeut te zijn?
  - b. Wat maakt dat jij een betere Fysiotherapeut zou zijn dan bijvoorbeeld je klasgenoten?

### Belangrijke afwegingen

De mate van structuur is een belangrijke afweging bij motivatiebrieven, de voorbeelden laten dat ook zien. Een minder gestructureerde opdracht kan goed aansluiten bij het doel om kandidaten zich in de opleiding te laten verdiepen en na te laten gaan of deze past bij hun interesse (Schreurs, Cleutjens, Collares et al., 2020). Een duidelijk gestructureerde brief past bij het doel de motivatie van kandidaten te beoordelen en mee te wegen in de selectie. Het is dan van belang dat kandidaten goed weten wat ze moeten doen en waar ze op beoordeeld zullen worden (Niessen & Neumann, 2022). Zonder een duidelijke instructie komen in de brieven vaak uiteenlopende thema's aan de orde, naast de reden voor het kiezen voor een bepaalde opleiding. In veel brieven gaan kandidaten bijvoorbeeld in op relevante levenservaring, toekomstige carriëredoelen en persoonlijkheidseigenschappen (Ding, 2007).

Of je de motivatiebrieven vooral gebruikt als een voorwaardelijke eis, of als instrument dat ook in de uiteindelijke beslissing meeweegt, is een andere belangrijke afweging. In het eerste geval moet de vraag die aan kandidaten gesteld wordt vooral zorgen dat zij zelf goed nadenken over hun keuze, en

zorgen dat ze zich verdiepen in de specifieke kenmerken van de opleiding. Studenten die met de opleiding stoppen kan bijvoorbeeld worden gevraagd naar de reden daarvoor, als basis voor de vragen die je stelt. De beoordeling zelf is in dat geval natuurlijk niet belangrijk, behalve dat er mogelijk bepaald moet worden of de opdracht voldoende serieus is uitgevoerd. Ook heldere communicatie over je verwachting op dit gebied richting kandidaten lijkt daarbij voor de hand te liggen, zoals in het eerste voorbeeld in dit hoofdstuk ook gebeurt. Daarnaast is hierbij een afweging of deze stap in de procedure voldoende opweegt tegen de tijdsinvestering die je van kandidaten vraagt extra belangrijk.

Als de motivatiebrief meetelt in de uiteindelijke selectiebeslissing, is het een belangrijke afweging hoe die beoordeeld wordt. Dat hangt allereerst af van het doel van de motivatiebrief in het geheel van de selectieprocedure. Op basis van wat je wilt meten en wat je van kandidaten kunt verwachten op basis van de opdracht die je meegeeft, ligt het dan voor de hand om vooraf helder te maken wat je precies wilt beoordelen, wat voor reacties je als goed of minder goed zou beoordelen, en hoe je die verschillende elementen mee wilt tellen. Dat kan bijvoorbeeld in de vorm van duidelijke rubrics.

Als motivatiebrieven duidelijker gestructureerd zijn, is het ook makkelijker de brieven op specifieke aspecten te scoren. Een overweging is daarbij ook of het een volledige brief moet zijn, of dat het los beantwoorden van enkele vragen over aspecten van motivatie beter is. Dat laatste structureert de beantwoording en kan het daarmee potentieel gemakkelijker maken om te beoordelen. Als er meerdere beoordelaars zijn, is een belangrijke afweging ook hoe je vorm kunt geven aan onderlinge afstemming en kalibratie.

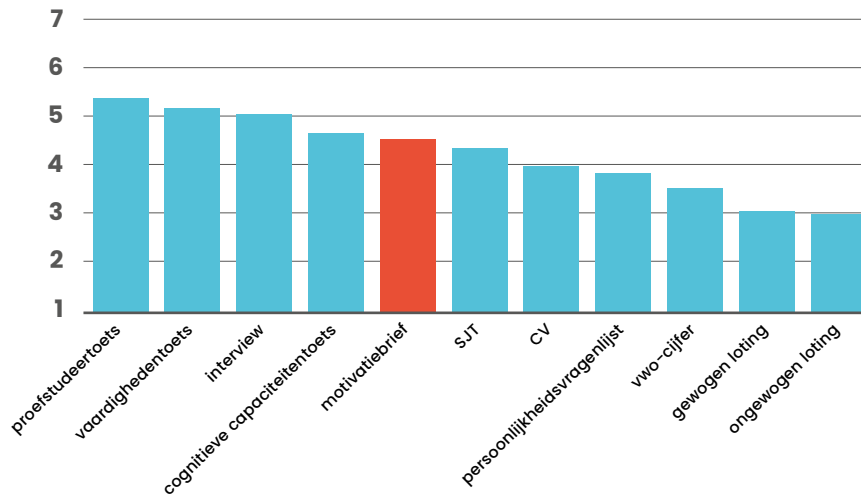
Een laatste afweging gaat over de condities waaronder je de motivatiebrief laat schrijven. Het gevaar van de motivatiebrief thuis laten schrijven, is dat je niet zeker weet dat de kandidaat de brief ook daadwerkelijk zelf heeft geschreven. Ook het gebruik van AI-tools voor het schrijven van motivatiebrieven kan in die afweging een rol spelen. Het is daarom een overweging om de motivatiebrief bijvoorbeeld tijdens een selectiedag te laten schrijven op locatie. Dat is natuurlijk een belangrijkere afweging als de motivatiebrief meetelt in de selectie dan wanneer dat niet het geval is.

## Wat vinden kandidaten?

Uit het onderzoek van Fikrat-Wevers en collega's (2023) blijkt dat kandidaten motivatiebrieven redelijk positief waarderen. Wat daarbij wel kan meespelen is dat motivatiebrieven een veelgebruikt instrument zijn: kandidaten zijn geneigd bekende instrumenten beter te waarderen dan instrumenten die ze niet kennen (Niessen & Neumann, 2022). Daarnaast geven



kandidaten aan behoefte te hebben om te laten zien wie ze zijn en wat ze in huis hebben, en dat in de motivatiebrief goed naar voren te kunnen brengen. Kandidaten stellen het op prijs om hun verhaal te kunnen vertellen (Wouters, Bakker, Van Wijk, Croiset & Kusrkar, 2014).



Figuur 15 • Wat vinden kandidaten

### Wat weten we uit onderzoek

Het bewijs voor de voorspellende waarde van motivatiebrieven is gemengd: in sommige onderzoeken wordt voorspellende waarde aangetoond, maar een groot deel van de onderzoeken laat zien dat de betrouwbaarheid en validiteit tekortschieten (Patterson, Knight et al., 2016). Ondanks de zorgen over de effectiviteit en mogelijke input van derden worden motivatiebrieven over de hele wereld veel gebruikt in selectie. Recent onderzoek bij de faculteit Geowetenschappen van de Universiteit Utrecht op basis van automatische tekstanalyse van motivatiebrieven laat zien dat het bespreken van bepaalde onderwerpen samenhangt met de resultaten die studenten in de opleiding behalen. Het bespreken van onderwerpen die inhoudelijk specifiek passen bij een bepaalde opleiding blijkt bijvoorbeeld significant samen te hangen met resultaten in de opleiding. Dat geldt in dit onderzoek bijvoorbeeld niet voor het bespreken van generieke doelen, extracurriculaire activiteiten of het willen hebben van maatschappelijke impact.

Wouters, Bakker, Van Wijk, Croiset en Kusrkar (2014) laten zien dat het vaststellen van autonome motivatie voor studeren vanuit een motivatiebrief lastig te gebruiken is in een selectieprocedure. Het ligt daarom minder voor de hand om motivatiebrieven te gebruiken voor meer dan voor het zoeken

naar een goede match tussen de inhoud en didactiek van de opleiding en de interesses van kandidaten.

### Potentiële bias

De hulp die kandidaten uit hun omgeving kunnen krijgen bij het schrijven van een motivatiebrief thuis is niet voor iedereen gelijk. Hierdoor kunnen specifieke groepen kandidaten voordeel hebben bij het gebruik van dit instrument. Een mogelijkheid is om de brief op locatie te laten schrijven, zodat de brief in elk geval door de kandidaat zelf geschreven is (Patterson, Knight et al., 2016). Daarnaast kan een meer gestructureerde brief helpen om bias te verkleinen. Bij een ongestructureerde brief wordt van kandidaten verwacht dat ze weten wat er in een motivatiebrief thuishoort. Dat veronderstelt een bepaald sociaal kapitaal, bijvoorbeeld om in te schatten of het een goed idee is om in te gaan op de locatie van de opleiding (leuke stad of mooi gebouw) of juist niet.

Ook in het beoordelen van motivatiebrieven schuilt potentie voor bias. Als bepaalde typen motivatie als minder waardevol worden beoordeeld, zou dat kunnen betekenen dat bepaalde groepen nadeel ondervinden. Een voorbeeld daarvan is wanneer opmerkingen over toekomstige baankansen of salaris negatief worden gewaardeerd, terwijl dergelijke overwegingen voor bepaalde groepen kandidaten van groot belang kunnen zijn. Ook kunnen bepaalde persoonlijkheidskenmerken (bijvoorbeeld extraversie) in motivatiebrieven meer of juist minder gewaardeerd worden. Als ingetogen reflecteren op eerdere prestaties bijvoorbeeld beter gewaardeerd wordt dan het breed uitmeten daarvan levert dat eveneens een potentiële bias op voor kandidaten met bepaalde persoonlijkheidskenmerken.



# 13 CV

## Wat is het?

Een CV (curriculum vitae) is een format waarin kandidaten hun eerdere ervaringen kunnen opnemen. Op de arbeidsmarkt is het CV een veelgebruikt instrument in sollicitatieprocedures voor een baan (Mogridge, 2019). In selectieprocedures voor opleidingen wordt de focus bij dit instrument vaak gelegd op de ervaringen die de kandidaat naast de vooropleiding heeft opgedaan ('extracurriculaire verdiensten').

Het vragen van een CV of een overzicht van extracurriculaire activiteiten kan verschillende doelen dienen binnen selectie. Allereerst kan een CV ingezet worden in de rol van een matchingactiviteit, waarbij het doel vooral is dat kandidaten zelf reflecteren op de verbinding tussen hun eerdere ervaringen en de inhoud of vorm van de opleiding. Een ander doel is om door het beoordelen van de activiteiten op het CV te bepalen of kandidaten beschikken over vaardigheden die van pas komen in de opleiding. Eventueel kan deze informatie dienen als input voor een interview, waarbij kan worden doorgevraagd naar wat er tijdens de activiteiten geleerd is en wat daarvan de link is met de opleiding.

Heeft een kandidaat naast de opleiding geen andere activiteiten gedaan, zoals een bijbaan, veel sporten of mantelzorg verlenen, dan heeft deze alle mogelijkheid gehad (of genomen) om zich op de vooropleiding te richten. Een kandidaat met dezelfde cijfers, maar daarnaast wel veel andere activiteiten, heeft mogelijk in minder tijd dezelfde resultaten behaald. Dit plaatst de behaalde cijfers in een ander perspectief en dat kan eventueel meegewogen worden.

De voorbeelden die hieronder beschreven staan, illustreren dat CV's en extracurriculaire activiteiten in verschillende vormen en op verschillende momenten in een selectieprocedure kunnen worden ingezet. Ook kunnen ze

op verschillende manieren worden vormgegeven. Voornamelijk verschillen in structurering van de manier waarop de informatie wordt gevraagd zijn duidelijk zichtbaar tussen verschillende opleidingen.

### Voorbeelden

Mondzorgkunde aan de HAN maakt gebruik van een CV als onderdeel van een voorbereidingsopdracht op de selectieprocedure. Het CV-onderdeel van de voorbereidingsopdracht wordt als volgt beschreven:

#### Deel 1: CV

Hierin vul je jouw gegevens in zoals in een curriculum vitae (CV). Zo hebben wij gelijk een goed overzicht van jouw gegevens, vooropleiding en eventuele werkervaring of andere activiteiten.

#### Wat levert het jou op?

Deze opdracht maken zorgt ervoor dat jij je uitgebreid oriënteert op verschillende aspecten van de opleiding en het beroep. En daarmee ben je goed voorbereid op de verdere selectie! Ook krijg je een goede indruk van de soort schriftelijke opdrachten die je als student op de opleiding Mondzorgkunde krijgt. Het volledig maken en indienen van de voorbereidingsopdracht is voorwaardelijk voor deelname aan de verdere selectie.

Het voltooien van de voorbereidingsopdracht is in dit geval een voorwaarde om mee te doen aan de selectieprocedure. De inhoud van het CV wordt niet beoordeeld. Dit is in lijn met het doel van deze opdracht, namelijk het oriënteren op de opleiding (en het latere beroep) waar de kandidaten zich voor inschrijven.

Bij mondzorgkunde aan de Hogeschool Utrecht vormt het CV ook een onderdeel van de selectie :

**Opdracht:** De leerling maakt de persoonlijkheidstest via de onderstaande link en beschrijft de uitkomsten op basis van onderstaande leerdoelen in een verslag van 1A4. Verder wordt er met het reflectieverslag een CV meegestuurd.  
*Persoonlijkheidstest: <https://www.jobpersonality.com/big-five-persoonlijkheidstest>*

#### Leerdoelen:

- De leerling betreft het 'Beroepsprofiel Mondhygiënist in Nederland' en de studieids 'opleiding Mondzorgkunde Hogeschool Utrecht' bij het uitwerken van geschiktheid voor de opleiding op basis van persoonlijkheidskenmerken uit de test verkregen
- De leerling onderbouwt de geschiktheid voor de studie en beroep, d.m.v. reflectie op sterke en minder sterke punten vanuit de persoonlijkheidstest





- De leerling onderbouwt zichzelf als gemotiveerde kandidaat
- De leerling reflecteert grondig op de persoonlijkheidskenmerken verkregen uit de persoonlijkheidstest
- De leerling schrijft de vorm van de reflectie en het CV in correct Nederlands

Hierin wordt een hele duidelijke structuur geboden aan de kandidaten en wordt ook deels meegegeven waarop de CV's beoordeeld zullen worden (correct Nederlands). Kandidaten weten wat ze moeten invullen en een van de doelen lijkt te zijn om bewustwording bij de kandidaten te bewerkstelligen. Het is niet helemaal duidelijk wat precies de inhoud van het CV moet zijn.

Een andere studie waarin naar onderscheidende vaardigheden wordt gevraagd is de studie geneeskunde aan de Erasmus Universiteit Rotterdam. In deze selectieprocedure wordt kandidaten gevraagd aan te geven op welke vaardigheden zij zich onderscheiden.

Op het deelnameformulier geef je informatie over je vooropleiding (inclusief extra schoolprestaties) en je motivatie voor de opleiding Geneeskunde in Rotterdam. Daarnaast kun je aangeven welke onderscheidende vaardigheden jij de afgelopen tweeënehalf jaar naast je voltijdschool of -opleiding hebt ontwikkeld. Het gaat hierbij om onderscheidende vaardigheden die je middels aantoonbare activiteiten, waaruit jouw ambitie en volharding blijken, hebt ontwikkeld. Deze vaardigheden kunnen je helpen bij je studie Geneeskunde in Rotterdam of het latere werk als arts.

Voorbeelden van geschikte activiteiten zijn:

- Werkervaring in de gezondheidszorg (maar ook andere sectoren kunnen relevant zijn);
- Inzet voor humanitaire hulpverlening of vrijwilligerswerk;
- Voorzitter of penningmeester van het leerlingenbestuur, (hoofd-)redacteur van de schoolkrant;
- Bijzondere prestatie op het gebied van wetenschap, kunst, literatuur of topsport. Topsportactiviteiten worden beoordeeld door Erasmus Topsport. Zie de website voor meer informatie, de voorwaarden en de vereiste bewijsstukken (NB: je dient als kandidaat zelf voor het vereiste bewijsmateriaal te zorgen).

Het ontwikkelen van vaardigheden kost tijd. We zijn dus op zoek naar activiteiten waarmee een substantiële tijdsinvestering gemoeid ging en waarbij je zelf actief aan de slag bent gegaan. Voor een aantal dagen meelopen in de verpleging/verzorging of meekijken met een huisarts of specialist in het ziekenhuis zal je een minimale score krijgen.

De activiteit moet ook onderscheidend zijn. Voorbeelden van niet onderscheidende activiteiten zijn oppassen of het behalen van een EHBO-diploma, zelf sporten niet behorend tot de categorie topsport en hobbymatig/recreatief muziek maken.

Je mag maximaal drie activiteiten vermelden op je deelnameformulier. Deze activiteiten moet je voorzien van bewijsmateriaal. Zonder bewijsstukken kunnen de activiteiten niet beoordeeld worden en tellen ze dus niet mee. Ook informatie en/of bewijsstukken die worden nagestuurd of te laat binnen komen, worden niet meer meegenomen in de beoordeling. In de handleiding die je ontvangt bij je deelnameformulier staat precies beschreven welke bewijsstukken wij van je nodig hebben.

In dit laatste voorbeeld is duidelijk dat er hoge eisen gesteld worden aan de activiteiten die de kandidaten noemen. Het doel van het instrument is hier in kaart te brengen op basis van welke activiteiten kandidaten zich onderscheiden van andere kandidaten. Het meegeven van veel informatie en structuur in de opdracht maakt het beoordelen in de volgende fase ook duidelijker. Een belangrijk onderdeel van de opdracht is dat er gevraagd wordt om bewijsmateriaal te leveren voor de activiteiten die genoemd worden. Ten slotte wordt er een eis van 'langdurigheid' opgelegd: kandidaten moeten een activiteit gedurende langere tijd uitvoeren willen ze meer punten kunnen krijgen voor die activiteit. De redenering daarachter is dat dergelijke activiteiten authentieker zijn en een flinke tijdsinvestering vragen (Stegers-Jager, 2018).

## Belangrijke afwegingen

De hoeveelheid structuur die je meegeeft aan kandidaten is ook hier een belangrijke afweging. Een goed gestructureerde uitvraag van het CV, waarbij het voor kandidaten duidelijk is wat ze moeten aanleveren en op welke manier, zorgt ervoor dat vergelijkbare informatie wordt aangeleverd. Bovendien zal de beoordeling ook sneller lukken als de uitvraag en de beoordeling goed gestructureerd zijn en heldere criteria beschikbaar zijn. Een ongestructureerd CV levert veel verschillende antwoorden op die lastig te vergelijken zijn. Dit kost meer tijd en zal moeilijker te interpreteren zijn. Bovendien is de kans groter dat de beoordeling afhangt van welke beoordelaar het CV bekijkt.

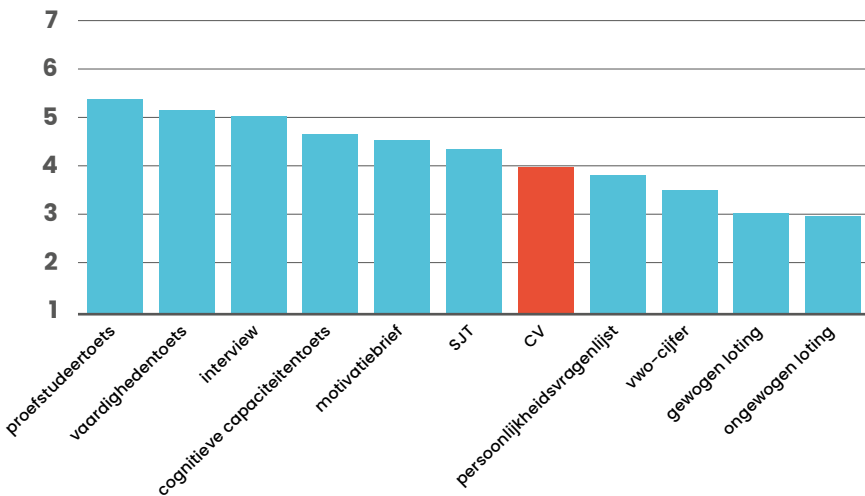
Een andere belangrijke afweging die hoort bij het meenemen van eerdere ervaring is dat het effect kan hebben dat kandidaten hierop gaan anticiperen: ze weten dat er in bepaalde procedures naar gevraagd wordt en gaan om deze reden op zoek naar extracurriculaire activiteiten (Stegers-Jager, 2018). Dat maakt het moeilijk om onderscheid te maken tussen kandidaten die deze

activiteiten ondernemen vanuit eigen motivatie en kandidaten die dit doen om meer kans te maken in de selectie. Het doen van extracurriculaire activiteiten zal in veel gevallen ook in dat laatste geval nog steeds bijdragen aan de ontwikkeling van de kandidaat, maar je meet net wat anders. Daarnaast kan het meenemen van extracurriculaire activiteiten zorgen voor extra druk op kandidaten. Extracurriculaire activiteiten hebben in veel gevallen een positief effect op latere prestaties binnen en buiten de opleiding (Urlings-Strop, Themmen & Stegers-Jager, 2017; Mahoney, Cairns & Farmer, 2003).

Een laatste belangrijke afweging is dat het beoordelen en meewegen van een CV fraudegevoelig is. Kandidaten kunnen allerlei activiteiten in hun CV opnemen waarvan niet duidelijk te beoordelen is of deze inderdaad zijn uitgevoerd. Dit is de reden dat veel opleidingen bij het CV vragen om bewijs, bijvoorbeeld in de vorm van referenten (Stegers-Jager, 2018; Urlings-Strop et al., 2017). Het daadwerkelijk natrekken van referenten vraagt wel extra tijd.

**Wat vinden kandidaten?**

Uit het onderzoek van Suzanne Fikrat-Wevers en collega's (2023) blijkt dat het meenemen van een CV niet heel positief gewaardeerd wordt door kandidaten. Kandidaten ervaren bij een CV veel ongelijkheid op basis van het hebben van een sociaal netwerk om relevante ervaring op te doen. Ook vragen kandidaten zich af of het wel realistisch is om van middelbare scholieren te verwachten dat zij al een CV hebben opgebouwd. In het internationale onderzoek van Patterson en Zibarras (2018) wordt juist genoemd dat de acceptatie door kandidaten hoog is. Dat lijkt in de Nederlandse context anders te zijn.



Figuur 16 • Wat vinden kandidaten

## Wat weten we uit onderzoek?

Er is veel onderzoek gedaan naar CV's binnen de arbeids- en organisatiepsychologie. Hieruit blijkt dat er veel informatie uit CV's te halen is, maar dat de betrouwbaarheid en voorspellende waarde ter discussie staan. Een verklaring kan zijn dat tot 50% van de sollicitanten overdrijft op hun CV (George & Marett, 2004). In de context van selectie van studenten wordt een CV vooral gebruikt om kandidaten te vragen naar hun extracurriculaire verdiensten. Onderzoek in de Nederlandse context laat zien dat het CV informatie kan geven over affectieve vaardigheden en dat het voorspellend kan zijn voor resultaten zoals uitval en klinische prestaties in de opleiding (Stegers-Jager, 2018).

## Potentiële bias

Het meenemen van extracurriculaire activiteiten kan, afhankelijk van hoe het wordt ingezet, een voordeel opleveren voor bepaalde groepen kandidaten. Als het doel is om na te gaan of kandidaten extracurriculaire activiteiten hebben gedaan die sterk gerelateerd zijn aan de opleiding, dan is de kans groot dat de toegang hiertoe niet voor alle kandidaten gelijk is. Kandidaten met bijvoorbeeld ouders die in de zorg werken, hebben meer kans hier vrijwilligerswerk of een baantje in te vinden. Ook is het voor kandidaten die vanwege de thuissituatie een betaalde bijbaan moeten hebben of mantelzorg moeten verlenen moeilijker om dat te combineren met vrijwilligerswerk. Door in de selectieprocedure de nadruk te leggen op bepaalde vaardigheden, zoals communicatie of samenwerking, in plaats van op ervaring in een specifieke sector, kan dit effect afnemen. Communicatievaardigheden ontwikkel je ook in een bijbaan achter de kassa, leiderschap kan je aantonen als je scoutingbegeleider bent geweest en planningsvaardigheden zijn belangrijk als je een grotere mantelzorgtaak hebt naast school.

De onderzoeksresultaten met het oog op bias lopen rondom het CV uit elkaar. Uit recent onderzoek (Fikrat-Wevers et al., 2023) komt naar voren dat mannelijke kandidaten en eerstegeneratiemigranten worden benadeeld door het gebruik van een CV. Ook kandidaten met een buitenlandse vooropleiding lijken nadeel te ondervinden bij het gebruik van dit instrument. Uit eerder onderzoek (Stegers-Jager, Themmen, Cohen-Schotanus & Steyerberg, 2015) komt dit niet naar voren. Het lijkt er dan ook op dat het vooral belangrijk is om te kijken naar hoe je het instrument vormgeeft en implementeert, en deze bias niet inherent bij het instrument hoort.

Aan de andere kant kan het meenemen van informatie uit een CV als achtergrond bij bijvoorbeeld vooropleidingscijfers er mogelijk voor zorgen dat kandidaten die thuis bijvoorbeeld een grote zorgtaak hebben of die vanwege



financiële redenen veel naast hun vooropleiding hebben moeten werken een meer vergelijkbare kans op toelating krijgen (Stegers-Jager et al., 2015). Hierbij is de procedure voor het meenemen van deze resultaten erg belangrijk, evenals de juridische kaders (bijvoorbeeld de Wet gelijke behandeling).



# 14 Aanbevelingsbrieven

## Wat is het?

In dit handboek wordt een brede definitie van het concept aanbevelingsbrieven gehanteerd. Het gaat hierbij om instrumenten waarbij middels referenten informatie wordt verzameld over de kwaliteiten van een kandidaat. Dat kan in de vorm van een brief zijn, maar ook in de vorm van een formulier of een meer kwantitatieve beoordeling. Aanbevelingsbrieven worden wereldwijd door iets minder dan de helft van de opleidingen aan topuniversiteiten gebruikt in selectieprocedures voor bacheloropleidingen (Steenman, 2018, pp. 77-78); voor selectie voor de master is het internationaal een zeer veel gebruikt instrument.

Opleidingen die aanbevelingen vragen doen dat op verschillende manieren. Zo treden er bijvoorbeeld verschillen op in:

- Van wie een aanbeveling wordt gevraagd;
- Welke informatie gevraagd wordt in de aanbeveling;
- In welke vorm de informatie gevraagd wordt (met name hoe gestructureerd).

De keuzes die bij de specifieke manier van inzetten van het instrument gemaakt worden, hebben invloed op de betrouwbaarheid en voorspellende waarde. Zoals we wel vaker concluderen: hoe beter je criteria zijn afgestemd op de inhoud en vorm van de opleiding en hoe beter hetgeen waarnaar gevraagd wordt in een aanbevelingsbrief daarbij past, hoe meer voorspellende waarde het instrument zal hebben voor de latere prestaties van kandidaten.

Via aanbevelingsbrieven kan informatie verzameld worden over vaardigheden van kandidaten die via de meeste andere mogelijke selectie-instrumenten moeilijk te verkrijgen is (Kuryshева & Krijgsheld, 2021). Het biedt de mogelijkheid een ander perspectief mee te nemen dan dat van de kandidaat

zelf en kan daarmee een goede toevoeging zijn aan een selectieprocedure. Dat is vooral het geval wanneer de referent de kandidaat goed kent in een relevante context.

Om een idee te geven van de grote verscheidenheid in hoe aanbevelingsbrieven ingezet kunnen worden, volgen hieronder een aantal voorbeelden.

### Voorbeelden

In de selectieprocedure voor de premaster geneeskunde aan de Universiteit Leiden wordt een aanbevelingsbrief gebruikt. Hieraan wordt de eis gesteld dat de aanbevelingsbrief geschreven wordt door iemand vanuit de bacheloropleiding die een kandidaat op dat moment volgt, of heeft gevolgd:

Een aanbevelingsbrief geeft ons inzicht in hoe anderen over jou denken. De aanbevelingsbrief laat je schrijven door iemand die jou kent vanuit je opleiding; die inzicht heeft en kan beoordelen wat jouw sterke kanten en jouw academische kwaliteiten zijn. De aanbeveling is opgesteld op briefpapier van de instelling en voorzien van een handtekening. De persoon die de aanbevelingsbrief schrijft mag geen vriendschappelijke, familie- of zakelijke banden met je hebben. Je stuurt één aanbevelingsbrief mee.

Voor de selectie voor het kleinschalig, internationaal bachelorprogramma in geneesmiddelenonderzoek dat binnen de bachelor Farmacie van de Universiteit Utrecht cursussen op honoursniveau aanbiedt, wordt ook gebruikgemaakt van aanbevelingsbrieven. Bij de vraag staat het volgende:

*“We kindly request that you ask somebody that knows you well in the school setting, preferably a (science) teacher or perhaps your mentor to fill in this form. If you are already studying, you may ask a former or current teacher/mentor. Please note: only one recommendation is needed and will be read”.*

Voor deze aanbevelingen wordt een specifiek format gesuggereerd:

#### **Recommendation form / Letter of recommendation**

By answering the following questions, you will greatly assist in the selection process. Also a ‘regular’ letter of recommendation that includes the information requested is very much appreciated.

Upon completion, the recommendation form can be returned to the applicant. However, if you wish to keep the recommendation confidential, you can mail it directly to [selectie.pharm@uu.nl](mailto:selectie.pharm@uu.nl).

Note to applicant: In the latter case, please upload a letter stating that the recommendation has been sent to this email address in Osiris Online Application.



Note to Dutch-speaking references: If you think you may represent your student better in Dutch, please answer the questions in Dutch instead of English.

### Reference's information

Name, Telephone Number, Email Address, Student's name, Relationship to the student, Date.

1. How well did the prospective student do in his/her studies?
2. Were there any circumstances that affected the student's performance?
3. What added value does this prospective student have to offer CPS and vice versa?
4. Can you briefly evaluate the student's:
  - a. Analytical capabilities? Is he/she capable of analysing and interpreting things?
  - b. Problem-solving abilities? Is he/she creative; does he/she take initiative; is he/she capable of making decisions?
  - c. Relational skills? Can he/she see things from another person's perspective; does he/she work well with others?
  - d. Self-directed learning capabilities? Is he/she adaptable and somewhat flexible; does he/she handle stress well; does he/she take responsibility for his/her behaviour; is he/she capable of critically evaluating his/her own behaviour; is he/she open to being evaluated by others?
5. Do you have other comments?

*Signature of reference*

De Radboud Universiteit Nijmegen vraagt kandidaten zelf een reflectie over de opleiding en het beroepsperspectief te schrijven, waarbij ook drie bekenden van de kandidaat mee moeten kijken. De referenten wordt gevraagd te benoemen waarom de kandidaat geschikt is voor de opleiding, maar ook minimaal één punt te noemen waarom de kandidaat minder geschikt zou zijn. Hierop wordt de kandidaat dan weer gevraagd te reflecteren:

Als huiswerkopdracht word je uitgedaagd om op jouw studiekeuze en het bijbehorende beroepsperspectief te reflecteren. Om te beginnen geef je in eigen bewoordingen een korte beschrijving van de opleiding Geneeskunde aan het RadboudUMC. Deze beschrijving leg je vervolgens voor aan drie personen die jou goed kennen. Je vraagt deze referenten om in een korte tekst te beschrijven waarom jij geschikt bent voor de opleiding Geneeskunde in Nijmegen. De referenten moeten ook minimaal één punt noemen waarom je minder geschikt zou zijn voor deze opleiding. Als je de drie referenties verzameld hebt, schrijf je in je eigen bewoordingen hierover een reflectie. Je gaat in op de mening van de referenten waaruit blijkt dat je geschikt bent voor deze opleiding en je gaat in op de mening van je referenten waaruit

blijkt dat je minder geschikt zou zijn. Je gaat dus in op je eigen geschiktheid voor de opleiding aan de hand van je eigen beschrijving en de opmerkingen van de referenten. Uiteindelijk lever je de volgende onderdelen in:

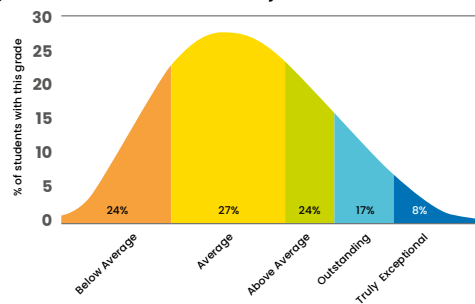
1. Jouw beschrijving van de studie Geneeskunde in Nijmegen, in het Nederlands in maximaal 500 woorden. Raadpleeg hiervoor het voorlichtingsmateriaal dat je vindt op de website van het RadboudUMC;
2. De drie referenties, elk van maximaal 500 woorden;
3. Jouw reflectie op de referenties, in het Nederlands in maximaal 500 woorden in totaal.

Je geeft blijk van voldoende motivatie voor deze opleiding door je huiswerkopdracht zodanig te maken dat je hier een voldoende voor behaalt. Daarnaast toon je je motivatie en aanleg voor de studie ook aan door je goed voor te bereiden op de twee selectietoetsen.

Bij de Graduate School of Life Sciences van de Universiteit Utrecht wordt gebruikgemaakt van een gestandaardiseerd scoreformulier voor de aanbeveling, waarbij de referent wordt gevraagd om de kandidaat in vergelijking met anderen te beoordelen op verschillende criteria. Een deel van dat formulier staat hieronder:

### Analytical and Research Skills

In this section, please indicate how well the applicant compares to other students at your institution for analytical skills:



1. The applicant has the skills to find, use and interpret specialist literature  
 Cannot assess    Below Average    Average    Above Average    Outstanding    Truly Exceptional
2. The applicant has skills in life sciences research methods  
 Cannot assess    Below Average    Average    Above Average    Outstanding    Truly Exceptional
3. The applicant has performed quantitative research methods and has data analysis skills  
 Cannot assess    Below Average    Average    Above Average    Outstanding    Truly Exceptional
4. The applicant is skilled in statistical methods  
 Cannot assess    Below Average    Average    Above Average    Outstanding    Truly Exceptional

## Belangrijke afwegingen

Veel van de belangrijke afwegingen zijn ook te zien in de keuzes die bij de verschillende voorbeelden gemaakt zijn, en hebben we aan het begin van dit hoofdstuk ook al kort genoemd. Als eerste is het belangrijk om helder te hebben van wie je een aanbeveling vraagt. Dat hangt sterk samen met de informatie die je kunt ophalen. Over het algemeen is het goed om af te bakenen van wie de informatie in de aanbeveling komt, vooral omdat de basis waarop de aanbeveling in dat geval gebaseerd is meer te vergelijken is. Je kunt bijvoorbeeld aangeven dat de referent een docent moet zijn die de kandidaat in een onderwijssetting moet hebben meegemaakt, met daarnaast bijvoorbeeld een afbakening van hoe recent dat moet zijn en hoe lang de periode waarin die ervaring is opgedaan moet zijn geweest. In dat geval kun je bijvoorbeeld informatie vragen over de houding en prestaties in die onderwijssetting. Bij het bepalen wie de aanbeveling schrijft hoort ook de vraag hoe je kunt achterhalen dat de aanbeveling inderdaad door diegene is geschreven, en daarmee hoe je fraude voorkomt.

De informatie die je vraagt en die je via dit instrument wilt verkrijgen is de tweede belangrijke afweging. Dat hangt sterk samen met het vorige punt; het is natuurlijk onverstandig om informatie te vragen waarvan het onzeker is of de respondenten die ook hebben. Daarnaast vormt wat relevant is voor de opleiding natuurlijk de basis voor deze afweging. Als samenwerken binnen de opleiding een belangrijke rol speelt, dan zou je naar relevante vaardigheden op dat vlak kunnen vragen. Als hogere orde cognitieve vaardigheden van belang zijn, zou je de referenten kunnen vragen de kandidaten hierop te beoordelen (bijvoorbeeld door iets te zeggen over hun kritisch denken of analytische vaardigheden). Ook of kandidaten bijvoorbeeld gedisciplineerd aan hun opleiding werken, wat iets lijkt te zeggen over de prestaties van kandidaten in kennis- en begriptoetsen binnen een bacheloropleiding, is een mogelijke vraag voor referenten. Het belangrijkste is dus dat je de referenten vraagt criteria te beoordelen die belangrijk zijn – maar waar ze ook inzicht in kunnen hebben vanuit hun rol ten opzichte van de kandidaat.

De vorm waarin de aanbeveling wordt gevraagd is eveneens belangrijk. Zoals al bij verschillende instrumenten naar voren is gekomen, lijkt ook hier te gelden dat een meer gestructureerde aanbeveling meer vergelijkbare informatie oplevert, en daarmee ook meer relevante informatie voor de selectiebeslissing. Een open aanbevelingsbrief kan, wanneer de manier van beoordelen wel goed gestructureerd is, een toevoeging aan de voorspellende waarde van de selectieprocedure leveren, voor vaardigheden die op andere manieren vaak lastig te voorspellen zijn (Kuncel, Kochevar & Ones, 2014), maar de beste resultaten op het gebied van voorspellende waarde voor studieresultaten in de Nederlandse context worden behaald door de referent te vragen om kandidaten voor bepaalde vaardigheden op een schaal te

beoordelen ten opzichte van een referentiegroep (bijvoorbeeld alle 6 vwo- of 5 havo-leerlingen). Een tussenvorm is om referenten specifieke vragen te stellen, die ze ofwel in een brief, ofwel in een formulier met losse velden kunnen beantwoorden.

Een andere belangrijke afweging is de hoeveelheid werk die je van referenten vraagt. Hierbij spelen twee elementen een rol. Allereerst de verhouding tussen de hoeveelheid werk die je van referenten vraagt en het gebruik van de informatie in het selectieproces. Wel om een aanbeveling vragen, maar deze uiteindelijk niet daadwerkelijk gebruiken in de procedure, lijkt daarmee vanuit dit oogpunt snel af te vallen. Een andere overweging is of je het referenten gemakkelijker kunt maken, bijvoorbeeld door de informatie voor te structureren, door aan te geven dat reacties in steekwoorden ook voldoende zijn, of door een schaal te gebruiken in plaats van tekst.

Een laatste belangrijke afweging is of kandidaten de aanbeveling zelf te zien krijgen. Het blijkt daarbij vooral belangrijk om dat voor alle kandidaten op dezelfde manier te doen, en daarover helder te communiceren. Onderzoek van Ceci en Peters (1984) laat zien dat bij aanbevelingen waarbij de referent een score geeft voor specifieke vaardigheden weliswaar positiever zijn wanneer kandidaten de aanbeveling te zien krijgen, maar dat de verschillen tussen de beoordelingen niet afnemen. De hele schaal schuift dus naar boven op.

### Wat vinden kandidaten?

In het onderzoek van Suzanne Fikrat-Wevers en collega's (2023) zijn aanbevelingsbrieven niet meegenomen, dus er is geen directe vergelijking met de andere selectie-instrumenten die in de Nederlandse context gebruikt worden. Volgens meerdere andere onderzoeken (bv. De Boer & Van Rijnsoever, 2021; Kuryshva & Krijgsheld, 2021; Patterson & Zibarras, 2018, p. 5), staat de meerderheid van de kandidaten positief tegenover het gebruik van aanbevelingsbrieven. Ook wanneer zij in de Nederlandse context gevraagd worden om zelf een selectieprocedure in te richten worden aanbevelingsbrieven regelmatig voorgesteld.

### Wat weten we uit onderzoek?

Literatuur over de voorspellende waarde van aanbevelingsbrieven levert een wisselend beeld op. Er zijn veel onderzoeken die geen voorspellende waarde vinden van ongestructureerde aanbevelingsbrieven (bv. Patterson & Zibarras, 2018; Creech, Cooper, Aplin-Kalisz, Maynard & Baker, 2018; Stedman, Hatch & Schoenfeld, 2009). Wel komt regelmatig naar voren dat aanbevelingen in deze vorm informatie opleveren om uitval te kunnen voorspellen (Van Vertloo, Burzette & Danielson, 2022).



Naarmate er meer gestructureerd informatie over kandidaten wordt verzameld, blijken aanbevelingsbrieven meer voorspellend voor de prestaties van een kandidaat binnen de opleiding. Meer gestructureerde aanbevelingsbrieven zijn efficiënter en makkelijker te beoordelen, de structuur verlaagt de kans dat de beoordelaar wordt beïnvloed door de schrijfstijl of het taalgebruik van de referent, en er kan specifiek gevraagd worden naar de informatie die je over de kandidaat te weten wilt komen (Kuryшева & Krijgsheld, 2021). Ook het vragen naar specifieke voorbeelden (Knouse, 1983), het trainen van beoordelaars aan de hand van duidelijke criteria en het goed nadenken over waarvoor en hoe je de aanbevelingsbrieven wilt gebruiken binnen je procedure zorgen voor een betere betrouwbaarheid en validiteit van dit instrument.

Wanneer referenten kandidaten scoren op een schaal ten opzichte van een referentiegroep kan de voorspellende waarde van aanbevelingen zelfs redelijk sterk zijn voor toekomstig presteren (McCarthy & Goffin, 2001). Dit onderzoek werd gedaan in Canada, maar deze resultaten blijken ook in de context van opleidingen in Nederland op te gaan (zowel voor Nederlandse als Engelstalige opleidingen). Die voorspellende waarde bestaat ook in aanvulling op andere instrumenten.

### Potentiële bias

De potentiële bias in aanbevelingen zit zowel in het verkrijgen ervan, in de inhoud ervan en in de beoordeling. Zoals bij meer selectie-instrumenten helpt het structureren van de aanbeveling en het aanbevelingsproces om bias te verminderen. Toegang tot mensen die een aanbeveling willen schrijven is niet altijd gelijk verdeeld, zeker niet wanneer het gaat om referenten met een ‘indrukwekkende’ achtergrond. De afbakening van de groep referenten kan helpen om de bias in deze situatie te verkleinen. Het geven van concrete handvatten voor hoe kandidaten een potentiële referent kunnen benaderen kan daarbij ook helpen.

Bij tekst in de aanbeveling blijkt dat het uitmaakt hoe goed de kandidaat en de referent elkaar kennen. Er is ook onderzoek dat laat zien dat mannen voor vrouwelijke kandidaten positievere brieven schrijven dan voor mannelijke kandidaten (Houser & Lemmons, 2018; Schmader, Whitehead & Wysocki, 2007). Ook de uiteindelijke beoordeling is potentieel gevoelig voor bias, vooral als die meer ruimte biedt voor de beoordelaar. Als er veel ruimte is voor de beoordelaar, gaan de eigen normen van de beoordelaars een grotere rol spelen (Kuryшева, Van Rijen & Dilaver, 2019). Bij het toekennen van scores door de referent lijken er over het algemeen geen structurele verschillen te zijn tussen scores voor kandidaten met bepaalde achtergronden, hoewel daarvoor meer onderzoek nog nodig is om dat met meer zekerheid te kunnen stellen.



# 15 Situational Judgement Tests

## Wat is het?

Een Situational Judgement Test (SJT) of situationele beoordelingstest is een instrument waarbij kandidaten hypothetische situaties voorgelegd krijgen en daarop moeten reageren. Het doel van een SJT is om in te schatten in hoeverre kandidaten verschillende (meestal affectieve) vaardigheden kunnen inzetten, zoals samenwerken, openmindedness of omgaan met onzekerheid. SJT's hebben een betere validiteit dan andere selectie-instrumenten die op grote schaal kunnen worden afgenomen voor affectieve vaardigheden en de inhoud van de SJT's kan worden gekoppeld aan vorm en inhoud van de opleiding (Cleland et al., 2023). Dit maakt de SJT's een goede optie voor verschillende vaardigheden.

De situaties in de test worden meestal voorgelegd in de vorm van een video of geschreven tekst. De manier waarop antwoord wordt gegeven kan verschillen, en ook de vaardigheden die worden gemeten lopen uiteen. Zo kunnen open vragen gebruikt worden waarbij ook argumentatie beoordeeld kan worden, meerkeuzeopties om begrip te testen, ordeningsvragen om de gepastheid van een antwoord te bevragen, et cetera. SJT's worden meestal voor een specifieke context ontwikkeld en afgenomen, met specifieke praktijksituaties waarop de kandidaat moet reageren. Hierdoor zijn er weinig algemene SJT's beschikbaar die in meerdere contexten zijn afgenomen.

Het voordeel is dat een SJT hiermee ook een informatieve functie kan hebben naar de kandidaat. De situaties die worden voorgelegd lijken op wat je tijdens de opleiding of in het latere beroep zou kunnen tegenkomen. Het aanpassen aan de specifieke context maakt ook dat er veel variaties zijn in context, type vragen en antwoordmogelijkheden. In dit hoofdstuk staan we stil bij de belangrijkste keuzes, maar zie voor een uitgebreider overzicht bijvoorbeeld het proefschrift van De Leng (2019).

Zowel binnen als buiten Nederland zijn voorbeelden te vinden van selectieprocedures waar SJT's worden gebruikt voor het meten van bepaalde affectieve vaardigheden. Hieronder worden een aantal voorbeelden beschreven.

### Voorbeelden

In de selectieprocedure voor de opleiding geneeskunde aan de Erasmus Universiteit Rotterdam werd een SJT gedaan om de integriteit van deelnemers te meten (dit ging om een pilot om te testen of dit een goede aanpak zou zijn, de test is nooit daadwerkelijk ingezet voor selectie). De scenario's waren korte beschrijvingen van een situatie, waarna kandidaten vier gegeven reacties moesten scoren op gepastheid op een schaal van 1-6.

#### Scenario

Rafiq komt er achter dat Marieke in het bezit is van het tentamen van volgende week. Marieke vertelt dat ze het tentamen al aan een aantal medestudenten heeft verkocht en vraagt aan Rafiq of hij het tentamen ook wil inzien.

Beoordeel voor elk van de onderstaande reacties hoe gepast ze zouden zijn voor Rafiq (omcirkel s.v.p.)

		<b>Zeer ongepast</b>				<b>Zeer gepast</b>
1.	Het tentamen inzien en het tegen niemand zeggen.	A	B	C	D	E F
2.	Het tentamen niet inzien en bij de docent melden dat het tentamen is uitgelekt zonder namen te noemen.	A	B	C	D	E F
3.	Het tentamen inzien omdat iedereen dat zou doen.	A	B	C	D	E F
4.	Het tentamen niet inzien omdat hij daar geen recht op heeft.	A	B	C	D	E F

De geneeskundeopleiding aan de Universiteit Maastricht zet een video-gebaseerde SJT in als onderdeel van de selectieprocedure. In deze toets worden meerdere vaardigheden gemeten: samenwerking, medisch en maatschappelijk bewustzijn, ethisch besef, empathie en reflectie.



Het doel van deze SJT opdracht is het peilen naar de aanleg voor de competenties Samenwerken en Communicatie.

Tijdens de bachelor en master Geneeskunde in Maastricht werk je regelmatig samen aan een opdracht en krijgt aan het einde iedereen hetzelfde cijfer. Verder moet je tekenen voor ieders bijdrage aan de opdracht. Bekijk het videofragment “tekstbericht” → ten behoeve van het handboek is het scenario zo goed mogelijk beschreven.

*Het is avond. Student zit achter laptop met aantekeningen en boek bij de hand, is druk aan het werk. Ontvangt dan het volgende tekstbericht van een groepsgenoot (collega student)*

*“Kan morgen niet, doe jij de presentatie? Jij zit er toch beter in. Het is een groepsbeoordeling. Voor jou ook beter! J.” Diepe zucht van student.*

Stel je bent de student die dit bericht ontvangt en je werkt samen met de afzender van dit bericht aan een presentatie. De afzender zou morgen de presentatie verzorgen.

Vraag 1: Hoe ga je reageren en waarom?

Onderstaande vraag uit de SJT die eerder gebruikt is voor de bacheloropleiding Zorg, Gezondheid en Samenleving aan de Universiteit Utrecht is in de volgende selectieronde verwijderd vanwege de grote invloed van kennis over de context van de gezondheidszorg op de potentiële beantwoording.

Je bent een arts in opleiding in een ziekenhuis. Jouw patiënt Meneer Hafeez is enige tijd geleden behandeld voor prostaatkanker en is weer opgenomen in het ziekenhuis voor controle. Vanmorgen heb je de uitslagen van de controle binnengekregen en het lijkt erop dat de kanker vermoedelijk terug is. Je wacht nog op de uitslag van de laatste test. Als je bezig bent met een andere patiënt op de zaal, buigt de heer Hafeez zich naar je toe en zegt: “Vertel eens eerlijk, is mijn kanker terug?”

Rangschik de volgende reacties in volgorde van geschiktheid van meest naar minst geschikt.

- 1 *Leg uit aan meneer Hafeez dat het waarschijnlijk is dat de kanker weer terug is*
- 2 *Informeer meneer Hafeez dat je vaart achter de testuitslagen zal zetten en vraag een van je senior collega's om met meneer Hafeez te bespreken*
- 3 *Nodig meneer Hafeez uit om samen met een senior verpleegkundige in een rustige kamer te gaan zitten en onderzoek wat zijn angsten zijn*

*4 Leg uit aan meneer Hafeez dat je nog niet alle testuitslagen binnen hebt, maar dat je meteen met hem komt praten als je de resultaten hebt*

*5 Stel meneer Hafeez gerust dat het goed zal komen*

In België (Vlaanderen) werd een SJT gebruikt voor een centraal georganiseerde selectie voor de geneeskundeopleidingen. De SJT was gericht op het meten van twee vaardigheden: het bouwen en in stand houden van relaties, en het uitwisselen van informatie. De SJT omvatte bijvoorbeeld situaties rond het tonen van aandacht en interesse, het overbrengen van slecht nieuws aan een patiënt, het reageren op de weigering van een patiënt om het voorgeschreven medicijn in te nemen, en het gebruik van de juiste taal om technische termen uit te leggen. Kandidaten werd uit vier antwoordopties gevraagd de meest effectieve reactie aan te kruisen.

## Belangrijke afwegingen

SJT's die worden ingezet in selectieprocedures worden meestal zelf ontwikkeld en afgestemd op de opleiding. Hieronder bespreken we een aantal afwegingen waarmee je te maken krijgt als je een SJT gaat ontwikkelen. Voor meer informatie over het ontwikkelen van een goede SJT verwijzen we ook naar hoofdstuk 4 van het handboek van Patterson en Zibarras (2018) en de AMEE Guide van Patterson, Zibarras en Ashworth (2016).

Een eerste belangrijke afweging is hoe de scenario's aan kandidaten worden voorgelegd. Meestal gebeurt dat bij SJT's in de vorm van korte video's of geschreven casussen. Onderzoek toont aan dat het gebruik van video's meestal een beter beeld geeft van de kandidaten en een hogere voorspellende waarde heeft. Bovendien ervaren deelnemers videofragmenten vaak positiever dan geschreven tekst (Patterson, Knight et al., 2016). Bij het gebruiken van video's wordt het echter wel noodzakelijk de toets op locatie af te nemen; de combinatie van een toetsprogramma met een proctoring-programma dat tegelijkertijd video's moet uploaden en downloaden, overstijgt de bandbreedte van de internetverbinding in veel huishoudens.

Het ontwikkelen van een goede SJT kost relatief veel tijd. Een video-gebaseerde SJT zal meer tijd kosten om te ontwikkelen dan een geschreven SJT. Wel zijn zeker geschreven SJT's qua uitvoering efficiënt: ze kunnen tegelijkertijd (online) worden afgenomen bij grote groepen kandidaten en automatisch worden beoordeeld (in het geval van gesloten vragen). Vooral voor opleidingen met veel aanmeldingen en relatief weinig plekken is een SJT daarmee mogelijk de moeite waard (Patterson, Knight et al., 2016; Schreurs, Cleland, Muijtjens, oude Egbrink & Cleutjens, 2018).

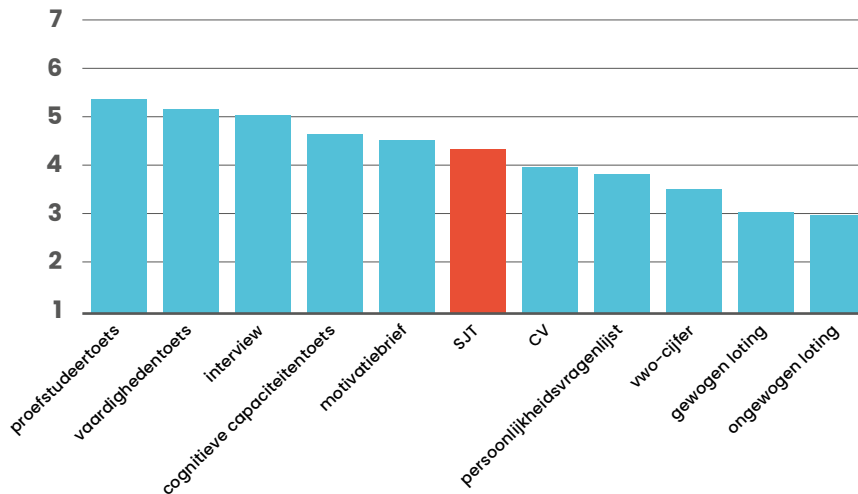
Een volgende belangrijke afweging is wat je wilt meten in een SJT. De vraag die je stelt naar aanleiding van het scenario bepaalt of je de kennis van kandidaten toetst, of juist het gedrag. Als je kandidaten vraagt naar de meest passende reactie in een voorgelegde situatie, test je of ze de kennis hebben over hoe je hoort te reageren. Je zou ook kunnen vragen naar hoe de kandidaten zelf zouden reageren. In dat geval schat je in welk gedrag zij in dergelijke situaties zouden vertonen. Vragen naar kennis lijkt minder goed gecoacht te kunnen worden (Patterson, Knight et al., 2016), maar vragen naar gedrag is meestal meer voorspellend voor hoe iemand zich in de opleiding gaat gedragen (Patterson & Zibarras, 2018). Vragen naar gedrag is wel meer vatbaar voor sociaal wenselijke beantwoording.

Zowel bij vragen naar kennis als naar gedrag zou je de vraag ook negatief kunnen stellen, dus hoe je in elk geval niet zou (moeten) reageren. Vragen naar wat je *niet* zou moeten doen kan een sterkere voorspeller zijn dan vragen wat je *wel* zou moeten doen (De Leng, Stegers-Jager, Born & Themmen, 2018). Andere mogelijkheden voor vragen die je stelt zouden het ordenen van de gepastheid van reacties kunnen zijn. Er kunnen ook open vragen worden overwogen. Dit vergt meer werk bij het beoordelen, maar kan ook meer informatie geven over de kandidaat. Hierbij is het van groot belang dat er van tevoren een uitgebreide en duidelijke antwoordsleutel opgesteld wordt om de betrouwbaarheid van de beoordeling te versterken.

De beoordeling van SJT's kan op verschillende manieren gebeuren en is een andere belangrijke afweging. Vaak wordt gebruikgemaakt van een expertgebaseerde score; een referentiegroep van experts uit het werkveld of docenten uit de opleiding beoordelen dan wat de juiste reactie zou zijn. Hierbij wordt ervan uitgegaan dat de experts de effectiviteit van bepaalde reacties in de praktijk kennen. Een nadeel kan zijn dat het soms ingewikkeld is om tot een consensus te komen, omdat verschillende groepen experts tot verschillende antwoorden kunnen komen. De empirische scoring gebruikt de scores van een grote groep als referentie. De voorwaarde hiervoor is dat je een voldoende grote groep hebt die de test doet, zodat deze representatief is voor een gemiddelde groep kandidaten. Onderzoek laat zien dat deze wijze van beoordelen in dat geval even betrouwbaar is als de expert-gebaseerde score (De Leng et al., 2017). Een derde manier die gebruikt wordt om SJT's te scoren is theoriegebaseerd. Hierbij wordt uitgegaan van een theoretisch kader dat in de antwoorden al dan niet wordt gespiegeld. Een nadeel van deze methode kan zijn dat het voor kandidaten makkelijker te doorzien is welk antwoord het 'juiste' is (Lievens & Schollaert, 2008). In de praktijk wordt ook vaak een combinatie van twee of drie van deze manieren van scoren gebruikt.

## Wat vinden kandidaten?

SJT's zijn meegenomen in het onderzoek van Fikrat-Wevers en collega's (2023), waarbij ze in de middenmoot scoren wat betreft hoe positief kandidaten over dit instrument zijn.



Figuur 17 • Wat vinden kandidaten

Hoewel kandidaten redelijk positief zijn over SJT's, zijn er vanuit kandidaten twijfels of de resultaten wel robuust genoeg zijn en of het niet te makkelijk is om je beter voor te doen (sociaal wenselijk te reageren). Hoewel onderzoek heeft laten zien dat deze zorgen niet gegrond zijn, leeft dit beeld wel onder kandidaten.

Internationale literatuur laat positieve geluiden van kandidaten over SJT's zien (Patterson, Knight et al., 2016). Kandidaten geven daar onder andere aan dat ze denken dat het gebruik van SJT's de genderdiversiteit in selectie zou kunnen verbeteren (Lievens, 2013).

## Wat weten we uit onderzoek?

Het voordeel van de mogelijkheden om SJT's zelf in te richten en de vele mogelijke combinaties die dat met zich meebrengt, is dat het instrument volledig aangepast kan worden aan het type vaardigheden en het niveau zoals dat verwacht wordt van kandidaten aan het begin van de opleiding. Het nadeel is dat uitkomsten uit onderzoek sterk afhankelijk zijn van de specifieke toets. In deze paragraaf worden meer algemene conclusies besproken, waarbij het goed is in gedachten te houden dat de precieze invulling en vorm van de SJT veel invloed hebben op de kwaliteitsmaten, zoals betrouwbaarheid, validiteit en voorspellende waarden.

Over het algemeen blijkt uit onderzoek dat SJT's een goede betrouwbaarheid hebben. Net als bij andere toetsen geldt wel dat hoe meer vragen je stelt om een bepaalde vaardigheid te meten, hoe betrouwbaarder het instrument wordt. Je kunt binnen een SJT verschillende vaardigheden meten, maar er moeten altijd voldoende vragen per vaardigheid in de test zitten.

Ook de validiteit van SJT's wordt in onderzoek positief beoordeeld (Patterson & Zibarras, 2018, p. 6). De SJT wordt gebruikt voor het meten van affectieve vaardigheden, maar er is overlap aangetoond met andere vaardigheden. Zo is er een samenhang met cognitieve vaardigheden van kandidaten. Dit geldt in het bijzonder als er in de SJT veel moet worden gelezen, begrepen en geïntegreerd. Ook is er overlap met uitkomsten van persoonlijkheidsvragenlijsten, met name met de kenmerken vriendelijkheid en zorgvuldigheid.

SJT's zijn in staat om latere, vooral affectieve, prestaties te voorspellen. Het ligt dus voor de hand dat een SJT vooral waardevol kan zijn in een selectieprocedure voor opleidingen waarin affectieve vaardigheden een grote rol spelen. SJT's voegen voorspellende waarde toe ten opzichte van andere instrumenten, zoals onder andere vooropleidingscijfers, gestructureerde interviews en kennis- en capaciteitstoetsen. Dit houdt in dat latere prestaties over het algemeen beter kunnen worden voorspeld door het toevoegen van een SJT. Evaluatie blijft belangrijk om na te gaan in hoeverre dit ook geldt binnen de specifieke context van de opleiding en verdere selectieprocedure.

Onderzoek bij geneeskunde aan de Universiteit Maastricht heeft aangetoond dat een SJT ook op de langere termijn prestaties kan voorspellen. Vaardigheden die tijdens de selectieprocedure voor de bachelor werden gemeten, voorspelden in toenemende mate prestaties op diezelfde competenties tijdens de master als studenten in de klinische praktijk werkten (Schreurs, Cleutjens, Cleland et al., 2020).

### Potentiële bias

Het lijkt erop dat er bij de SJT als instrument minder sprake is van een bias tegen bepaalde groepen kandidaten dan bij kennis- en capaciteitstoetsen (Patterson, Knight et al., 2016). Een gerelateerd voordeel is dat een SJT zich niet/nauwelijks laat coachen, waardoor het al dan niet kunnen (of willen) betalen van training ook geen effect kan hebben op de uiteindelijke score. Omdat SJT's zich vaak beter lenen voor vaardigheden waarover vrouwelijke kandidaten vaak op dat moment meer beschikken, komt dat verschil ook beperkt in onderzoek naar voren.

Wel hangt het af van de inhoud van de scenario's in hoeverre bepaalde groepen een voordeel kunnen hebben. Als een SJT bijvoorbeeld gebruikmaakt

van scenario's uit het medische werkveld, hebben kandidaten met ouders die zelf arts zijn een voordeel bij het beantwoorden van de vragen. Om dit te voorkomen is het aan te raden situaties te gebruiken die alle kandidaten herkennen uit hun eigen leven, zoals schoolsituaties. Dat doet wel af aan het matchingseffect van de test. Een andere optie kan zijn om te zorgen dat het niveau heel goed in lijn is met wat je van alle kandidaten kunt verwachten. Probeer voor jezelf (en met een diverse groep collega's) duidelijk na te gaan of er vragen zijn die makkelijker te beantwoorden zijn als je specifieke mensen in je sociaal netwerk hebt, en geef die informatie er vervolgens bij.

# 16 Interviews

## Wat is het?

Interviews zijn mondelinge gesprekken, op locatie of digitaal, tussen een kandidaat en een vertegenwoordiger van de opleiding (bijvoorbeeld een docent). In een interview wordt geprobeerd informatie te achterhalen over de kandidaat, hoe goed deze bij de opleiding past en over welke inhoudelijke kennis en vaardigheden deze beschikt. Verder leent een interview zich mogelijk ook voor het inschatten van bijvoorbeeld probleemoplossend vermogen en creativiteit (aspecten van HOCS) (Steenman, 2018). Interviews worden wereldwijd veel ingezet binnen selectieprocedures in het hoger onderwijs (Steenman, 2018).

Interviews kunnen in verschillende vormen worden afgenomen, waarbij de tijdsduur, opzet en inhoud van het gesprek sterk kunnen variëren. Zo zijn er ongestructureerde interviews waar de inhoud niet vastligt en varieert per kandidaat, maar ook zeer gestructureerde Multiple Mini Interviews (MMI) waarbij de kandidaat een circuit doorloopt van meerdere (minimaal 6 maar liever 8) verschillende korte stations met verschillende interviewers op elk station (Eva, Rosenfeld, Reiter & Norman, 2004).

Het gebruik van ongestructureerde interviews voor selectie levert weinig betrouwbare informatie op, en er is geen aanwijzing dat dat type interview enige voorspellende waarde heeft voor toekomstige prestaties van kandidaten. Ze kunnen wel werken voor de binding tussen opleiding en kandidaat, en voor bepaalde aspecten van matching. Meer gestructureerde vormen van interviews kunnen passen bij bepaalde vaardigheden die ook in de opleiding naar voren komen. Bijvoorbeeld in de situatie waarin het ontwikkelen of afnemen van schriftelijke toetsen minder voor de hand ligt of wanneer een combinatie tussen het sociale aspect van het instrument (binding met de opleiding tijdens het toelatingsproces) en inhoud belangrijk wordt gevonden.

Hieronder worden een aantal voorbeelden beschreven van manieren waarop interviews worden meegenomen in selectieprocedures.

### Voorbeelden

De bachelor medische hulpverlening aan de Hogeschool Rotterdam neemt een interview af, waarbij resultaten van een eerdere gemaakte toets als input worden gebruikt om het interview te structureren:

#### **Opdracht 1: Toets 'Critical Thinking'**

Voorafgaand aan de tweede ronde maak je thuis de digitale toets 'Critical Thinking'. De informatie voor het kunnen maken van deze toets krijg je na ronde 1 toegestuurd. De uitkomst van deze toets wordt gebruikt als input voor het motivatiegesprek.

#### **Opdracht 2: Motivatiegesprek**

Je voert een gesprek van 15 minuten met een docent van de opleiding op basis van de uitkomsten van de toets Critical Thinking. De gesprekken zullen online via Microsoft Teams plaatsvinden.

De scores van deze 2 opdrachten bij elkaar opgeteld vormen de totaalscore over ronde 2. Met de totaalscores over ronde 1 en 2 bij elkaar opgeteld wordt bepaald welke plaats je krijgt op de ranking.

In dit interview krijgen de kandidaten van tevoren te horen van welke vaardigheden zij voorbeelden moeten meenemen en hoe zij deze in het gesprek goed kunnen toelichten. Kandidaten krijgen in dit voorbeeld dus duidelijke instructies met informatie en praktische tips om zich goed voor te bereiden. De master Arts-Klinisch Onderzoeker (AKO) aan de Universiteit Maastricht is ook selectief, en zij gebruiken 'Multiple Mini Interviews' (MMI) in hun selectieprocedure :

Je score van Ronde 1 bepaalt of je wordt uitgenodigd voor Ronde 2.

Tijdens Ronde 2 heb je een aantal korte interviews met leden van de selectiecommissie. In deze gesprekken staan je motivatie, past performance, communicatieve en andere relevante vaardigheden centraal. De selectiegroep kijkt ook naar de realiseerbaarheid van de opleiding.

Deze selectieprocedure resulteert in een rangorde van kandidaten.





## Belangrijke afwegingen

Uit de voorbeelden hierboven komen in de verschillen al enkele belangrijke afwegingen naar voren. Bij het opzetten en uitvoeren van interviews zijn veel opties mogelijk (bv. Patterson, Knight et al., 2016; Patterson, Pre-scott-Clements et al., 2016; Patterson et al., 2018; Steenman, 2018). Uiteindelijk hangen die keuzes net als in voorgaande hoofdstukken steeds sterk samen met het doel van het instrument in de selectieprocedure als geheel.

De eerste belangrijke afweging is wie er wordt geïnterviewd. Daarin speelt naast de rol van het interview in het verzamelen van relevante informatie binnen de selectieprocedure als geheel, ook de capaciteit voor het afnemen van interviews mee. Bij een groot aantal aanmeldingen ten opzichte van de beschikbare plaatsen ligt het meer voor de hand om niet alle kandidaten te interviewen. Daarbij kan het een keuze zijn om alleen kandidaten te interviewen die in de eerste ronde een hoge score hebben behaald (bijvoorbeeld tot maximaal twee keer het aantal beschikbare plaatsen). Een andere optie is om alleen kandidaten te interviewen waarover twijfel bestaat. Bij die laatste optie is een afweging wel dat je daarmee die twijfel ook communiceert naar kandidaten.

Naast wie er worden geïnterviewd is een belangrijke afweging ook wie de interviews afneemt. Dat kunnen bijvoorbeeld docenten van de opleiding zijn, de leden van de selectiecommissie, mensen uit de praktijk, of mensen met specifieke expertise in het afnemen van interviews. Daarbij is ook de vraag of de interviews door één iemand worden afgenomen, of juist door meerdere beoordelaars. De betrouwbaarheid van interviews neemt toe bij meerdere beoordelaars; de invloed van subjectieve observaties neemt immers af. Aan de andere kant neemt de druk op kandidaten door het gesprek met meerdere mensen te voeren ook toe. Je kunt er ook voor kiezen om kandidaten langs verschillende interviewers te geleiden, waarmee ze steeds korte gesprekjes voeren (of kleine taken uitvoeren). Voor deze 'Multiple Mini Interviews' is het meeste bewijs beschikbaar dat ze voorspellende waarde hebben voor prestaties binnen de opleiding.

Een derde belangrijke afweging is de mate van structuur in een interview. Dat kan uiteenlopen van ongestructureerd tot sterk gestructureerd, waarbij vooraf vastligt welke vragen er aan alle kandidaten gesteld gaan worden. Het grote voordeel van een gestructureerd interview is dat dat voor alle kandidaten vergelijkbare informatie oplevert, die voor de selectie dus ook vergeleken kan worden. Ongestructureerde interviews zijn niet geschikt als selectie-instrument omdat mensen niet goed in staat blijken te zijn om op basis van een open gesprek betrouwbare, consistente inschattingen te maken over de vaardigheden van andere mensen. Lastig daarbij is dat veel mensen wel denken dat ze dat kunnen (Kahneman, 2011). Een goed gestructureerd

interview met gestandaardiseerde vragen voor alle kandidaten en getrainde beoordelaars kunnen betrouwbare en valide instrumenten zijn. Echter, hoe meer gestructureerd het interview, hoe meer voorbereidingen dit met zich meebrengt. Als je interviews wilt meenemen in de selectieprocedure is het structureren van het interview dus belangrijk. Een mogelijk waardevolle methode om interviewvragen te structureren is de STARR (Situatie, Taak, Actie, Reactie, Reflectie)-methode. In [bijlage 3](#) zit een korte uitwerking van die methode.

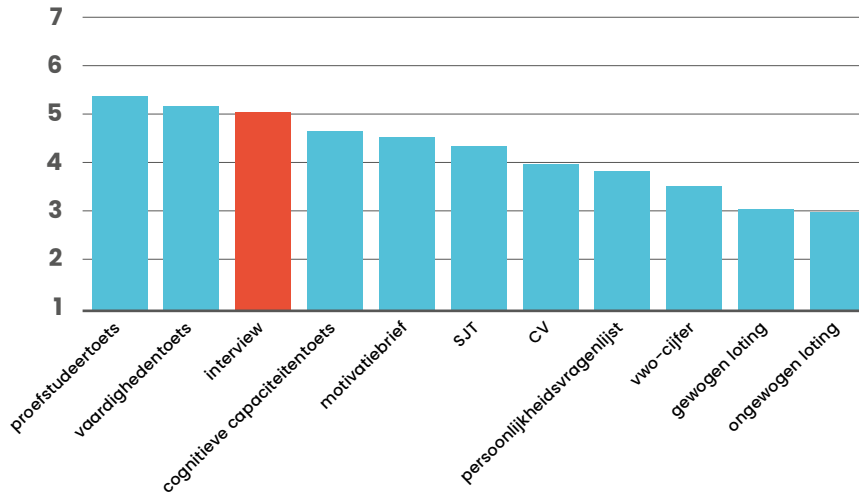
Ook voor het beoordelen van de gesprekken is het structureren een belangrijke afweging. De keuzes die je daarin maakt zijn gebaseerd op wat je hoopt in het interview te kunnen meten. Vooraf goed nadenken welke reacties van kandidaten op vragen welke score opleveren ligt daarbij voor de hand en kunnen ook erg behulpzaam zijn in het bepalen van de vragen zelf. Een afwijking daarbij is ook of de interviewers zelf de score bepalen, of dat zij feitelijk verslag doen van het gesprek (wat is er gezegd, of hoe is er gereageerd) en de beoordeling wordt bepaald door anderen. De interviewer heeft mogelijk meer informatie beschikbaar uit het gesprek die ook relevant kan zijn voor de beoordeling, maar kan ook beïnvloed worden door de ervaren sfeer tijdens het gesprek.

De inhoud van het interview is een laatste belangrijke afweging. Die afweging hangt sterk samen met de beoordeling en met de rol van het interview in de procedure als geheel. Een interview kan zich richten op de vraag of een kandidaat bij de opleiding past, waarbij bijvoorbeeld gevraagd wordt naar de achtergrond van de keuze, match tussen inhoud en vorm van de opleiding en motivatie, of de waarden en doelen van kandidaten. Het kan zich ook richten op welke vaardigheden een kandidaat heeft. Het kan hierbij gaan over academische vaardigheden, maar vaak wordt ook juist breder gevraagd naar vaardigheden als communicatie en samenwerking. Ook vaardigheden rond zelfreflectie kunnen een rol spelen in het gesprek. Afhankelijk van wat je allemaal wilt achterhalen in het gesprek variëren interviews doorgaans tussen de 15 en 60 minuten in lengte.

## Wat vinden kandidaten?

Interviews zijn meegenomen in het uitgebreide onderzoek van Fikrat-Wevers en collega's (2023) en worden door kandidaten relatief hoog gewaardeerd. Interviews worden veel gebruikt, niet alleen in de toelating tot opleidingen, maar ook bij het solliciteren naar een baan. Het is dus een bekende methode. Dat soort instrumenten worden over het algemeen beter gewaardeerd door kandidaten dan onbekende instrumenten. Kandidaten waarderen over het algemeen ook dat ze iets over zichzelf en hun motivatie kwijt kunnen in een interview.

In het onderzoek is geen onderscheid gemaakt in het type interview of de hoeveelheid structuur in het gesprek. Uit internationaal onderzoek blijkt dat kandidaten meer gestructureerde interviews hoger waarderen dan minder gestructureerde methoden (Patterson, Knight et al. (2016).



Figuur 18 • Wat vinden kandidaten

### Wat weten we uit onderzoek?

De betrouwbaarheid en validiteit van een interview hangt sterk samen met de mate van structuur in de vragen en beoordeling. Onderzoek is eenduidig over de onbetrouwbaarheid van ongestructureerde interviews. Het is niet duidelijk wat er precies gemeten wordt en het voorspelt niet hoe een kandidaat gaat presteren in de studie (of in het werk). In een internationaal onderzoek bleek bij ongestructureerde interviews 56% van de verschillen in scores tussen kandidaten te kunnen worden verklaard door wie de beoordelaar was (Patterson, Knight et al., 2016; Patterson & Zibarras, 2018).

Zelfs bij gestructureerde interviews blijft het uitdagend om te bepalen wat het instrument precies meet. Een duidelijke beoordelingslijst (rubric) waarmee de kandidaten beoordeeld worden kan helpen. Hiermee kan per vaardigheid informatie uit de antwoorden van de kandidaat worden gestructureerd (voor meer informatie over rubrics en het opstellen en beoordelen ervan, zie vooral het artikel van Van Strien & Joosten-ten Brinke, 2016). In het geval van meerdere beoordelaars is het belangrijk dat zij een gemeenschappelijk begrip van de criteria hebben. Hoe meer gestandaardiseerd de beoordeling is, hoe hoger de betrouwbaarheid zal zijn. Ook wijst onderzoek uit dat een langer interview betrouwbaarder is dan een kort interview, omdat het meer informatie geeft. De voorspellende waarde blijft vaak lager dan die

van andere instrumenten. Aan de andere kant laat onderzoek ook zien dat een interview kan zorgen voor een persoonlijkere band tussen de kandidaat en het programma (Patterson & Zibarras, 2018).

### Potentiële bias

Er is voor interviews bewijs dat suggereert dat de prestaties van kandidaten aanzienlijk kunnen worden beïnvloed door coaching (Patterson, Knight et al., 2016). Als er aanbod is voor dat soort coaching is het aannemelijk dat er sprake kan zijn van een bias op basis van de mogelijkheid om voor dergelijke coaching te betalen. Dat kan ook belangrijk zijn in het bepalen hoe het interview wordt ingericht.

Bij interviews kan er bij de beoordelaar mogelijk sprake zijn van een (onbewuste) bias tegen bepaalde kandidaten. Zeker bij ongestructureerde interviews die door de interviewer worden beoordeeld blijkt dat kandidaten waarin de interviewer zichzelf meer herkent vaak beter worden beoordeeld. Indien er meerdere beoordelaars zijn, zorgt diversiteit in achtergrond van de beoordelaars voor minder kans op (onbewuste) bias. Sowieso wordt dit effect kleiner als elk gesprek door meerdere beoordelaars gescoord wordt. Er zijn, waarschijnlijk in verband daarmee, aanwijzingen dat MMI's minder gevoelig zijn voor bias.



# 17 Psychomotorisch

## Wat is het?

Het derde domein in de taxonomie van Bloom is het psychomotorische domein. Het gaat hierbij om fysieke beweging, coördinatie en het gebruik van motorische vaardigheden in het algemeen. Het ontwikkelen van deze vaardigheden vereist oefening en wordt gemeten in termen van snelheid, precisie, afstand, en procedures of technieken in uitvoering (Bloom, 1956). Er zijn voor sommige opleidingen duidelijke redenen om al bij aanvang van de opleiding een goede basis in bepaalde psychomotorische vaardigheden te verwachten. Denk bijvoorbeeld aan het bespelen van een muziekinstrument voor het conservatorium of het beoefenen van sporten voor de opleiding tot sportopleiders. Het ligt dan voor de hand om deze vaardigheden mee te nemen in de selectieprocedure.

Hieronder volgen twee voorbeelden van opleidingen die psychomotorische vaardigheden meenemen in de selectie.

### Voorbeelden

Bij de bachelor psychomotorische therapie en bewegingsagogie aan de Hogeschool Windesheim worden motorische vaardigheden door middel van een bewegingsassessment getest. Dit onderdeel van een veel bredere selectieprocedure wordt met name gebruikt om de toelaatbaarheid van kandidaten vast te stellen. In dit instrument wordt naast beweegvaardigheden, ook gekeken naar leerbaarheid en sociale componenten. Deze vereisten worden op dit moment herzien:

**Je motorische vaardigheden (bewegingsassessment)** op het vlak van spel, turnen, atletiek, zwemmen, bewegen en muziek op basisniveau. De uitslag van het assessment is bindend. Een negatieve uitslag betekent dat je niet kunt starten. Deelname is slechts één keer per studiejaar mogelijk, zelfs

wanneer je je bij Windesheim voor meerdere sportopleidingen inschrijft. Op dit onderdeel kun je maximaal 5 punten behalen.

Let op! Je dient naast deze procedure ook een sportmedische keuring te ondergaan.

In de selectieprocedure voor de conservatoriumopleiding tot beiaardier aan de Hogeschool voor de Kunsten wordt kandidaten gevraagd een aantal muziekstukken voor te bereiden. Deze stukken moeten voorgespeeld worden tijdens het praktijkexamen, waarbij de uitvoering meetelt voor de toelating (de exacte procedure verschilt per instrument; dit geldt specifiek voor de beiaardiersopleiding).

Repertoire Enkele technisch eenvoudige stukjes, bijvoorbeeld:

- Engelse (of Franse) volksliedjes (bewerking L. 't Hart)
- Etudes van B. Winsemius (Etudenboek I)
- Etudes uit 'Ludus Modalis' van G. D'hollander
- Vier Dansen van G.Hengeveld
- Thema met 13 variaties van W.Vogel
- Tema Ostinato van W.Vogel

Van blad spelen enkele zeer eenvoudige stukjes van de moeilijkheidsgraad van bijvoorbeeld:

- Bonset, Thema uit Hernhutters Avondlied
- Staf Nees, Thema (uit: Thema met variaties, 1927)

Aandachtspunten

- Voldoende technische en muzikale aanleg voor het instrument
- Coördinatie manueel en pedaal
- Het duidelijk kunnen maken van zinstructuren, ritmiek, dynamiek, et cetera

## Belangrijke afwegingen

De meest centrale afweging bij het gebruik van psychomotorische tests is wat een kandidaat al bij aanvang van de opleiding moet kunnen, en wat ontwikkeld kan worden tijdens de opleiding. Een belangrijke rol daarin speelt de vraag of algemene 'handigheid' belangrijk is, of dat het gaat om een specifieke vaardigheid waarmee een kandidaat al veel oefening gehad moet hebben voor aanvang van de opleiding. Als een opleiding draait om het kundig worden in een psychomotorische vaardigheid, bereik je een veel hoger eindniveau als studenten aan het begin van de opleiding al een deel



van die vaardigheid ontwikkeld hebben. Om een expert te worden, heb je veel uren aan oefening nodig (Ericsson, Krampe & Tesch-Römer, 1993). Het is dan ook een voordeel als studenten al ervaren hebben dat ze daarvoor de motivatie kunnen opbrengen en ze deze uren al deels achter de rug hebben. Hierbij speelt natuurlijk de vraag wat het basisniveau is dat ten minste wordt verwacht, en helpt het om dat zo objectief mogelijk te kunnen formuleren en beoordeelbaar te maken. Hoe dat kan hangt natuurlijk sterk af van de vaardigheid die het betreft en is voor het bespelen van een instrument heel anders dan voor het beoefenen van een sport.

Voor opleidingen waarin psychomotorische vaardigheden vanaf het startpunt gaan worden aangeleerd, zou een psychomotorische test gebruikt kunnen worden om te beoordelen of kandidaten voldoende psychomotorische basisvaardigheden hebben zodat zij geen problemen hebben om het gewenste eindniveau te kunnen halen. In dat geval ligt de focus dus meer op het *uitselecteren*, of om een vorm van ‘harde’ selectie.

### Wat vinden kandidaten

In het onderzoek van Suzanne Fikrat-Wevers en collega's (2023) zijn geen psychomotorische instrumenten meegenomen. Er is weinig bekend over hoe kandidaten staan tegenover het gebruik van deze instrumenten.

### Wat weten we uit onderzoek?

De literatuur is niet helemaal eenduidig over het meten van psychomotorische vaardigheden voor selectiedoeleinden. Dat lijkt met name te zitten in het uitgangspunt dat veel van deze vaardigheden aan te leren zijn, in plaats van voort te komen uit een algemene aanleg voor dit soort type vaardigheid (Giuliani et al., 2007). Het onderzoek naar het gebruik van psychomotorische vaardigheden brengt daarnaast veel uitdagingen met zich mee, bijvoorbeeld in het afnemen van tests voor sterk uiteenlopende psychomotorische vaardigheden bij dezelfde groep mensen (Carretta & Ree, 2000). De analyse van Carretta en Ree (2000) geeft ook enkele aanwijzingen dat er mogelijk toch een algemene aanleg voor psychomotorische vaardigheden (een soort maat voor ‘handigheid’) bestaat als het onderzoek anders ingericht wordt.

Groenier, Groenier, Miedema en Broeders (2015) laten zien dat assistenten met betere psychomotorische vaardigheden een laparoscopische ingreep sneller en met betere resultaten onder de knie kregen (Groenier, Groenier, Miedema & Broeders, 2015). Meer oefenen, ook mogelijk voor de start van de opleiding, lijkt dus te werken. Dat wordt versterkt doordat er ook bewijs is dat psychomotorische vaardigheden generaliseren voorbij de specifieke

vaardigheid die geoefend is. Zo is er onderzoek dat laat zien dat je psychomotorische vaardigheden kunt oefenen door middel van videogames, en dat die vaardigheden je later helpen bij chirurgische ingrepen (Kennedy, Boyle, Traynor, Walsh & Hill, 2011). Een dergelijk effect wordt ook in de selectiecontext gevonden. De HAM-Man-test wordt bijvoorbeeld succesvol bij een geneeskundeopleiding in Duitsland gebruikt om te voorspellen hoe snel kandidaten nieuwe psychomotorische vaardigheden kunnen uitvoeren (Kothe, Hissbach & Hampe, 2014).

Ook bij selectie voor de pilotenopleiding is het gebruik van psychomotorische toetsen al redelijk ingeburgerd. Om vast te stellen of nieuwe psychomotorische tests daar toegevoegde waarde zouden hebben in een selectieprocedure, is bijvoorbeeld onderzocht of ervaren piloten het op die toetsen beter doen (Ostoin, 2007). Dat blijkt inderdaad het geval te zijn.

Evans en Dirks (2001) concluderen aan de andere kant dat in het opleiden voor tandtechnische laboratoriumtechniek, het reeds bezitten van psychomotorische vaardigheden geen absolute voorspeller is van prestaties op psychomotorische vaardigheden in de studie. Zij concluderen wel dat psychomotorische vaardigheden binnen selectie kunnen dienen om kandidaten uit te sluiten waarvan verwacht kan worden dat zij de psychomotorische vaardigheden niet zullen kunnen aanleren (Evans & Dirks, 2001).

## Potentiële bias

In het geval van psychomotorische instrumenten waarbij al een bepaald niveau van een vaardigheid verwacht wordt (zoals bij muziek-, sport- of kunstopleidingen), hebben mogelijk niet alle kandidaten gelijke toegang gehad tot lessen en ondersteuning bij het aanleren en oefenen. Deze instrumenten veroorzaken hiermee mogelijk een bias tegenover kandidaten uit gezinnen met minder financiële armslag, of kandidaten die bijvoorbeeld meer tijd hebben moeten besteden aan zorgtaken. Instrumenten die bekijken hoe snel kandidaten nieuwe psychomotorische vaardigheden aanleren, zijn minder gevoelig voor coaching en bias.



# 18 Loting

## Wat is het?

Loting als selectie-instrument heeft een heel ander uitgangspunt dan de andere instrumenten die we in dit handboek bespreken. Het gaat om het toelaten van studenten op basis van kans, in plaats van op basis van hun kwaliteiten. Op het moment van schrijven is loten voor toelating tot opleidingen in het hoger onderwijs niet toegestaan in Nederland, maar is een wetsvoorstel dat dat wel mogelijk maakt voor bacheloropleidingen met een numerus fixus door de Tweede Kamer aangenomen. Daarom bespreken we het hier wel binnen de kaders die in dat wetsvoorstel gegeven zijn. Het is dus belangrijk om nog juridisch advies in te winnen over wat er definitief wel en niet is toegestaan.

Er zijn in het wetsvoorstel drie mogelijke manieren om loting in te zetten:

- Ongewogen loting, waarbij alle kandidaten die aan de instroomvereisten voldoen evenveel kans maken om toegelaten te worden;
- Gewogen loting, waarbij de kans om toegelaten te worden toeneemt aan de hand van de score op andere selectie-instrumenten;
- Loting binnen een subgroep, waarbij (al dan niet gewogen) geloot wordt binnen een groep kandidaten, nadat andere kandidaten direct zijn toegelaten of zijn afgewezen.

Of loten een betere of rechtvaardigere manier is om plekken in het hoger onderwijs te verdelen dan andere instrumenten, is een normatieve vraag en geen vraag die op basis van empirisch onderzoek te beantwoorden is (zie ook Steenman, 2020). Een manier om hierover na te denken is door te reflecteren op de aard van wat je aan het verdelen bent. Past het bij plaatsen in het hoger onderwijs om die te verdelen op basis van kans, of past het juist om die te verdelen op basis van bepaalde kwaliteiten van kandidaten? Een andere vraag is of je voldoende onderscheid kunt maken tussen kandidaten met behulp van selectie-instrumenten.

Voor dat laatste speelt vooral de verhouding tussen het aantal plaatsen en het aantal kandidaten een grote rol. Voor opleidingen waarvoor het verschil tussen plaatsen en aanmeldingen klein is, kan het soms lastig zijn om instrumenten te vinden die voldoende onderscheid maken tussen kandidaten. Vooral vanuit psychologieopleidingen in Nederland komt daarom de roep om een combinatie te kunnen maken tussen het selecteren van de hoogst en laagst scorende kandidaten, en binnen andere groepen te loten (Visser, 2019; Van der Smagt, Tick & Van Ooijen, 2022). Een nog beperktere vorm is om te loten tussen kandidaten die een sterk vergelijkbare score hebben om hun precieze plaats te bepalen binnen de uiteindelijke ranking.

Gewogen loting is grotendeels te categoriseren als selecteren op het instrument of de instrumenten waarmee je de weging bepaalt. Het grote verschil is dat alle kandidaten kans houden om toegelaten te worden, ook als hun uiteindelijke score op die instrumenten erg laag is. Het is daarbij vooral belangrijk dat je helder kunt uitleggen waarom je toch vindt dat iedereen een kans moet maken om toegelaten te worden, maar waarom je die kans voor bepaalde groepen toch groter wilt hebben.

## Belangrijke afwegingen

Veel belangrijke afwegingen gelden niet voor ongewogen loting. De praktische uitvoering van loting is in alle gevallen waarbij het lot een rol speelt in de toelating van belang. Bedenk goed op welke manier je gaat loten (bijvoorbeeld met het gebruik van betrouwbare software), hoe je gewichten toekent, hoe je daarbij voorkomt dat dezelfde kandidaat vaker wordt toegelaten, et cetera. Nu centrale loting niet terug lijkt te komen, is het nog onduidelijk welke wettelijke vereisten zullen gaan gelden voor het lotingsproces voor opleidingen. Of de procedure onder toezicht van een notaris moet vallen, zoals bij loterijen gebruikelijk is, is bijvoorbeeld nog onduidelijk.

Bij gewogen loting is een belangrijke afweging hoeveel verschil in kansen je wilt introduceren. Hoe meer gewicht je toekent aan de selectiescore, hoe kleiner de kans dat kandidaten met een lage score binnenkomen. Bij het gebruik van instrumenten om de gewichten te bepalen spelen natuurlijk de afwegingen een rol die bij die instrumenten in de vorige hoofdstukken naar voren komen. In het wetsvoorstel zoals dat door de Tweede Kamer is aangenomen moet het gewicht voor elke kandidaat bepaald worden op basis van ten minste twee criteria (die niet beide gebaseerd mogen zijn op vooropleidingscijfers).

Bij het loten binnen een subgroep is een belangrijke afweging vooral waar de grenzen liggen tussen de groep die direct wordt toegelaten en de groep die aan de loting moet deelnemen. Van der Linden (2018) suggereert dat

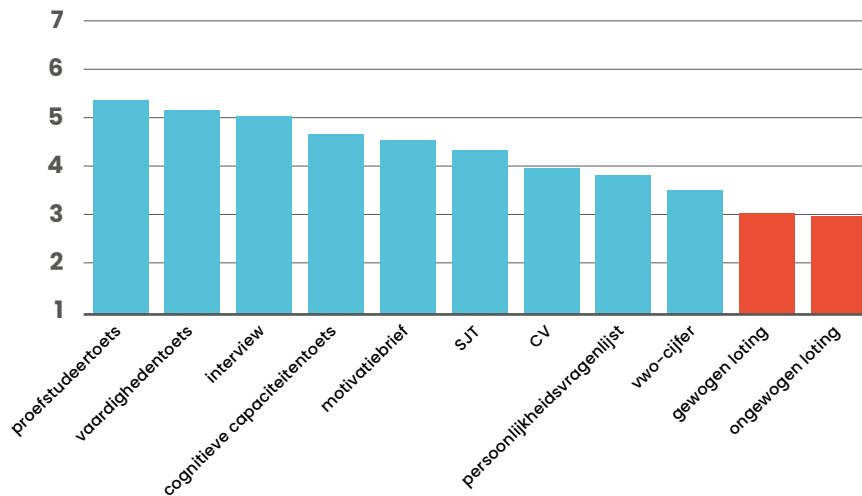
signaaldetectie behulpzaam kan zijn in het bepalen van de grens tussen directe toelating en de categorie waarbinnen loting passend zou kunnen zijn. Signaaldetectie is een methode om voor verschillende grenswaarden en voor verschillende instrumenten te berekenen hoeveel toegelaten studenten goed of juist niet goed presteren binnen de opleiding, en hoeveel van de studenten die je zou afwijzen goed of juist niet goed zouden presteren. De verhouding tussen die percentages kan helpen bij het gesprek over grenswaarden.

De ruimte voor opleidingen om te bepalen wat de ondergrens is waaronder kandidaten direct afgewezen worden, lijkt door toezeggingen van UNL en de Nederlandse Federatie van Universitair Medische Centra (NFU) beperkt te zijn. Deze organisaties hebben toegezegd dat alleen kandidaten die niet hebben voldaan aan een inspanningsverplichting in de selectieprocedure zullen worden uitgesloten van de loting binnen de groep die niet direct wordt toegelaten.

Ook bij loting binnen een subgroep moet het bepalen van de groep die direct wordt toegelaten en de groep die nog moet loten in het huidige wetsvoorstel gebeuren aan de hand van ten minste twee criteria, waarbij niet beide criteria gebaseerd mogen zijn op vooropleidingscijfers. Bij de groep die nog aan de loting deelneemt, mag ook weer gebruikgemaakt worden van gewogen loting, waarbij sommige kandidaten meer kans hebben om toegelaten te worden dan andere kandidaten. Het gewicht moet dan eveneens gebaseerd zijn op twee criteria, en ook die mogen niet beide volledig steunen op vooropleidingscijfers. Welke instrumenten je gebruikt voor het bepalen van de groep die direct wordt toegelaten, en welke instrumenten je gebruikt voor het wegen zijn daarmee de laatste belangrijke afweging. De voorgaande hoofdstukken bieden meer informatie voor die afweging.

### Wat vinden kandidaten?

Uit het onderzoek van Suzanne Fikrat-Wevers en collega's (2023) blijkt dat kandidaten loting het minst waarderen van alle selectie-instrumenten. De gewogen loting wordt nipt beter gewaardeerd dan de ongewogen loting. Kandidaten geven aan dat het voor hen eerlijker voelt als ze zelf invloed hebben op of ze al dan niet toegelaten worden.



Figuur 19 • Wat vinden kandidaten

### Wat weten we uit onderzoek?

In het geval van ongewogen loting is er geen sprake van betrouwbaarheid of validiteit. Bij gewogen loting geldt voor de validiteit en betrouwbaarheid vooral wat we weten van de instrumenten die gebruikt zijn voor de weging. Onderzoek dat is uitgevoerd in de tijd dat bij verschillende geneeskundeopleidingen zowel kandidaten werden toegelaten via selectie als via op vwo-cijfers gewogen loting, blijkt dat kandidaten die via de selectie werden toegelaten beter presteerden binnen de opleiding. Zelfs kandidaten die waren afgewezen voor de selectie en via loting waren binnengekomen vielen minder vaak uit en studeerden sneller dan de groep die zonder eerst aan de selectie deel te nemen via loting binnen kwam (Schreurs, Cleutjens, Cleland et al., 2020; Schreurs, Cleutjens et al., 2018; Schripsema, 2017; Urlings-Strop, Stijnen, Themmen & Splinter, 2009; Urlings-Strop, Themmen, Stijnen & Splinter, 2011).

Hierbij kun je de vraag stellen of dat een effect is van de selectie zelf, of dat kandidaten die zich aanmeldden om via selectie binnen te komen geen willekeurige groep waren. Uit onderzoek blijkt echter dat op alle variabelen die meegenomen konden worden, de groepen vergelijkbaar waren. Het gemiddelde vooropleidingscijfer was een van die variabelen: de lotinggroep en selectiegroep (zowel toegelaten als afgewezen) hadden een vergelijkbaar gemiddeld vooropleidingscijfer. Verder blijkt uit een onderzoek bij een geneeskundeopleiding dat de studierendementen als alle plaatsen via selectie worden toegekend daar beter zijn dan wanneer diezelfde opleiding alle studenten toelaat via loting (Schreurs, Cleland et al., 2018). Wat blijft staan is dat kandidaten die zich voor de selectieprocedure aanmeldden, mogelijk

meer gemotiveerd zijn voor de opleiding. Aan de andere kant laat onderzoek zien dat selectie niet in alle gevallen betere resultaten laat zien dan loting. De proefschriften van Wouters (2017) en Schripsema (2017) laten dat voor specifieke selectieprocedures zien.

## Potentiële bias

Ongewogen loting heeft als vanzelfsprekend geen onterechte bias (en ook geen mogelijk terechte bias voor bijvoorbeeld kandidaten die bijzonder goed bij de opleiding passen), en werkt in vergelijking met selectie-instrumenten in het voordeel van kandidaten die op die instrumenten slecht zouden scoren. Bij gewogen loting geldt natuurlijk de bias van de instrumenten waarop de gewichten in de loting zijn gebaseerd.

Onderzoek van de onderwijsinspectie laat zien dat verschillende groepen scholieren een grotere drempel ervaren om zich aan te melden voor een opleiding die loot dan voor een opleiding die niet loot of selecteert (Inspectie van het Onderwijs, 2022). Figuur 20 geeft een overzicht voor welke groepen dat geldt. Over het algemeen lijkt loten iets meer effect te hebben op het aanmeldgedrag van specifieke groepen dan selecteren, in elk geval voor groepen die de onderwijsinspectie aanduidt als ‘risicogroep’. Sommige risicogroepen lijken zich voor selecterende opleidingen juist sneller aan te melden dan voor opleidingen die niet selecteren of loten. Dat geldt bijvoorbeeld voor mbo’ers en havisten met een niet-westerse migratieachtergrond. Zie verder ook het rapport van de Inspectie van het Onderwijs (2022).

Het effect van loting op de samenstelling van de studentenpopulatie en de representatie daarin van specifieke groepen studenten hangt verder vooral samen met welke andere procedure vervangen wordt door loting. Het vervangen van een slechte selectieprocedure met veel bias zal de representatie van specifieke groepen zeker kunnen verbeteren. Aan de andere kant laat Amerikaans onderzoek van Baker en Bastedo (2021, p. 7) “dramatische en negatieve potentiële effecten van loting op de toelating van gekleurde studenten, studenten met een laag inkomen en mannen” zien. In dat fictieve experiment wordt een selectieve procedure vervangen door een loting met een ondergrens die je zou kunnen vergelijken met de instroomeis van het juiste niveau in het voortgezet onderwijs, dus havo of vwo in de Nederlandse context. Mulder et al. (2022) betogen dat de ongelijke representatie in het hoger onderwijs met name ligt in de stap van primair naar voortgezet onderwijs, waar leerlingen in het vmbo, havo en vwo gesorteerd worden, veel meer dan in de stap naar aanmelding en toelating (p. 794).

Model	Kenmerken	Mbo 4 - studenten	Havisten	Vwo'ers
1	Achtergrond	<p>Studenten met een westerse migratie-achtergrond</p> <p>Studenten met een laag gemiddeld rapportcijfer</p> <p>Studenten uit de lagere inkomensgroepen</p>	<p>Mannelijke scholieren</p> <p>Scholieren met ouders die geen opleiding hebben</p> <p>Scholieren met ouders met een mbo 3- of 4- of vergelijkbare opleiding</p>	Scholieren met een psychische of lichamelijke beperking
2	Sociaal netwerk	Studenten met ouders/verzorgers die weinig stimuleren tot verdere studie in het hoger onderwijs	Scholieren met vrienden die een negatieve houding ten opzichte van het hoger onderwijs hebben	Scholieren met ouders/verzorgers die weinig stimuleren tot verdere studie in het hoger onderwijs
3	Persoonlijkheid	Scholieren met risicoaversie	Scholieren met risicoaversie	Scholieren die leenaversie hebben
4	School			
5	Overig	Studenten die in de regio Noord of Zuid wonen	Scholieren met bijzondere omstandigheden	

Bron: Inspectie van het Onderwijs (2022), p.40.

*Figuur 20 · Groepen die zich minder snel aanmelden wanneer een opleiding loot, ten opzichte van wanneer een opleiding niet loot en niet selecteert.*

## 19 Afsluiting

In dit handboek hebben we geprobeerd de kennis die op dit moment in de Nederlandse context beschikbaar is over het vormgeven van selectie en de verschillende instrumenten die daarbij gebruikt worden zo toegankelijk mogelijk weer te geven. We hebben daarbij steeds belangrijke afwegingen gegeven, en argumenten om binnen afwegingen zorgvuldig keuzes te kunnen maken. Als vanzelfsprekend hebben we ook regelmatig gekozen om details weg te laten. De literatuurlijst biedt veel mogelijkheden tot verdere verdieping. De leden van het OnderzoekersNetwerk zijn daarnaast bereid om mee te denken over specifieke vragen. Dergelijke vragen kunnen worden gestuurd naar de coördinatoren van het OnderzoekersNetwerk: Kitty Cleutjens, Karen Stegers-Jager en Sebastiaan Steenman.

Het centrale uitgangspunt in alle delen van dit handboek is het idee van alignment, of afstemming van selectie op de inhoud en didactiek van de opleiding. Daarom adviseren we in dit handboek ook niet één procedure, maar bespreken we het selectieproces in het algemeen en de verschillende inhoudelijke en procesmatige keuzes die daarbij horen. Het uitgangspunt van alignment van selectie betekent dat verschillende opleidingen op verschillende manieren ‘goed’ kunnen selecteren. Aan de andere kant zijn er ook afwegingen die inherent bij een instrument of bij het inrichten van een selectieprocedure horen.

Ook de context van een opleiding speelt bij de afwegingen rond het inrichten van een selectieprocedure een belangrijke rol. In verschillende hoofdstukken staan we bijvoorbeeld stil bij de relevantie van de verhouding tussen het aantal plekken en het aantal aanmeldingen, de aanwezigheid van internationale instroom, en andere diversiteit in vooropleidingen.

Aan de andere kant is het voor kandidaten ook ingewikkeld wanneer voor een vergelijkbare opleiding verschillende selectieprocedures gebruikt worden. Die complexiteit kan toenemen wanneer loting als potentieel instrument wordt toegevoegd aan het repertoire van opleidingen. Kandidaten zullen dan

nog meer strategische afwegingen gaan maken tussen verschillende procedures op basis van hun inschatting binnen welke procedure ze de meeste kansen maken. In het algemeen onderstreept dit het belang van transparant communiceren over de selectieprocedure richting kandidaten en richting aanleverende scholen. Daarbij kan het ook waarde hebben om expliciet te maken welke keuzes een opleiding niet heeft gemaakt, en die keuzes inhoudelijk te onderbouwen.

Het perspectief van de kandidaat dwingt ook tot afweging op het gebied van de inspanning die gevraagd wordt tijdens het selectieproces. Sommige instrumenten vragen veel meer tijd van kandidaten dan andere. In de afwegingen per instrument staan we daar ook bij stil. Het aantal criteria, en zelfs het aantal instrumenten is dus niet direct verbonden aan de tijdsdruk die kandidaten ervaren door de selectieprocedure. Het is een overweging die gaat over het precieze gebruik van specifieke instrumenten.

Een daarmee samenhangende afweging is dat het belangrijk is om je ervan bewust te zijn dat het meten op één moment altijd nadelen heeft. Het heeft daarom vanuit het belang van een betrouwbare selectie de voorkeur om criteria op meerdere momenten met meerdere meetinstrumenten in kaart te brengen. Het zoeken naar mogelijke combinaties van instrumenten waarin voor een deel overlap mogelijk is, maar die elkaar ook aanvullen in hun verschillende kwaliteiten is daarom de moeite waard.

Als laatste is het goed en gestructureerd evalueren van de procedure en van de verschillende instrumenten belangrijk. In het handboek bespreken we de algemene uitgangspunten van de procedure en de instrumenten. Dat zijn gestaaftde conclusies, maar de manier waarop de procedure en de instrumenten bij een specifieke opleiding werken is altijd afhankelijk van de precieze invulling van de selectie en van de manier waarop er binnen de opleiding getoetst wordt (dat is immers meestal de afhankelijke variabele in de evaluatie van de selectieprocedure). Nadrukkelijke aandacht voor die evaluatie, verder wetenschappelijk onderzoek en het delen van de resultaten daarvan met anderen, maakt dat selectieprocedures steeds kunnen worden verbeterd.









# Literatuur

Ab Latif, R., & Mat Nor, M. Z. (2020). Using the ADDIE Model to Develop a Rusnani Concept Mapping Guideline for Nursing Students. *Malays Journal of Medical Science*, 27(6), 115-127. doi:10.21315/mjms2020.27.6.11

Baker, D. J., & Bastedo, M. N. (2021). What If We Leave It Up to Chance? Admissions Lotteries and Equitable Access at Selective Colleges. *Educational Researcher*, 51(2), 134-145. 10.3102/0013189X211055494

van Berkel, H., Bax, A., & Joosten-ten Brinke, D. (2017). *Toetsen in het hoger onderwijs*. Bohn Stafleu van Loghum.

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher education*, 32(3), 347-364.

Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. David McKay Company, Inc.

Carretta, T. R., & Ree, M. J. (2000). General and specific cognitive and psychomotor abilities in personnel selection: The prediction of training and job performance. *International Journal of Selection and Assessment*, 8(4), 227-236.

Ceci, A. J., & Peters, D. (1984). Letters of reference: A naturalistic study of the effects of confidentiality. *American Psychologist*, 39(1), 29-31.

Cleland, J., Blitz, J., Cleutjens, K., oude Egbrink, M., Schreurs, S., & Patterson, F. (2023). Robust, defensible, and fair: The AMEE guide to selection into medical school: AMEE Guide No. 153. *Medical teacher*, 1-14.

Collares, C. F. (2021). Cognitive Diagnostic Modelling in Healthcare Professions Education: An Eye-opener. *Advances in Health Sciences Education*, 27(2), 427-440.

de Boer, T. (2020). *Master selection and admission in the faculty of geosciences at UU*. Paper gepresenteerd bij de VSNU Learning Community, Utrecht.

de Boer, T., & van Rijnsoever, F. J. (2021). In search of valid non-cognitive student selection criteria. *Assessment & Evaluation in Higher Education*, 47(5), 783-800. <https://doi-org.proxy.library.uu.nl/10.1080/02602938.2021.1958142>

de Boer, T., & Van Rijnsoever, F. J. (2022). One field too far? Higher cognitive relatedness between bachelor and master leads to better predictive validity of bachelor grades during admission. *Assessment & Evaluation in Higher Education*, 1-14. <https://doi-org.proxy.library.uu.nl/10.1080/02602938.2022.2158453>

de Leng, W. E. (2019). *Measuring Integrity for Selection into Medical School: Development of a Situational Judgement Test*. Erasmus University Rotterdam.

de Leng, W. E., Stegers-Jager, K. M., Husbands, A., Dowell, J. S., Born, M. P., & Themmen, A. P. N. (2017). Scoring method of a Situational Judgment Test: influence on internal consistency reliability, adverse impact and correlation with personality? *Advances in Health Sciences Education*, 22(2), 243-265. doi:10.1007/s10459-016-9720-7

de Leng, W. E., Stegers-Jager, K. M., Born, M. P., & Themmen, A. P. (2018). Integrity situational judgement test for medical school selection: judging 'what to do' versus 'what not to do'. *Medical education*, 52(4), 427-437. Geraadpleegd via: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5901405/pdf/MEDU-52-427.pdf>

de Visser, M., Fluit, C., Fransen, J., Latijnhouwers, M., Cohen-Schotanus, J., & Laan, R. (2017). The effect of curriculum sample selection for medical school. *Advances in Health Sciences Education*, 22(1), 43-56. doi:10.1007/s10459-016-9681-x

de Visser, M., Laan, R., Engbers, R., Cohen-Schotanus, J., & Fluit, C. (2018). Selection as a learning experience: an exploratory study. *Advances in medical education and practice*, 9, 345-355.

de Vries, A., Reinout, E., Born, M. P., & van den Berg, R. H. (2014). Persoonlijkheid als voorspeller van werkprestatie en contraproductief werkgedrag: Het belang van specifieke persoonlijkheidsmetingen. *Gedrag & Organisatie*, 27(4).

Deci, E. L., & Ryan, R. M. (2012). Self-determination theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Red.), *Handbook of theories of social psychology* (pp. 416-436). Sage Publications Ltd. <https://doi-org.proxy.library.uu.nl/10.4135/9781446249215.n21>

Ding, H. (2007). Genre analysis of personal statements: Analysis of moves in application essays to medical and dental schools. *English for specific purposes*, 26(3), 368-392.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review*, 100(3), 363-406. <https://doi.org/10.1037/0033-295X.100.3.363>

Eva, K. W., Rosenfeld, J., Reiter, H. I., & Norman, G. R. (2004). An admissions OSCE: the multiple mini-interview. *Medical education*, 38(3), 314-326.

Evans, J. G., & Dirks, S. J. (2001). Relationships of admissions data and measurements of psychological constructs with psychomotor performance of dental technology students. *Journal of Dental Education*, 65(9), 874-882.

Ferguson, E., James, D., & Madeley, L. (2002). Factors associated with success in medical school: systematic review of the literature. *BMJ*, 324(7343), 952-957.

Fikrat-Wevers, S. (2022). *Cijfers als toelatingmethode: diversiteit en percepties van kandidaten*. Paper gepresenteerd bij Bijeenkomst NVMO werkgroep selectie.

Fikrat-Wevers, S., Stegers-Jager, K. M., Afonso, P. M., Koster, A. S., Van Gestel, R. A., Groenier, M., . . . Woltman, A. M. (2023). Selection tools and student diversity in health professions education: a multi-site study. *Advances in health sciences education*, 1-26.

Fikrat-Wevers, S., Stegers-Jager, K. M., Groenier, M., Koster, A., Ravestloot, J., Van Gestel, R., . . . Woltman, A. (2023). Applicant perceptions of selection methods for health professions education: Rationales and subgroup differences. *Medical education*, 57(2), 170-185.



- Fossati, A., Borroni, S., Marchione, D., & Maffei, C. (2011). The big five inventory (BFI). *European Journal of Psychological Assessment*, 27(1), 50-58. <https://doi-org.proxy.library.uu.nl/10.1027/1015-5759/a000043>
- George, J., & Marett, K. (2004). The truth about lies. *HR Magazine*, 49(5), 87-91.
- Giuliani, M., Lajolo, C., Clemente, L., Querqui, A., Viotti, R., Boari, A., & Miani, C. (2007). Is manual dexterity essential in the selection of dental students? *British Dental Journal*, 203(3), 149-155.
- Griffin, B., & Wilson, I. G. (2012). Faking good: self-enhancement in medical school applicants. *Medical Education*, 46(5), 485-490. doi:10.1111/j.1365-2923.2011.04208.x
- Griffin, B., Yeomans, N. D., & Wilson, I. G. (2013). Students coached for an admission test perform less well throughout a medical course. *Internal Medicine Journal*, 43(8), 927-932. doi:10.1111/imj.12171
- Groenier, M., Groenier, K. H., Miedema, H. A., & Broeders, I. A. (2015). Perceptual speed and psychomotor ability predict laparoscopic skill acquisition on a simulator. *Journal of surgical education*, 72(6), 1224-1232.
- Hooker, E. A., Mallow, P., Downes, C., & Baidwan, S. (2022). Abandoning the GRE and GMAT Improved Diversity in a Graduate Program for Health Administration. *The Journal of Health Administration Education*, 38(4), 895.
- Houser, C., & Lemmons, K. (2018). Implicit bias in letters of recommendation for an undergraduate research internship. *Journal of Further and Higher Education*, 42(5), 585-595.
- Inspectie van onderwijs. (2022). *Onbedoelde zelfselectie: drempels die gekwalificeerde jongeren ervan weerhouden om een specifieke opleiding in het hoger onderwijs te kiezen*. Rapport.
- Inspectie van onderwijs. (2023). *Selectie in het hoger onderwijs: criteria, instrumenten en de borging van kansengelijkheid: Een onderzoek naar de onderbouwing van selectieprocedures*. Rapport.
- Kahneman, D. (2011). *Thinking fast and slow*. Straus and Gioux.
- Kass, D., Kim, J. S., LaMacchia, C., & Bekisz, A. (2020). Does the GMAT matter? The impact of a GMAT waiver policy on MBA student performance. *Journal of Education for Business*, 95(7), 432-438.
- Kennedy, A.-M., Boyle, E., Traynor, O., Walsh, T., & Hill, A. (2011). Video gaming enhances psychomotor skills but not visuospatial and perceptual abilities in surgical trainees. *Journal of surgical education*, 68(5), 414-420.
- Knouse, S. B. (1983). The letter of recommendation: Specificity and favorability of information. *Personnel Psychology*, 36(2), 331-341.
- Koczwara, A., Patterson, F., Zibarras, L., Kerrin, M., Irish, B., & Wilkinson, M. (2012). Evaluating cognitive ability, knowledge tests and situational judgement tests for postgraduate selection. *Medical Education*, 46(4), 399-408. doi:10.1111/j.1365-2923.2011.04195.x
- Kohnstamm, R. (2009). Persoonlijkheid in wording. In *Kleine ontwikkelingspsychologie I* (pp. 195-215): Springer.

- Kothe, C., Hissbach, J., & Hampe, W. (2014). Prediction of practical performance in pre-clinical laboratory courses—the return of wire bending for admission of dental students in Hamburg. *GMS Zeitschrift für Medizinische Ausbildung*, 31(2).
- Krathwohl, D. R. (2002). A revision of Bloom’s taxonomy: An overview. *Theory into practice*, 41(4), 212–218.
- Kuncel, N. R. (2011). *Measurement and meaning of critical thinking*. Paper gepresenteerd bij de Draft report for the National Research Council’s 21st century skills workshop.
- Kuncel, N. R., & Borneman, M. J. (2007). Toward a new method of detecting deliberately faked personality tests: The use of idiosyncratic item responses. *International Journal of Selection and Assessment*, 15(2), 220–231. doi:DOI 10.1111/j.1468–2389.2007.00383.x
- Kuncel, N. R., Kochevar, R. J., & Ones, D. S. (2014). A Meta-analysis of Letters of Recommendation in College and Graduate Admissions: Reasons for hope. *International Journal of Selection and Assessment*, 22(1), 101–107. doi:10.1111/ijsa.12060
- Kuryшева, A., & Krijgsheld, P. (2021). *Why do we need a transfer to standardized letters of recommendation?* Utrecht University.
- Kuryшева, A., van Rijen, H. V. M., & Dilaver, G. (2019). How do admission committees select? Do applicants know how they select? Selection criteria and transparency at a Dutch University. *Tertiary education and management*, 25, 367–388.
- Lievens, F. (2013). Adjusting medical school admission: assessing interpersonal skills using situational judgement tests. *Medical Education*, 47(2), 182–189. doi:10.1111/medu.12089
- Lievens, F., Sackett, P. R., & De Corte, W. (2021). Weighting admission scores to balance predictiveness–diversity: The Pareto–optimization approach. *Medical Education*. doi:10.1111/medu.14606
- Liu, O. L., Bridgeman, B., & Adler, R. M. (2012). Measuring learning outcomes in higher education: Motivation matters. *Educational Researcher*, 41(9), 352–362.
- Mahoney, J. L., Cairns, B. D., & Farmer, T. W. (2003). Promoting interpersonal competence and educational success through extracurricular activity participation. *Journal of educational psychology*, 95(2), 409–418. <https://doi-org.proxy.library.uu.nl/10.1037/0022-0663.95.2.409>
- McCarthy, J. M., & Goffin, R. D. (2001). Improving the validity of letters of recommendation: An investigation of three standardized reference forms. *Military Psychology*, 13(4), 199–222.
- McManus, I., Woolf, K., Dacre, J., Paice, E., & Dewberry, C. (2013). The Academic Backbone: longitudinal continuities in educational achievement from secondary school and medical school to MRCP (UK) and the specialist register in UK medical students and doctors. *BMC medicine*, 11(1), 1–27.
- Meeus, W. (2019). Iemand worden: ontwikkeling van de persoonlijkheid. In *Vallen en opstaan in de adolescentie* (pp. 41–53): Springer.
- Meijer, R. R., Neumann, M., Hemker, B. T., & Niessen, A. S. M. (2019). A Tutorial on Mechanical Decision-Making for Personnel and Educational Selection. *Frontiers in Psychology*, 10, 3002. doi:10.3389/fpsyg.2019.03002



- Mogridge, C. D. (2019). Recruitment, selection, and assessment: Are the CV and interview still worth using? *Psychreg Journal of Psychology*, 3(3), 40–45.
- Molenda, M. (2003). In search of the elusive ADDIE model. *Performance improvement*, 42(5), 34–37.
- Mulder, L., Wouters, A., Twisk, J. W., Koster, A. S., Akwiwu, E. U., Ravestloot, J. H., . . . Kusurkar, R. A. (2022). Selection for health professions education leads to increased inequality of opportunity and decreased student diversity in The Netherlands, but lottery is no solution: A retrospective multi-cohort study. *Medical teacher*, 44(7), 790–799. <https://doi-org.proxy.library.uu.nl/10.1080/0142159X.2022.2041189>
- Niessen, A. S. M., Meijer, R. R., & Tendeiro, J. N. (2016). Predicting Performance in Higher Education Using Proximal Predictors. *PLoS One*, 11(4), e0153663. doi:ARTNe015366310.1371/journal.pone.0153663
- Niessen, A. S. M., Meijer, R. R., & Tendeiro, J. N. (2017a). Applying organizational justice theory to admission into higher education: Admission from a student perspective. *International Journal of Selection and Assessment*, 25(1), 72–84.
- Niessen, A. S. M., Meijer, R. R., & Tendeiro, J. N. (2017b). Measuring non-cognitive predictors in high-stakes contexts: The effect of self-presentation on self-report instruments used in admission to higher education. *Personality and individual differences*, 106(1), 183–189. <https://doi.org/10.1016/j.paid.2016.11.014>
- Niessen, A. S. M., & Neumann, M. (2022). Using personal statements in college admissions: An investigation of gender bias and the effects of increased structure. *International Journal of Testing*, 22(1), 5–20.
- Ostoin, S. D. (2007). *An assessment of the performance-based measurement battery (PBMB), the Navy's psychomotor supplement to the Aviation Selection Test Battery (ASTB)*. Naval Postgraduate School Monterey CA.
- Patterson, F., Ferguson, E., & Knight, A. (2010). Selection for medical education and training. In Swanwick, T. (Red.). *Understanding medical education: Evidence, theory and practice*: (pp. 352–365).
- Patterson, F., Ferguson, E., & Thomas, S. (2008). Using job analysis to identify core and specific competencies: implications for selection and recruitment. *Medical Education*, 42(12), 1195–1204. doi:10.1111/j.1365-2923.2008.03174.x
- Patterson, F., Knight, A., Dowell, J., Nicholson, S., Cousans, F., & Cleland, J. (2016). How effective are selection methods in medical education? A systematic review. *Medical education*, 50(1), 36–60. <https://onlinelibrary.wiley.com/doi/10.1111/medu.12817>
- Patterson, F., Prescott-Clements, L., Zibarras, L., Edwards, H., Kerrin, M., & Cousans, F. (2016). Recruiting for values in healthcare: a preliminary review of the evidence. *Advances in Health Sciences Education*, 21(4), 859–881.
- Patterson, F., Roberts, C., Hanson, M. D., Hampe, W., Eva, K., Ponnampuruma, G., . . . Cleland, J. (2018). 2018 Ottawa consensus statement: Selection and recruitment to the healthcare professions. *Medical Teacher*, 40(11), 1091–1101. doi:10.1080/0142159X.2018.1498589

- Patterson, F., & Zibarras, L. (2018). *Selection and Recruitment in the Healthcare Professions: Research, Theory and Practice*. Palgrave Macmillan.
- Patterson, F., Zibarras, L., & Ashworth, V. (2016). Situational judgement tests in medical education and training: Research, theory and practice: AMEE Guide No. 100. *Medical Teacher*, 38(1), 3-17. doi:10.3109/0142159X.2015.1072619
- Pinxten, M., Van Soom, C., Peeters, C., De Laet, T., & Langie, G. (2019). At-risk at the gate: prediction of study success of first-year science and engineering students in an open-admission university in Flanders—any incremental validity of study strategies? *European journal of psychology of education*, 34(1), 45-66.
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological bulletin*, 135(2), 322-338.
- Prideaux, D., Roberts, C., Eva, K., Centeno, A., McCrorie, P., McManus, C., . . . Wilkinson, D. (2011). Assessment for selection for the health care professions and specialty training: consensus statement and recommendations from the Ottawa 2010 Conference. *Medical Teacher*, 33(3), 215-223. doi:10.3109/0142159X.2011.551560
- Ramaekers, S., Kremer, W., Pilot, A., van Beukelen, P., & van Keulen, J. (2009). Determination of the power to decide in uncertainty with the method of Script Concordance Test. *Pedagogische Studien*, 86(3), 230-245.
- Reinbold, S. (2013). Using the ADDIE model in designing library instruction. *Medical reference services quarterly*, 32(3), 244-256. doi:10.1080/02763869.2013.806859
- Royal, K. D. (2017). Four tenets of modern validity theory for medical education assessment and evaluation. *Advances in medical education and practice*, 8, 567-570.
- Sackett, P. R., Walmsley, P. T., Koch, A. J., Beatty, A. S., & Kuncel, N. R. (2016). Predictor content matters for knowledge testing: Evidence supporting content validation. *Human Performance*, 29(1), 54-71. doi:10.1080/08959285.2015.1120307
- Sam, A. H., Field, S. M., Collares, C. F., van der Vleuten, C. P., Wass, V. J., Melville, C., . . . Meeran, K. (2018). Veryshortanswer questions: reliability, discrimination and acceptability. *Medical education*, 52(4), 447-455.
- Schmader, T., Whitehead, J., & Wysocki, V. H. (2007). A linguistic comparison of letters of recommendation for male and female chemistry and biochemistry job applicants. *Sex roles*, 57(7), 509-514.
- Schreurs, S. (2020). Selection for Medical School: the quest for validity. Universiteit Maastricht.
- Schreurs, S., Cleland, J., Muijtjens, A. M. M., oude Egbrink, M. G. A., & Cleutjens, K. (2018). Does selection pay off? A cost-benefit comparison of medical school selection and lottery systems. *Medical Education*, 52(12), 1240-1248. doi:10.1111/medu.13698
- Schreurs, S., Cleutjens, K., Collares, C. F., Cleland, J., & oude Egbrink, M. G. A. (2020). Opening the black box of selection. *Advances in Health Sciences Education*, 25(2), 363-382. doi:10.1007/s10459-019-09925-1





- Schreurs, S., Cleutjens, K. B. J. M., Cleland, J., & oude Egbrink, M. G. A. (2020). Outcome-Based Selection Into Medical School: Predicting Excellence in Multiple Competencies During the Clinical Years. *Academic Medicine*, 95(9), 1411–1420. doi:10.1097/ACM.0000000000003279
- Schreurs, S., Cleutjens, K. B. J. M., Muijtjens, A. M. M., Cleland, J., & oude Egbrink, M. G. A. (2018). Selection into medicine: the predictive validity of an outcome-based procedure. *BMC Medical Education*, 18(1), 214. doi:10.1186/s12909-018-1316-x
- Schripsema, N. R. (2017). Medical student selection: Effects of different admissions processes. Rijksuniversiteit Groningen.
- Schripsema, N. R., van Trigt, A. M., Lucieer, S. M., Wouters, A., Croiset, G., Themmen, A. P. N., . . . Cohen-Schotanus, J. (2017). Participation and selection effects of a voluntary selection process. *Advances in Health Sciences Education*, 22(2), 463–476. doi:10.1007/s10459-017-9762-5
- Schripsema, N. R., van Trigt, A. M., van der Wal, M. A., & Cohen-Schotanus, J. (2016). How Different Medical School Selection Processes Call upon Different Personality Characteristics. *PLoS One*, 11(3), e0150645. doi:ARTNe015064510.1371/journal.pone.0150645
- Steenman, S. C. (2018). *Alignment of Admission: An Exploration and Analysis of the Links Between Learning Objectives and Selective Admission to Programmes in Higher Education*. Universiteit Utrecht.
- Steenman, S. C. (2020). Hoe verdelen we de chocoladetaart? Toelating en selectie in de strategische agenda. *Th&ma Hoger Onderwijs*, 2020(1), 32–36.
- Steenman, S. C., Bakker, W. E., & van Tartwijk, J. W. F. (2016). Predicting different grades in different ways for selective admission. Disentangling the first-year grade point average. *Studies in Higher Education*, 41(8), 1408–1423.
- Stegers-Jager, K. M. (2018). Lessons learned from 15 years of non-grades-based selection for medical school. *Medical Education*, 52(1), 86–95. doi:10.1111/medu.13462
- Stegers-Jager, K. M., Themmen, A. P., Cohen-Schotanus, J., & Steyerberg, E. W. (2015). Predicting performance: relative importance of students' background and past performance. *Medical Education*, 49(9), 933–945. doi:10.1111/medu.12779
- Stegers-Jager, K. M. (2018). Lessons learned from 15 years of nongradesbased selection for medical school. *Medical education*, 52(1), 86–95.
- Urlings-Strop, L. C., Themmen, A. P., & Stegers-Jager, K. M. (2017). The relationship between extracurricular activities assessed during selection and during medical school and performance. *Advances in Health Sciences Education*, 22(2), 287–298.
- Urlings-Strop, L. C., Stijnen, T., Themmen, A.P., & Splinter, T.A. (2009). Selection of medical students: a controlled experiment. *Medical Education*, 43(2), 175–183.
- Urlings-Strop, L. C., Themmen, A. P., Stijnen, T., & Splinter T. A. (2011). Selected medical students achieve better than lottery-admitted students during clerkships. *Medical Education*, 45(10), 1032–1040.

- van der Linden, L. (2018). *Prediction of Study Success—Creation of Magic Zones*. Universiteit Utrecht.
- van der Smagt, M. J., Tick, A. & Van Ooijen, L. Eerlijke toelatingsprocedure combineert decentrale selectie en loting. *Scienceguide*. Opinie. <https://www.scienceguide.nl/2021/07/eerlijke-toelatingsprocedure-combineert-decentrale-selectie-en-loting/>
- van Strien, J., & Joosten–ten Brinke, D. (2016). Het beoordelen van de kwaliteit van rubrics. *Examens, Tijdschrift voor de toetspraktijk*, 2016(1), 33–37.
- van Vertloo, L. R., Burzette, R. G., & Danielson, J. A. (2022). Predicting Academic Difficulty in Veterinary Medicine: A Case–Control Study. *Journal of veterinary medical education*, 49(4), 524–530.
- Visser, K. (2017). Kantel de selectie in het hoger onderwijs. *Scienceguide*, 23 augustus. Opinie. <https://www.scienceguide.nl/2017/08/kantel-de-selectie-in-het-hoger-onderwijs/>
- Vulperhorst, J., Lutz, C., de Kleijn, R., & van Tartwijk, J. (2018). Disentangling the predictive validity of high school grades for academic success in university. *Assessment & Evaluation in Higher Education*, 43(3), 399–414.
- Woods, S. A., & Anderson, N. R. (2016). Toward a periodic table of personality: Mapping personality scales between the five–factor model and the circumplex model. *Journal of applied psychology*, 101(4), 582.
- Wouters, A. (2017). *Effects of medical school selection*. Vrije Universiteit Amsterdam.
- Wouters, A., Bakker, A. H., van Wijk, I. J., Croiset, G., & Kusurkar, R. A. (2014). A qualitative analysis of statements on motivation of applicants for medical school. *BMC Medical Education*, 14(1), 200. doi:10.1186/1472-6920-14-200
- Zwick, R. (2017). *Who gets in? Strategies for fair and effective college admissions*: Harvard University Press.

# Bijlagen



# Bijlage 1 • Beoordelen van toetsen

De checklist hieronder geeft een overzicht van wat standaard onderdeel is

PRIORITEIT	ELEMENT	MOGELIJKHEDEN	ANALYSE
HOOG	Ervaringen van kandidaten	(Korte) vragenlijst, bijvoorbeeld of deelnemers het idee hebben dat de beoogde vaardigheid inderdaad is getoetst.	In hoofdstuk 6, onder het kopje 'procedure', zijn enkele voorbeeldvragen opgenomen die door middel van een korte vragenlijst aan kandidaten voorgelegd kunnen worden. Als interactie met deelnemers belangrijk is, dan kunnen ook interviews of focusgroepen gebruikt worden.
	Betrouwbaarheid	Hoe consistent een toets iets meet. Dit kan door een toets twee keer af te nemen, of door de toets in tweeën te splitsen en de consistentie ertussen te berekenen.	Je kunt zelf een <i>test-heretest</i> betrouwbaarheid meten, door dezelfde toets tweemaal af te nemen en de correlatie tussen de afnames te berekenen. Nadeel hierbij is dus dat deelnemers leren van de eerste keer dat ze de toets doen, en dit dus scheve resultaten oplevert. <i>Split half betrouwbaarheid</i> is de andere mogelijkheid, dan splits je de toets in 2 delen (die het meest equivalent zijn) en bereken je de correlatie tussen beide helften. Die splitsing is echter vrij arbitrair en deze methode heeft ook veel haken en ogen.
		Als je de toetsen één keer afneemt geven statistische maten zoals de alfa of omega informatie.	<i>Cronbach's alfa</i> kun je gebruiken om de consistentie tussen vragen binnen een toets te berekenen, dit is echter een omstrepen methode, het is daarom belangrijk om bij het gebruik hiervan goed te letten op de manier waarop je het gebruikt en de interpretatie van resultaten. De beste optie, psychometrisch gezien, is de Omega om betrouwbaarheid vast te stellen, in plaats van alfa. Niet alle statistische software kan deze maat echter gemakkelijk berekenen.
		De consistentie tussen beoordelaars is belangrijk voor toetsen waar zij een grote rol spelen. In een selectiecontext is streven naar minimaal een betrouwbaarheid van 0,8, maar liever hoger voor de hand liggend.	Verder is ook de beoordelaarsbetrouwbaarheid van groot belang, zowel de interbeoordelaars- als de intrabeoordelaarsbetrouwbaarheid: in hoeverre twee verschillende beoordelaars eenzelfde item hetzelfde beoordelen, versus in hoeverre dezelfde beoordelaar hetzelfde item op een ander moment beoordeelt. Dit geldt vooral voor open vragen, maar is ook van toepassing op het beoordelen van bijvoorbeeld interviews.
	<u>Itemanalyse</u>	In een itemanalyse ga je naar de individuele items (vragen) binnen de toets kijken. Vragen die je hierbij wilt beantwoorden zijn: - Zijn er items die te moeilijk zijn of juist te makkelijk zijn? - Zijn er items die goed gemaakt worden door kandidaten die andere vragen slecht maken of die slecht gemaakt worden door kandidaten die andere vragen goed maken? - Zijn er items (meerkeuzevragen) met afleiders die te veel gekozen worden, en waarbij het mogelijk is dat het verkeerde antwoord als goed is aangemerkt?	Standaard toetsanalyse: • p, moeilijkheid van een vraag. De p-waarde geeft weer hoeveel procent van de kandidaten de vraag goed heeft gemaakt (bij een meerkeuzevraag) of wat de gemiddelde score van alle kandidaten was op een open vraag. De p-waarde ligt bij meerkeuzevragen tussen 0 en 1. Bij 0 hebben alle kandidaten de vraag fout beantwoord. Bij 1 hebben alle kandidaten de vraag goed beantwoord. Dus hoe makkelijker de vraag hoe hoger de p-waarde. In een toets is het belangrijk dat er een goede variatie in moeilijkheid is, vooral rond het afkappunt is het belangrijk om voldoende vragen te hebben zodat de toets goed onderscheid kan maken tussen kandidaten. • Rir & Rar waarden: Rir en Rar waarden zeggen bij meerkeuzevragen iets over het onderscheidend vermogen van een vraag (Rir-waarde) of een afleider (Rar-waarde). Onderscheidend vermogen wil zeggen of een vraag of een afleider een goed onderscheid maakt tussen de goed scorende en de minder goed scorende kandidaten. Een hoge Rir-waarde wil zeggen dat de betere kandidaten de vraag juist beantwoorden en de minder goede kandidaten de vraag onjuist beantwoorden. Een afleider doet zijn werk goed als de Rar-waarde van die afleider negatief is en hoe meer negatief hoe beter. Dit is een aanwijzing dat juist de minder goede kandidaten kiezen voor de afleider. Negatieve Rir-waarden zijn ongewenst en duiden erop dat juist de betere kandidaten een vraag onjuist beantwoorden en de kandidaten die verder minder goed scoren een bepaalde vraag juist goed beantwoorden. De Rir- en Rar-waarden liggen tussen -1 en +1. Bij voorkeur is de Rir-waarde groter dan 0,2, maar nog liever groter dan 0,4, en is de Rar-waarde negatief; en hoe negatiever hoe beter. • a-waarde: De a-waarde is de aantrekkelijkheid van een afleider. De a-waarde ligt tussen 0 en 1. Bij 0 kiest geen enkele kandidaat de afleider. Bij 1 kiezen alle kandidaten de afleider. Als een afleider een hele hoge a-waarde heeft, is de p-waarde dus laag. Dat is ongewoon omdat de p-waarde hoger dient te zijn dan een a-waarde.



**PRIORITEIT ELEMENT MOGELIJKHEDEN ANALYSE**

Al deze elementen moeten in relatie tot elkaar gezien worden, zie daarvoor ook deze [tabel](#) :

Level of difficulty (P-value)	Distinctive capacity (Rir-value)	
	Rir < 0,1: Question is not distinctive and with negative values even opposite to expectations: high performing candidates score low on this question.	Rir > 0,1: Acceptable to good (Rir > 0,3) distinctive question
Hard question P' < 0,3	The question has been poorly answered, also by the better performing candidates of the test. <i>Question: Is the answer key correct? Could more answers be approved? Tip: compare the alternatives to high A-values and positive Rar-values.</i>	Generally, the question has been answered incorrectly, but it still distinguishes between better and poor performing candidates on the test. <i>Is the question too hard, too complex? Is it a tricky question? Make sure there are not too many hard questions in the test.</i>
Practicable question 0,3 < P' < 0,8	The question has been answered pretty well but distinguishes unsatisfactory between poor and high performing candidates. <i>Perhaps other questions can be -partly- approved? Tip: compare the alternatives to high A-values and positive Rar-values.</i>	The question has been answered fairly good and the distinctive capacity is all right. <i>In this case there is no need for further action.</i>
Easy question P' > 0,8	The question was easy to answer for the majority of the candidates, regardless of their performance on the test. <i>Is the question an unintentional giveaway and could it be solved with common sense?</i>	The question has been answered well and distinguished between high and lower performing candidates. <i>In this case there is no need for further action.</i>

**Convergente validiteit**      **Concurrent** (gelijktijdig)  
Berekenen van correlatie met een ander element dat gemeten wordt in de selectie

Er zijn twee opties, controleren of er inderdaad een correlatie is met een variabele waarmee je samenhang verwacht (convergent), maar ook controleren of er geen correlatie is met een variabele waarmee je geen relatie verwacht (divergent).  
Controleren of er geen correlatie bestaat met een ongerelateerde variabele kan vaak binnen de procedure, bijvoorbeeld tussen kennistoets en een maat voor samenwerkingsvaardigheden verwacht je een lage correlatie.

**Voorspellend**  
Berekenen van correlatie met een ander element dat later in de opleiding gemeten wordt

In de analyse fase is uitgedacht welke delen van de procedure welke resultaten in de opleiding zouden moeten voorspellen, dit controleer je voor de toets. Ook dit kan twee kanten op: convergent of er inderdaad correlatie is tussen de toets en de uitkomst waarvan je in de analyse-fase samenhang verwachtte, bij divergent controleer je of uitkomsten in het curriculum waarmee je geen samenhang verwacht inderdaad lage correlatie vertonen met die toets.

**WANNEER VAN TOE-PASSING**      Pass/Fail

Is de pass/fail standaard wel goed gezet?

Voor selecties waarin een toets is opgenomen waar kandidaten een bepaald resultaat moeten behalen, is dit een belangrijk criterium (bijvoorbeeld een kennistoets die kandidaten moeten halen). Hierbij moet de grens natuurlijk goed liggen en passend zijn bij de opleiding.

**Structuur van de toets**

Hier controleer je of de structuur die je gepland hebt in de toets daadwerkelijk terugkomt in de data

Als je meerdere elementen meet, maar verschillende items meten steeds wel 1 enkel element (multidimensionaliteit tussen items), dan is een factoranalyse of Omega een goede methode om de structuur van de toets te evalueren. Controleer dan of de items die je bij verschillende schalen verwacht ook bij die verschillende schalen terechtkomen.  
Als je meerdere elementen meet, en de items kunnen in zichzelf ook meerdere elementen meten (multidimensionaliteit tussen én binnen items), dan is Cognitive Diagnostic Modeling de enige optie. Deze analyse geeft aan welke items misschien toch (deels) andere competenties meten.

## Bijlage 2 • Evalueren van de selectieprocedure

Deze handreiking is tot stand gekomen naar aanleiding van de learning community toelating en selectie op 2 november 2020. De learning communities worden georganiseerd door de Vereniging van Universiteiten (VSNU) in samenwerking met Centre of Expertise in Admission (UU).

Deze handreiking is geschreven door een aantal leden van het onderzoekersnetwerk toelating en selectie:

- Dr. Sebastiaan Steenman (Universiteit Utrecht)
- Dr. Kitty Cleutjens (Universiteit Maastricht)
- Dr. Sanne Schreurs (Universiteit Utrecht)
- Timon de Boer, MSc (Universiteit Utrecht)
- Gijsbert van Elven (Universiteit Utrecht)

Voor goede selectie is betekenisvolle evaluatie van groot belang. Inmiddels is er in de Nederlandse context op verschillende plekken ervaring opgedaan met die evaluatie. Die ervaring is voldoende om een aanzet te maken voor een handreiking waarin de belangrijke uitgangspunten van goede evaluatie zijn uitgewerkt. De basis van dit document ligt bij de bijeenkomst van de VSNU learning community toelating en selectie van 2 november 2020, waarin ervaringen en wetenschappelijk onderzoek rond het evalueren van selectie zijn uitgewisseld. Deze basis is uitgebreid met wetenschappelijk onderzoek naar dit thema. De voorbeelden in de tekstwolkjes zijn voorbeelden uit de learning community van 2 november.

In deze handreiking komen acht uitgangspunten van goede evaluatie van selectie naar voren. Deze worden besproken, samen met handvatten om de uitgangspunten vorm te geven en enkele korte voorbeelden vanuit de bijeenkomst van de learning community voor hoe dat nu op instellingen gebeurt. Deze acht uitgangspunten zijn bedoeld om een handreiking voor goede evaluatie van selectie te geven, en om opleidingen en instellingen die de evaluatie van selectie willen vormgeven te ondersteunen.

Het onderzoekersnetwerk selectie (ONS) kan vanuit wetenschappelijk onderzoek helpen bij specifieke vragen bij de vormgeving van evaluatie en de selectie zelf.

### **Uitgangspunt 1: heldere aansluiting instrumenten bij opleiding**

Voor de evaluatie van selectie moet eerst helder zijn wat de verschillende instrumenten die tijdens de selectie worden gebruikt beogen te voor-



spellen. Vwo-cijfers kunnen bijvoorbeeld goed gebruikt worden om iets te zeggen over de kennis- en begripstoetsen in de opleiding, terwijl andere instrumenten gebruikt worden om te kijken of de specifieke didactiek in de opleiding bij de kandidaat past. Dat resulteert in een onderbouwde theorie over hoe selectie is verbonden aan zowel de doelen van de opleiding waarvoor de selectie wordt ontwikkeld als het onderwijs en de toetsing binnen die opleiding. Dit heet alignment van selectie.

### **Select-UU**

In het Select-UU-project worden de selectieprocedures van alle 12 selecterende bacheloropleidingen aan de Universiteit Utrecht jaarlijks systematisch onderzocht en geëvalueerd vanuit het principe van alignment. In dit onderzoek worden resultaten van de verschillende selectie-instrumenten gekoppeld aan de getoetste vaardigheden in de opleiding. Op basis van de resultaten uit dit onderzoek worden de selectieprocedures waar nodig aangepast. Interventie wordt zo direct gekoppeld aan de feedback.

### **Effectiviteitsonderzoek selectietoets opleiding Psychologie (UvA)**

In dit onderzoeksproject werd aan een groep studenten gevraagd de effectiviteit van de selectietoets bij de opleiding Psychologie te onderzoeken. In dit onderzoek werd een een-op-een vergelijking gemaakt tussen cijfers voor de selectietoets en verschillende cijfers in de opleiding. Bovendien werd er in dit onderzoek onderscheid gemaakt tussen Nederlandse en internationale studenten.

Voor de evaluatie moet deze link tussen selectie en meetbare prestaties binnen (of buiten/na) de opleiding helder zijn, dat is de enige manier om te kunnen toetsen of de selectie ook daadwerkelijk valide is. Het is daarbij belangrijk dat studenten op verschillende manieren succesvol kunnen zijn. Niet elk instrument is bedoeld om hetzelfde te voorspellen. Het ligt dan ook voor de hand dat selectieprocedures meerdere instrumenten gebruiken om verschillende relevante vaardigheden van kandidaten mee te kunnen nemen. In de evaluatie moet dan ook oog zijn voor die differentiatie.

### **The quest for validity: onderzoek naar selectie voor opleiding Geneeskunde (UM)**

In het promotieonderzoek van Sanne Schreurs werd de validiteit van de selectieprocedure van de opleiding Geneeskunde onderzocht. Deze selectieprocedure is opgezet volgens het principe van alignment en volgt een strakke blauwdruk. Naast het koppelen van de resultaten van de verschillende selectie-instrumenten aan de getoetste vaardigheden in zowel de bachelor als de master is ook de interne structuur en item functioning van de diverse selectieonderdelen onderzocht.

In de evaluatie betekent dit dat de prestaties van studenten op verschillende selectie-instrumenten wordt vergeleken met verschillende prestaties binnen de opleiding. Dat betekent dat het niet enkel vergeleken wordt met het gemiddelde cijfer dat in het eerste jaar wordt behaald, of met het al dan niet uitvallen van studenten.

#### **Uitgangspunt 2: geïnformeerd door wetenschappelijke kennis**

Selectie-instrumenten behoren te zijn afgestemd op de beschikbare kennis over toelatings- en selectiebeleid. Het periodiek toetsen van de assumpties achter instrumenten aan de state-of-the-art kennis op dit gebied, zowel internationaal als specifiek in de Nederlandse context, is essentieel. Hierin kan het onderzoekersnetwerk een vraagbaakrol vervullen.

#### **Uitgangspunt 3: toetsen op bias**

Selectie-instrumenten worden getoetst op impliciete 'biases'. Specifieke groepen kandidaten worden niet onevenredig beoordeeld of benadeeld door de gebruikte selectie-instrumenten. Dit kan zowel door de instrumenten zelf tegen het licht te houden, als door prestaties van verschillende groepen op instrumenten te evalueren.

Kijk bijvoorbeeld of in instrumenten voorkennis wordt vereist die voor de ene groep directer aanwezig is dan voor de andere groep. Stel liever geen vragen waarbij het een groot voordeel oplevert wanneer ouders dezelfde opleiding gevolgd hebben of in het werkveld van de opleiding werken. Ook voorbeelden die consequent uitgaan van een bepaalde achtergrond kunnen bias veroorzaken. Evalueer ook of de beoordeling van het werk van kandidaten zo objectief mogelijk gebeurt. Veel individuele interpretatieruimte introduceert snel bias.

Een ongestructureerde beoordeling door 1 persoon van een motivatiebrief van kandidaten, of het gebruiken van voor een deel van de kandidaten onbekende of onrealistische casus kan bijvoorbeeld zorgen voor dergelijke biases.



#### **Uitgangspunt 4: de organisatie van de selectie**

Aansluitend op het vorige punt is de organisatie rondom de selectie onderwerp van evaluatie. Daarbij speelt mee dat de procedure geen ruimte moet geven voor persoonlijke voorkeuren, scores zo objectief mogelijk zijn, meerdere mensen betrokken zijn en een selectiecommissie breed is samengesteld. Studentbetrokkenheid bij de vormgeving van selectie heeft toegevoegde waarde.

#### **Uitgangspunt 5: regels en richtlijnen**

In de organisatie van de selectieprocedure wordt blijvend geëvalueerd of het beleid voldoet aan de regelgeving en de richtlijnen die er zijn rondom toelatings- en selectiebeleid. Hierbij is ook een goede verantwoording voor alle keuzes die er in de selectieprocedure worden gemaakt steeds beschikbaar. Daarbij is ook aandacht nodig voor data-beveiliging en privacy van de kandidaten en voor de inspraak van medezeggenschap.

#### **Toepassing Kader Toelating Master Radboud Universiteit**

Aan de Radboud Universiteit is een toepassing van het kader toelating master ontwikkeld. Dit bestaat uit een handleiding met bijbehorende checklist voor het inrichten, structureren en evalueren van selectieprocedures en helpt bovendien bij het aanwakkeren van dialoog over de selectieprocedure.

#### **Uitgangspunt 6: effectiviteit**

In de evaluatie van een selectieprocedure is het ook van belang om te evalueren of de selectie-instrumenten niet te veel dezelfde vaardigheden, eigenschappen of kenmerken van kandidaten bevragen. Het is een belangrijk criterium om de selectieprocedure effectief vorm te geven, ook om kandidaten en opleidingen niet onnodig veel te belasten. Hierbij hoort de vraag of meerdere selectie-instrumenten bedoeld zijn om dezelfde vaardigheden van kandidaten te meten, terwijl minder instrumenten dat ook voldoende betrouwbaar doen (zie ook het eerste punt). Daarnaast hoort hierbij de vraag of er valide alternatieven zijn die minder tijd van kandidaten of van de opleiding vragen.

Kijk goed of het toegevoegde waarde heeft om bepaalde selectiemethoden (naast elkaar) te gebruiken.

Meerdere toetsen gebruiken is goed voor de betrouwbaarheid. Vwo-cijfers bestaan al uit een groot aantal toetsen. Als vwo-cijfers en de kennistoets dezelfde functie hebben in de selectie, en de kennistoets gaat niet ook bijvoorbeeld over een specifieke vorm van informatieverwerking die aansluit bij de opleiding, dan is het goed om te kijken of het gebruik van beide voldoende toegevoegde waarde heeft.

**Uitgangspunt 7: kosten en baten**

In de evaluatie van een selectieprocedure is het van belang om naast het meenemen van de kosten, ook de baten mee te nemen. Denk dus zowel aan de kosten voor de organisatie van de selectie, als aan de opbrengsten van een goede aansluiting tussen student en opleiding.

**Kosten-baten analyse**

Bij de opleiding geneeskunde van de Universiteit Maastricht zijn relevante kosten en baten meegenomen in een evaluatie van de kosteneffectiviteit van de selectieprocedure. Naast de directe kosten van de selectie, is daarbij ook het voorkomen van uitval en studievertraging meegenomen. Na het vergelijken van de kosten en opbrengsten gedurende de gehele bacheloropleiding blijkt dat de uitgebreide en relatief dure selectieprocedure kosteneffectief is ten opzichte van de voor de opleiding “gratis” lotingsprocedure.

**Uitgangspunt 8: creativiteit en vernieuwing**

Het toetsen van de efficiëntie en validiteit van instrumenten en assumpties vraagt om enige creativiteit en vernieuwende evaluatiemethodes. Daarvoor kan het nodig zijn om onderzoekers vanuit de instelling te betrekken of te informeren bij het onderzoekersnetwerk. Wanneer het ook met creatieve, geavanceerde oplossingen niet lukt om instrumenten en assumpties te toetsen, ook over meerdere cohorten, moet dat in principe beteken dat die niet langer mee worden genomen in de selectieprocedure.

**Mastertoelating Geowetenschappen**

Bij de mastertoelating bij de faculteit Geowetenschappen aan de Universiteit Utrecht wordt in de evaluatie bijvoorbeeld gewerkt met automatische tekstanalyse van motivatiebrieven en worden de effecten van de cognitieve afstand tussen bacheloropleiding en masteropleiding onderzocht.

## De ontwerpprincipes van deze handreiking

De acht uitgangspunten in deze handreiking hangen sterk met elkaar samen. Ze zijn breed ingestoken, met niet alleen oog voor de selectie-instrumenten en criteria, maar ook voor het selectieproces en de tijdsinvestering van zowel kandidaten als opleiding. In alle gevallen vragen zij om een heldere link tussen evaluatie en gevolgen daarvan voor de toekomst van de selectiepraktijk. Het zijn algemene uitgangspunten die voor alle opleidingen gelden, met behoud van ruimte in hoe de selectie eruitziet vanwege de specifieke eigenschappen van de opleiding. Wanneer immers voor een andere opleiding wordt geselecteerd zullen ook andere vaardigheden nodig zijn. Selectie is ondersteunend om de juiste student op de juiste plek te krijgen. Goede evaluatie versterkt dit uitgangspunt en deze handreiking helpt daarbij.

Een brede selectie heeft waarde omdat academische opleidingen verschillende vaardigheden van studenten vragen, maar alleen die vaardigheden die valide en betrouwbaar gemeten worden hebben een plaats in een selectieprocedure. Als laatste is de handreiking uitgesproken in het belang van zo goed mogelijke, door beschikbaar onderzoek en evaluatie geïnformeerde selectie om bias tegen te gaan.

Bewijs voor	Wat?	Hoe?
Doel	De reden voor selecteren	Leg uit waarom selectie nodig is. Welke doelen zijn bij de selectie gesteld? – Dit komt uit de analysefase (Hoofdstuk 3)
	De te beoordelen kennis, vaardigheden en attitudes	Voer een analyse uit wat kandidaten al bij aanvang van de opleiding kunnen of weten of zouden moeten kunnen of weten en zoek ondersteuning uit de literatuur over de te meten kennis, vaardigheden en attitudes – Dit komt uit de analysefase (Hoofdstuk 3)
	De interpretaties van en beslissingen op basis van de scores	Leg uit hoe de resultaten van de meting moeten worden geïnterpreteerd (wanneer is een score goed, hoe worden die gecombineerd en wat zegt dat over je kandidaten) Leg uit hoe ‘high-stakes’ de beslissing is die op de meting moet worden gebaseerd – voor een selectie is dit al relatief vroeg duidelijk: het is relatief belangrijk voor de toekomst van de kandidaten, hetgeen betekent dat er hoge eisen gesteld dienen te worden aan de betrouwbaarheid en validiteit van de meting.
Inhoud	Een blauwdruk die de link legt tussen het testdomein en echte prestaties	Leg de logische, theoretische en/of empirische relaties uit tussen de items, (sub)schalen en het te meten domein – Dit komt uit de analyse- en design fases (Hoofdstuk 3 en 4)
	Representativiteit	Bepaal de relevantie en authenticiteit van alle items voor het domein en de doelgroep: hebben stakeholders het gevoel dat de meting inderdaad een meting is van de vaardigheid die je ook wilde meten?
	Hoe goed wordt het domein bestreken	Zijn alle relevante delen van het domein (zoals gedefinieerd in het doel) voldoende vertegenwoordigd?
	Kwaliteit van instrumenten	Was er voldoende expertise beschikbaar om de instrumenten goed te ontwikkelen? Waren de onderdelen van de instrumenten bijvoorbeeld: begrijpelijk, relevant, aanvaardbaar, haalbaar en volledig?
Reacties	Scoringmodellen	Leg uit hoe de onderdelen van de instrumenten werken en hoe ze gescoord moeten worden (bijvoorbeeld een ranking vraag in een toets, wordt die in delen beoordeeld of in één geheel), met een redenering daarachter (waarom worden items op een bepaalde manier gescoord). Verder van belang voor dit onderwerp is de constructie van scorecriteria en rubrics: hoe wordt bepaald hoeveel deelnemers punten krijgen voor verschillende onderdelen? Past dat bij de inhoud van de onderdelen en bij het gebruikte antwoordformat? Hiervoor is het van belang dat voor bijvoorbeeld open vragen een duidelijke antwoordsleutel gemaakt is. Na afname van de selectie kunnen verschillende scoringmodellen ook vergeleken worden en berekend worden wat het effect zou zijn op uitkomsten van de procedure.
	Samenvatten en wegen van de scores	Rationale en procedure voor het combineren van scores, bijvoorbeeld het omrekenen naar z-scores en daar het gemiddelde van nemen. Welke onderdelen worden gecombineerd voorafgaande aan het berekenen van een z-score? Daarnaast horen hierbij de overwegingen op welke moment in het proces de verschillende scores samengebracht worden en via welke procedure. Tenslotte is er de weging, die zou in de blauwdruk al in grote lijnen duidelijk moeten zijn (hoofdstuk 3 en 4), maar bij het berekenen van de scores zijn er vaak nog beslissingen te nemen, bijvoorbeeld of items ten opzichte van elkaar ook gewogen moeten worden, of dat alleen de criteria ten opzichte van elkaar gewogen moeten worden.

Bewijs voor	Wat?	Hoe?
Reacties	Corresponderen het item en de antwoordmogelijkheden?	Zie ook hoofdstuk 5, onder antwoordformats. Het gaat er hier om dat het observatieformaat (bijvoorbeeld de rubrics) of type antwoordformat logisch aansluit bij wat er gemeten wordt (bijvoorbeeld voor een redenering gebruik je meestal een open vraag, terwijl feitenkennis meestal makkelijk in een meerkeuzevraag gesteld wordt). Het belangrijkste is om je af te vragen wat je wilt dat de kandidaten doen: hoe simpeler dat is, hoe simpeler ook het antwoordformat kan blijven. Daarbij geldt echter dat het gebruik van verschillende formats een meer betrouwbaar en valide beeld geeft, net zoals het gebruik van meerdere soorten toetsen dat doet op het niveau van de gehele procedure. Om dit verder te onderzoeken kun je ook evaluatievragen gebruiken die je aan de kandidaten stelt (zie ook hoofdstuk 6 onder procedure, en in bijlage 2).
	Kwaliteitscontrole	Hieronder vallen alle processen die je in acht neemt om de kwaliteit van de selectie te waarborgen. De evaluatie in zijn algemeenheid is hier een belangrijk element in, maar ook wie hierbij betrokken is, in hoeverre de evaluatie inderdaad meegenomen wordt in het verbeteren van de selectie voor latere jaren, etc. Let hierbij ook op de elementen uit bijlage 1.
	Dataveiligheid	Documentatie van procedures rond het waarborgen van privacy (bijvoorbeeld hoe je hierover communiceert naar kandidaten, waar informatie opgeslagen wordt, of en hoe deze gedeeld wordt met collega's).
Interne structuur	Betrouwbaarheid	Testbetrouwbaarheid (consistentie binnen het instrument en in herhaling van de instrumenten): zie bijlage 1 – dit blijft natuurlijk ook in de procedure belangrijk.
	Item functioneren	Verken items met speciale aandacht voor itemdistributies, variatie, moeilijkheidsgraad en discriminatie, en vloer- en plafondefecten; zie bijlage 1. Verder is het in de gehele procedure ook nodig het itembelang vast te stellen door te kijken naar de inhoud: wat zijn de meningen van de experts en wat is de theoretische waarde van de verschillende items? Functioneren de belangrijke items wel ook voldoende?
	Fairness / eerlijkheid	Vooraf: neem de items door en controleer ze op culturele bias, een expert hiervoor inroepen kan daarbij veel toegevoegde waarde hebben. Evaluatie: is de controleprocedure goed gelopen? Controleer de resultaten op Differential Item Functioning (DIF); voer subgroepanalyses uit om te controleren op onverklaarbare verschillen tussen groepen, en controleer of over de hele procedure heen verschillende groepen dezelfde mogelijkheden krijgen.
	Psychometrische modellen	Dimensionaliteit moet geëvalueerd worden als de procedure afgelopen is – dit is een analyse die binnen een toets gebeurt, en is beschreven in bijlage 1. Ook als er geen toets gebruikt wordt kun je goed op dimensionaliteit letten, maar dan zit dit vaak in de structuur van het instrument. Voor een interview zijn de dimensies voornamelijk al vooraf bepaald in het beoordelingsformulier, hetzelfde geldt voor referenten/aanbevelingsbrieven etc.
Relatie met andere variabelen	Gelijktijdige en/of voorspellende en convergente en/of divergente maatregelen, correlatiepatronen met deze maatregelen (zoals in bijlage 1)	Ga na of metingen die hetzelfde criterium (dezelfde vaardigheid) zouden moeten meten inderdaad positief correleren met elkaar en/of metingen die een ander criterium moeten meten inderdaad niet correleren met elkaar. Dit kan zowel gaan om metingen die op hetzelfde moment plaatsvonden, als met metingen op enig moment in de toekomst.

Bewijs voor	Wat?	Hoe?
Consequenties	Bedoelde en onbedoelde, positieve en negatieve, gevolgen op korte en lange termijn	Een overzicht van de impact die de testcores/-resultaten en de daarop gebaseerde beslissingen hebben op elk van de belanghebbenden die bij het proces betrokken zijn (voornamelijk de kandidaten, maar ook de huidige studenten, de universiteit, etc.); In hoeverre wegen de positieve gevolgen op tegen de negatieve gevolgen? Dit kan ook een afweging van de financiële kosten en baten (zie bijvoorbeeld Schreurs et al, 2018). De belangrijkste bron van informatie hiervoor zijn de betrokkenen zelf (dus voornamelijk de kandidaten). Vragenlijsten kunnen hiervoor informatie geven (zie ook hoofdstuk 6).
	Pass/Fail standaard, bij selectie gaat dit om de verdedigbaarheid van de grens van afwijzen en toelaten	De kernvraag hier is of de grens tussen toelaten en afwijzen voldoende scherp is. Om dit te onderzoeken kun je kandidaten rondom de toelatingsgrens vergelijken (de laatste kandidaten die direct zijn toegelaten in vergelijking met een vergelijkbaar aantal kandidaten die via de wachtlijst binnen zijn gekomen). Een andere optie is om deskundigen de definitieve beslissingen over kandidaten te laten vaststellen, dit werkt voornamelijk goed voor kleine programma's waar dezelfde personen alle kandidaten zien en kunnen vergelijken, of selecties waar veel informatie verzameld wordt die samen geïnterpreteerd kan worden en vergeleken kan worden met het daadwerkelijke resultaat.
	Implicaties van het niet behalen van de toets	Wat gebeurt er als een aanvrager/kandidaat als deze niet succesvol door de procedure komt en is dit een redelijk resultaat? Hoe meer implicaties een meting heeft, hoe meer ondersteuning er moet worden verzameld om de constructvaliditeit te ondersteunen.
	Vals positieve / negatieve resultaten	Hoe vaak vals positieve en vals negatieve uitkomsten voorkomen (waarbij we rekening houden met het feit dat vals-negatieve uitkomsten een onvermijdbare consequentie van selectie zijn). Welke kandidaten zijn toegelaten die eigenlijk beter niet toegelaten hadden kunnen worden, en is er een specifieke reden waardoor deze kandidaten zijn toegelaten (goede prestaties op een specifiek onderdeel)? Vals negatieve uitkomsten zijn moeilijk vast te stellen omdat zij niet toegelaten worden en omdat ze bij een numerus fixus onvermijdelijk zijn (zie ook hoofdstuk 3).

Naar: Schreurs (2020)

## Bijlage 3 • STARR

De STARR methoden is bedoeld om aan de hand van vragen over een specifieke situatie kandidaten te bevragen. Het helpt daarbij met het systematisch bevragen van de ervaringen van kandidaten, en helpt ook om concreet te maken waar het over gaat. Bij de STARR methode ligt de nadruk op een specifieke situatie, die in de eerste stap door de kandidaat wordt verhelderd. De volgende stappen geven inzicht in hoe de kandidaat handelt en reflecteert in dergelijke situaties. Het is belangrijk om vooraf helder te formuleren naar welk type situatie je zoekt (bijvoorbeeld een situatie waarin er een conflict ontstond in een groepssamenwerking). Hieronder is elk van de stappen aan de hand van een aantal vragen kort uitgewerkt:

### 1 Situatie – Beschrijf de situatie

- Wat was de situatie?
- Wanneer en waar vond de situatie plaats?
- Wat gebeurde er?
- Wie was er betrokken?

### 2 Taak – Beschrijf wat je rol of taak was

- Wat was jouw verantwoordelijkheid in deze situatie?
- Wat wilde je bereiken?
- Wat werd er van jou verwacht?
- Wat verwachtte je van jezelf in deze situatie?
- Wat was je persoonlijke doel?

### 3 Acties – Loop stap voor stap door wat je deed

- Wat was je aanpak?
- Wat deed of zei je?
- Wat gebeurde er daarna?
- Hoe reageerden anderen richting jou?

### 4 Resultaten – Beschrijf de resultaten

- Wat was het resultaat van je acties?
- Werkte het? En waarom wel of niet?
- Wat was de uiteindelijke uitkomst?

### 5 Reflectie – Reflecteer op je ervaring

- Hoe kijk je terug op je eigen acties?
- Was je tevreden met de uitkomst?
- Wat heb je er van geleerd?
- Wat zou je een volgende keer anders aanpakken?
- Kun je wat je hebt geleerd ook in een andere type situatie toepassen?





## Verklarende woordenlijst

- Alignment** De afstemming tussen de procedure, criteria en instrumenten die voor selectie worden gebruikt en de inhoud en vorm van de opleiding, zie hoofdstuk 2.
- Bias** Het onbedoeld benadelen van kandidaten met bepaalde achtergrondkenmerken. In dit handboek vaak inclusief het gebruik van instrumenten waarvan duidelijk is dat bepaalde kandidaten daar structureel lager op scoren (adverse impact), zie hoofdstuk 3 en hoofdstuk 7.
- Criteria** De vaardigheden die in een selectieprocedure gemeten worden, waarop kandidaten worden beoordeeld, zie hoofdstuk 3.
- Domeinen van vaardigheden** Op basis van de taxonomie van onderwijsdoelen van Bloom (1956) de verschillende gebieden waarop doelen kunnen worden gesteld: het cognitieve domein, het affectieve domein en het psychomotor domein, zie hoofdstuk 2.
- Harde selectie** Selectie zonder dat er een bepaald maximum aan het aantal plaatsen dat beschikbaar is is gesteld. Afwijzing van een kandidaat betekent in die context een harde uitspraak over passendheid bij de opleiding en niet een uitspraak over dat andere kandidaten beter passen, zie hoofdstuk 3.
- HOCS (Higher order cognitive skills)** Hogere orde cognitieve vaardigheden, vaardigheden in het cognitieve domein, in het bijzonder op het gebied van analyseren, evalueren en creëren, zie hoofdstuk 2.
- Instrumenten** Manieren waarop binnen een selectieprocedure bepaald wordt in hoeverre kandidaten bepaalde vaardigheden bezitten, aan bepaalde criteria voldoen, zie deel 2 van het handboek.
- Kwalitatieve criteria** De term die de wet gebruikt om aan te geven dat criteria moeten gaan over kwaliteiten van de kandidaat, zie hoofdstuk 3.
- LOCS (Lower order cognitive skills)** Lagere orde cognitieve vaardigheden, vaardigheden in het cognitieve domein, in het bijzonder op het gebied van kennen en begrijpen, zie hoofdstuk 2.
- Numerus fixus** Een besluit dat het totaal aantal kandidaten dat tot een opleiding toegelaten kan worden beperkt is tot het in dat besluit bepaalde aantal, zie hoofdstuk 3.
- SES** Sociaal-economische status, wordt ook wel omschreven als iemands maatschappelijke positie, indicatoren die vaak worden gebruikt zijn opleidingsniveau, inkomensniveau en beroepsniveau.
- UNL** Universiteiten van Nederland, de organisatie waarin de veertien Nederlandse universiteiten samenwerken.



## Over de auteurs

**Sanne Schreurs** werkt als postdoctoraal onderzoeker voor het project Select-UU aan de Universiteit Utrecht en als universitair docent bij het departement Onderwijsonderzoek en ontwikkeling van de faculteit Health, Medicine and Life Science aan de Universiteit Maastricht. Ze promoveerde in Maastricht op een onderzoek naar selectie voor geneeskundeopleidingen en is voorzitter van de selectiecommissie van de opleiding Biomedische wetenschappen. Ze houdt zich in onderzoek en praktisch bezig met toetsing.

**Sebastiaan Steenman** is universitair hoofddocent en onderwijsdirecteur bij het departement Bestuurs- en Organiseringswetenschap van de Universiteit Utrecht en lid van de Expertgroep Toegankelijkheid Hoger Onderwijs (ETHO). Zijn onderzoek richt zich op de alignment van selectie en de ontwikkeling van (onder andere cognitieve) vaardigheden. Hij coördineert het OnderzoekersNetwerk Selectie en is (co)projectleider van het project Select-UU.

**Ada Kool** is onderwijskundig adviseur en trainer bij de afdeling Onderwijsadvies & Training van de Universiteit Utrecht. Ze is gepromoveerd op onderzoek naar studentkwaliteiten die samenhangen met studiesucces en opbrengsten van honoursprogramma's en doet onderzoek naar selectie in het hoger onderwijs. Ze is (co)projectleider van het project Select-UU en is daarnaast betrokken bij de academische lerarenopleiding voor het voortgezet onderwijs.

**Karen Stegers-Jager** is universitair docent bij het Erasmus Medisch Centrum. Haar onderzoek richt zich op toetsing en selectie in het medisch onderwijs, met een sterke focus op rechtvaardigheid en diversiteit. Ze is verantwoordelijk voor de selectie voor de opleiding geneeskunde van het Erasmus MC, coördinator van het OnderzoekersNetwerk Selectie en lid van de Expertgroep Toegankelijkheid Hoger Onderwijs (ETHO).

**Kitty Cleutjens** is universitair hoofddocent bij de afdeling pathologie van de faculteit Health, Medicine and Life Sciences van de Universiteit Maastricht. Ze is oud-voorzitter van de selectiecommissie en voorzitter van de examencommissie van de opleiding geneeskunde van de Universiteit Maastricht en doet onderzoek naar selectie voor geneeskundeopleidingen. Ze is coördinator van het OnderzoekersNetwerk Selectie.

Een goede selectieprocedure is van groot belang voor eerlijke toegang tot het hoger onderwijs. Het vormgeven van een goede selectieprocedure is echter geen gemakkelijke taak. Het doel van dit handboek is om elke stap in het vormgeven van een selectieprocedure te voorzien van handvatten, op basis van onderzoek en ervaring. Dit boek richt zich daarbij primair op de mensen die bij het ontwerpen van de selectieprocedure betrokken zijn op opleidings- en instellingsniveau. Daarbij richt het zich op selectie in de breedte van het hoger onderwijs, zowel bachelor als master, en zowel hbo als wo.

Het handboek bestaat uit twee delen. In het eerste deel wordt aan de hand van een procesmodel het ontwikkelen en evalueren van een selectieprocedure besproken. Het tweede deel gaat over de verschillende mogelijke methoden of selectie-instrumenten. Dit deel gaat in op belangrijke afwegingen bij die instrumenten, voorbeelden uit de praktijk en de potentiële nadelige effecten voor bepaalde groepen kandidaten.

