



Uitgebreide Analyse aan de hand van de Kwaliteitscyclus Toetsing

Eva Klop

HRM expert, coach, facilitator

Hotel Management School Leeuwarden

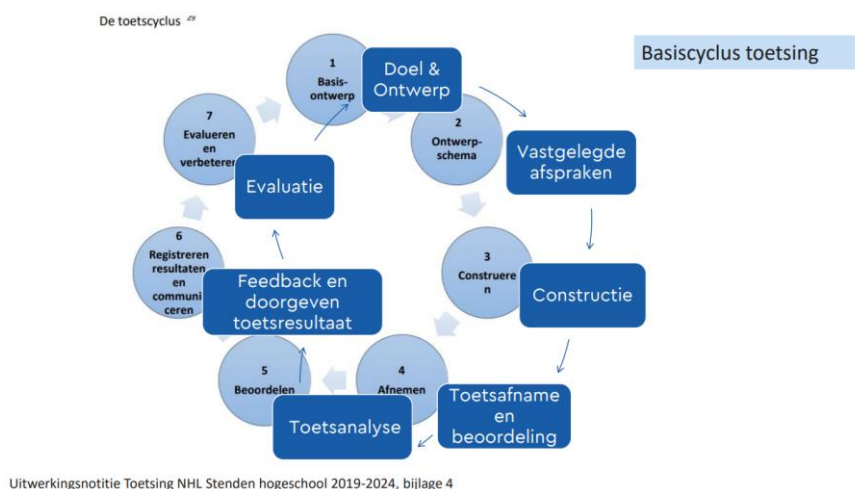
Inhoudsopgave

Inleiding.....	3
De betreffende toets	3
Fase 1: Doel en ontwerp.....	4
Het toetsbeleid van de HMSL en de functie van de betreffende toets	4
Programmatisch toetsen	4
De functie van de toets.....	4
Plaats in het toetsprogramma en de eindkwalificatie	5
Evaluatie van de betreffende toets	6
Conclusie	6
Fase 2: vastgelegde afspraken	7
Aantal kansen	7
Toetsinhoud en niveau	7
Afspraken m.b.t. beoordeling	8
Fase 3: Constructie	8
De gekozen toetsvorm	8
De opdrachtomschrijving.....	9
Richtlijnen voor nakijken en feedback	9
Rubric	10
Collegiale screening/ kalibratie	10
Fase 4: Toetsafname en beoordeling	10
Locatie en omstandigheden	10
Objectiviteit?.....	11
Beoordelaarseffecten	11
Plagiaatchecker.....	11
Fase 5: Toetsanalyse	11
Kwalitatieve analyse:	12
De vier kwaliteitscriteria	12
Validiteit.....	12
Betrouwbaarheid	12
Transparantie:	13
Efficiëntie	14
Kwantitatieve analyse	14
Fase 6: Feedback en doorgeven toetsresultaat.....	16
Fase 7: Evalueren	16
Conclusie en geformuleerde verbeteracties.....	16
Bijlages	18

Inleiding

Een belangrijk onderdeel van het docentschap is het competent zijn in toetsen en beoordelen. Toetsen en beoordelen is een breed begrip. Dat wordt wel duidelijk als je je verdiept in het onderwerp; er is een scala aan wetenschappelijk onderzoek te vinden en alleen al binnen de NHL Stenden Hogeschool zijn talloze documenten te vinden die gaan over toetsen en beoordelen. Waar te beginnen als startend docent met weinig know-how over het onderwerp?

Binnen atelier Beoordelen vonden we onze houvast in de *Kwaliteitscyclus Toetsing*. Het doorlopen van de toetscyclus, ook wel basiscyclus Toetsing genoemd, helpt als je toetsen op een goede en kwalitatieve manier wil ontwikkelen, afnemen en beoordelen. Het helpt je als startende docent ook het bestaande beleid te ontdekken én langzamerhand een eigen visie te vormen. Want het doorlopen van de cyclus vraagt om verdieping, spiegeling en een kritische houding. De cyclus kent verschillende fasen, of stappen, die je doorloopt om de kwaliteit van toetsing te onderzoeken, acties te formuleren en zo de kwaliteit te waarborgen:



De toetscyclus is in eerste instantie bedoeld om te gebruiken als je nog helemaal geen bestaande toets hebt en deze wil gaan ontwerpen. Hij is echter ook bruikbaar als je kritisch naar een (eigen) bestaande toets wil kijken en wil beoordelen of deze in al zijn facetten in orde is. Ik heb de toetscyclus helemaal doorlopen en toegepast op één van de toetsen waar ik in de praktijk mee werk. Middels het doorlopen van de verschillende fasen van de cyclus, de toepassing hiervan op een (eigen) toets, en de uitwerking ervan in dit document, heb ik de kennis en kunde die ik heb opgedaan in – en rondom – het atelier Beoordelen in de praktijk gebracht. Het leidde tevens tot het formuleren van aanbevelingen en het concreet uitwerken van één van die aanbevelingen (verbetering rubrics). De aanbevelingen zijn te vinden onder *Geformuleerde verbeteracties*, het laatste hoofdstuk van dit document. De concrete uitwerking *verbetering rubrics* is apart opgenomen in mijn portfolio.

De betreffende toets

Het onderwerp van deze analyse is een schrijfoopdracht¹ over Teamdynamiek, een 'professional assignment' voor tweedejaars studenten van de Hotel Management School Leeuwarden (HMSL). Deze opdracht, ontwikkeld door mijn collega Mariët Hiemstra, maakt deel uit van het vak Human Resource Management en wordt in groepsverband (DC groepjes) uitgevoerd. Binnen de HMSL werken we in Learning Communities (LC's), waarin studenten en docenten zijn ingedeeld in vier vaste groepen. Dit bevordert een vertrouwde leeromgeving. Elke LC heeft een eigen HRM expert. Als HRM-expert in LC2 geef

¹ *De termen 'toets, opdracht, assignment' verwijzen in dit document allemaal naar deze specifieke toets. Meestal zal ik het hebben over 'de opdracht.'

ik de HRM-lessen en beoordeel ik de bijbehorende opdrachten.

Ik heb deze toets gekozen omdat ik tijdens het nakijken van de opdracht in mijn eerste jaar als docent regelmatig tegen vragen en onzekerheden aanliep. Zo merkten meerdere studenten op dat een onderdeel van de opdrachtoomschrijving niet duidelijk was. Daarnaast miste ik zelf houvast tijdens het nakijken; er leek veel ruimte te zijn voor eigen interpretatie en invulling. Was dat de bedoeling? Hoe gingen mijn collega's hiermee om? En zou dit LC-overstijgend wel leiden tot eerlijke beoordelingen? In andere woorden: is er wel sprake van kwalitatieve toetsing en beoordeling bij dit vak?

Fase 1: Doel en ontwerp

In deze fase kijk je naar de kaders waarbinnen je een toets moet construeren (Hogeschool Utrecht, z.d.-c): Kijk je naar het toetsbeleid waarbij je toets moet aansluiten. Bepaal je de functie van de toets. Bepaal je op welke eindkwalificaties je toets betrekking heeft. Kijk je naar de plaats van de toets in het toetsprogramma, waarbij je bepaalt van welk niveau (o.a. in mate van complexiteit en zelfstandigheid) er sprake zou moeten zijn.

De opdracht is een bestaande toets en hoeft niet opnieuw ontworpen te worden. Toch is het waardevol om ook bij een bestaande toets de bovengenoemde aspecten te bekijken. Dit biedt inzicht in het doel van de toets in relatie tot andere toetsen in het programma, de kwalificaties waarop de toets zich richt en de aansluiting van de toets bij het beleid van de opleiding.

Het toetsbeleid van de HMSL en de functie van de betreffende toets

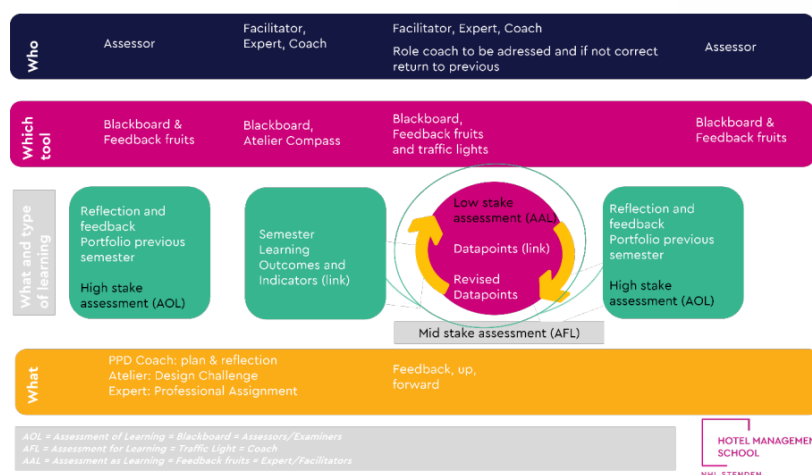
Programmatisch toetsen

Binnen de HMSL is het toetsingsframework vastgelegd in het *Assessment Work Plan Bachelor HMS 2023-2024*. Dit document beschrijft wat, waarom en hoe er binnen de bacheloropleiding wordt getoetst. Daarnaast legt het uit hoe het NHL Stenden-brede beleid van 'programmatisch toetsen' binnen de HMSL is vertaald.

Een belangrijk uitgangspunt is dat toetsen niet langer slechts een 'afvinkmoment' is, maar een middel om studenten te ondersteunen bij het structureren, monitoren en bijsturen van hun leerproces. Bij programmatisch toetsen toont de student bekwaamheid aan door bewijs te verzamelen in een portfolio. De voortgang wordt gemonitord via verschillende bewijsstukken (NHL Stenden, z.d.-a).

De functie van de toets

In het *Work Plan* wordt duidelijk dat HMSL bij programmatisch toetsen gebruikmaakt van low, mid en high-stake assessments. Een low-stake assessment, zoals een presentatie, diagnostische toets of professional assignment, leidt niet tot een pass/fail-beslissing of ECTS-toekenning. Dit gebeurt pas bij een high-stake assessment, waarbij de student een portfolio inlevert met alle datapunten en reflecties, gevolgd door een portfolio-interview. Dit proces is schematisch weergegeven in het volgende figuur (overgenomen uit het *Assessment Work Plan*):



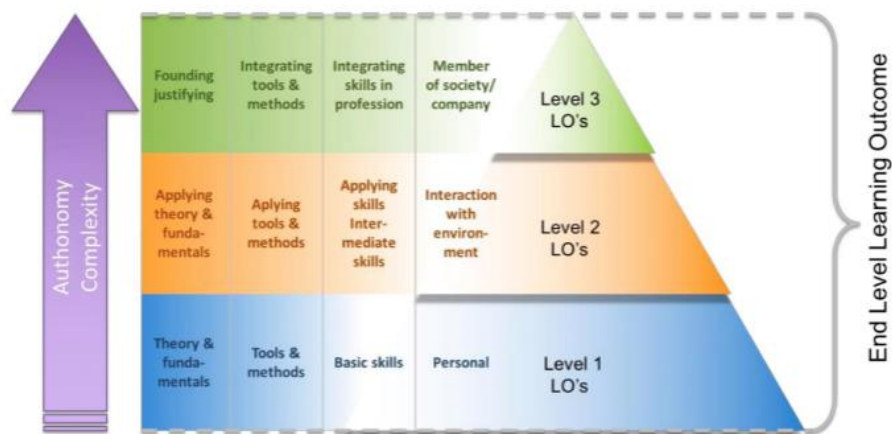
Low-stake assessments komen regelmatig voor binnen een semester en hebben een summatieve, ontwikkelgerichte functie. Ze bieden inzicht in de voortgang van de student ten opzichte van een leeruitkomst. Een leeruitkomst is “de uitkomst van het leren, oftewel datgene wat studenten dienen aan te tonen aan het einde van een leerperiode” (Peeters, 2024).

In de praktijk ontvangen studenten voor deze toetsen (binnen de HMSL datapunten genoemd) feedback van de docent en soms een score (bijvoorbeeld een aantal punten op een schaal van 100) als indicatie van de kwaliteit van het product. Dit product, samen met de feedback en eventuele score, kan door de student als datapunt (bewijslast) worden opgenomen in het portfolio, bij de betreffende leeruitkomst. De high-stake assessments, de portfolio-interviews, vinden in principe één keer per semester plaats. Studenten die hun portfolio-interview niet behalen, hebben de mogelijkheid om in het volgende semester een herkansing te doen.

Plaats in het toetsprogramma en de eindkwalificatie

De opdracht is een low-stake assessment met een formatieve functie. Het *Work Plan* stelt dat een professional assignment moet kunnen worden gekoppeld aan een leeruitkomst. Dit is logisch, aangezien het einddoel van het semester het behalen van de semester-leeruitkomsten is. Studenten werken hier continu naartoe via low-stake assessments. De feedback die zij ontvangen helpt hen om te bepalen waar ze staan in hun leerproces.

De indicatoren in een opdracht moeten aansluiten bij het expertisegebied én de bijbehorende leeruitkomst. Bij de HMSL zijn de leeruitkomsten verdeeld over drie niveaus: Level 1, Level 2 en Level 3. In het *Compass van de Main Phase*, een document waarin studenten essentiële informatie over studie en toetsing vinden, is dit schematisch weergegeven:



Noot. Figuur overgenomen uit “Main Phase Compass 2023- 2024.” Door HMSL.

Een leeruitkomst wordt per level opgebouwd met toenemende eisen aan autonomie en complexiteit, waarbij de taxonomie van Bloom duidelijk herkenbaar is (HMSL, 2023). Zo vraagt een leeruitkomst op Level 2 om meer complexiteit en zelfstandigheid dan op Level 1. Dit betekent dat een Level 2-toets méér moet toetsen dan een Level 1-toets.

Binnen het beleid van de HMSL is vastgesteld dat toetsen op Level 1 zich richten op het onthouden en begrijpen van kennis, terwijl toetsen op Level 2 de nadruk leggen op het toepassen van deze kennis. Dit benadrukt hoe een toets is ingebed in het (toets)programma en welke bijdrage deze levert aan de leerontwikkeling en voortgang van de student.

Samengevat moet bij het formuleren van het doel en ontwerp van een Level 2-toets binnen de bacheloropleiding van de HMSL rekening worden gehouden met de volgende aspecten:

Uitgangspunten doel en ontwerp van toetsen op Level 2	De indicatoren van de toets moeten gerelateerd zijn aan de indicatoren van de Leeruitkomst die bij het niveau en het expertisegebied past.
---	--

	De toets moet studenten uitnodigen hun opgedane kennis en vaardigheden daadwerkelijk toe te passen.
--	--

Evaluatie van de betreffende toets

Een evaluatie van de syllabus en opdrachtbeschrijving toont aan dat de ontwerpers hebben voldaan aan de hierboven beschreven eisen. De opdracht is bedoeld voor studenten die zich op Level 2 van hun studieloopbaan bevinden. De kwalificatie waarop de opdracht betrekking heeft is Leeruitkomst 6.2 van de bacheloropleiding, die als volgt luidt:

Leeruitkomst	6.2 Tactical Leadership Identifies and understands individuals in a team and interacts with them in the role of leader at the tactical level. <u>Indicators:</u> The student identifies and explains the team dynamics of an existing team. The student interacts with team members in the role of leader at the tactical level.
---------------------	---

Is er rekening gehouden met de eerder genoemde twee uitgangspunten? Ja. In de opdrachtbeschrijving (zie bijlage 1) werd de eerste indicator van Leeruitkomst 6.2 letterlijk overgenomen:

a). Identify and explain the team dynamics in an existing team of your choice. For example at your job or your design challenge group. b). use one of the tools/models (e.g. Mapstell or the levels of maturity in a team) we provided in class or choose your own tool or model from literature.

In de derde en vierde stap van de opdracht wordt vervolgens gevraagd:

c). Explain how you as a tactical manager might act in this situation and why. What interventions would you undertake? Explain for at least one intervention, what the intervention might look like, and why? Support your choices/reasoning with sources. D). Max. 800 words. Use at least academic 2 sources (APA-referencing) to support your reasoning.

Hoewel niet letterlijk overgenomen, wordt duidelijk dat de tweede indicator van Leeruitkomst 6.2 hier wordt getoetst. De opdracht vraagt de student in de rol van leidinggevende te stappen en te beschrijven hoe hij/zij in een dergelijke situatie zou handelen en waarom. Daarnaast wordt de student gevraagd een interventie te ontwerpen, die logischerwijs interactie tussen de teamleider en het team stimuleert. Dit sluit aan bij de eisen voor toetsing op Level 2 binnen de HMSL, waarbij studenten worden uitgedaagd hun opgedane kennis en vaardigheden in de praktijk toe te passen.

Conclusie

Uitgangspunten doel en ontwerp van toetsen op Level 2	- De indicatoren van de toets zijn gerelateerd aan de indicatoren van de Leeruitkomst die bij het niveau en het expertisegebied past. - De toets nodigt studenten uit hun opgedane kennis en vaardigheden daadwerkelijk toe te passen.
---	--

- De toetsing van het niveau is in lijn met wat er in het **beleid** van de HMSL is voorgeschreven over de plaatsing van de toets in het **toetsprogramma** en het daarbij behorende doel.
- Validiteit is één van de kwaliteitscriteria en gaat over: meet de toets wat hij hoort te meten? Het gegeven dat de indicatoren van de toets zijn afgestemd op de indicatoren van de Leeruitkomst biedt een belangrijke basis voor de **validiteit** van de toets. De toets is immers bedoeld om de leeruitkomsten te toetsen en als bewijsstuk te dienen dat de student aan de leeruitkomst

voldoet. Als de inhoud van de toets nauw aansluit bij de leeruitkomst, komt dit de validiteit ten goede.

- Bovenstaande punten weerspiegelen ook het principe van **constructive alignment** waarbij de leeruitkomsten, de leeractiviteiten en de toetsing nauw op elkaar moeten zijn afgestemd (Last & Jongen, 2021).

Fase 2: vastgelegde afspraken

In fase twee van de toetscyclus maak je afspraken, of kijk je of er afspraken gemaakt zijn die de standaardisering van het toetsen kunnen bevorderen. In dit hoofdstuk wordt weergegeven welke afspraken er gemaakt zijn rondom de opdracht.

Toetsafnamecondities, hulpmiddelen, randvoorwaarden.

De opdracht is een groepsopdracht die studenten thuis uitvoeren. Omdat het een schriftelijke opdracht betreft, zijn er geen vastgestelde toetsafnamecondities, toegestane hulpmiddelen of andere randvoorwaarden. Er is ook geen specifieke tijdslimiet waarin de opdracht voltooid moet worden, en de inlevermomenten kunnen verschillen per Learning Community (LC). Elke HRM-expert bepaalt namelijk de deadline binnen zijn of haar eigen LC.

Daardoor kan het voorkomen dat een groep in de ene LC meer tijd heeft om de opdracht te maken dan een groep in een andere LC. Studenten worden echter aan het begin van het semester geïnformeerd over de deadline, zodat zij hier zelf hun planning op kunnen afstemmen.

Aantal kansen

Er zijn duidelijke afspraken over het aantal pogingen dat studenten binnen een semester mogen doen: elk groepje heeft recht op twee attempts. Dit draagt bij aan de efficiëntie (=kwaliteitscriterium) van de toets, omdat studenten worden gestimuleerd hun kennis optimaal toe te passen binnen deze beperkte pogingen. Het voorkomt dat ze onbeperkt kunnen blijven proberen en zorgt tegelijkertijd voor een efficiëntere werklust voor de HRM-experts, die hierdoor minder vaak hoeven na te kijken.

Daarnaast bevordert het maximum aantal pogingen de transparantie (=kwaliteitscriterium). Studenten weten precies wat er van hen wordt verwacht en begrijpen dat zij deze twee kansen zorgvuldig moeten benutten om een waardevol bewijsstuk voor hun portfolio te creëren.

Toetsinhoud en niveau

Binnen de HMSL zijn duidelijke afspraken gemaakt over de toetsinhoud en het niveau van low-stake assessments. De opdracht moet een bewijsstuk zijn van Leeruitkomst 6.2 en zowel kennis als toepassing toetsen. Uit de evaluatie in Fase 1 blijkt dat aan deze afspraken wordt voldaan.

Weging, cesuur, puntenverdeling

Studenten kunnen de opdrachtbeschrijving inzien op Blackboard. Ook kunnen zij enige informatie terug vinden over hoe ze uiteindelijk beoordeeld zullen worden:

The rubrics your assignment will be graded on:	
<i>Identification and explanation of team dynamics</i>	<i>Above/at/below level</i>
<i>Analysis using tools and models</i>	<i>Above/at/below level</i>
<i>Comments on effective solutions/strategies</i>	<i>Above/at/below level</i>
<i>Link to course readings, effective use of sources and APA</i>	<i>Above/at/below level</i>

Noot: tabel overgenomen uit de opdrachtomschrijving voor studenten. Zie ook bijlage 1.

De student krijgt voor de gehele opdracht een score op een schaal van 0 tot 100. Omdat het een low-stake assessment is, ontbreekt een duidelijke cesuur. Studenten begrijpen echter wel dat een hogere score de opdracht waardevoller maakt als bewijsstuk voor hun portfolio.

In de tabel is enige informatie beschikbaar voor studenten, maar details over de weging, cesuur en puntenverdeling ontbreken. Deze informatie is pas te vinden door naar de inleverpagina op Blackboard te

gaan en op 'view rubrics' te klikken. Hier worden de prestatieniveaus en weging per onderdeel (criteria 1 t/m 4) weergegeven: Below Level: 5%, At Level: 17,5%, Above Level: 25%. Voor docenten is dit duidelijk, omdat zij de ingeleverde opdrachten in Blackboard openen om feedback te geven. Voor studenten is dit echter minder inzichtelijk, aangezien zij vaak pas op de inleverpagina kijken bij het inleveren van de opdracht.

Aanbeveling	Om de transparantie (=kwaliteitscriterium) te verbeteren, zou het benoemen van de criteria met prestatieniveaus, beschrijvingen per niveau (indicatoren), en een overzicht van de weging in de opdrachtbeschrijving een waardevolle toevoeging zijn.
-------------	--

Afspraken m.b.t. beoordeling

Afspraken over de beoordeling van de toets zijn vrijwel niet vastgelegd. De enige bestaande afspraken zijn dat HRM-experts drie werkweken de tijd hebben om de opdracht na te kijken. Daarnaast is vastgelegd dat de opdracht wordt beoordeeld op vier onderdelen (criteria), die elk worden gescoord als Below, At, of Above Level (prestatieniveaus). Verder bieden de rubrics weinig houvast;

- er zijn geen indicatoren die duidelijk maken wanneer een student below, at of above level krijgt op een onderdeel. Indicatoren zijn beschrijvingen die een bepaald prestatieniveau concreet maken en zorgen dat duidelijk is welke specifieke aspecten eigenlijk beoordeeld worden binnen het onderdeel. Ze maken dus duidelijk wat er wordt verwacht op welk niveau.
- Door het gebrek hieraan moeten docenten zelf tijdens het beoordelen interpretaties maken, wat de efficiëntie (=kwaliteitscriterium) kan verminderen.
- Ook zijn er geen afspraken over het toevoegen van feedback of de omvang daarvan, waardoor docenten steeds opnieuw moeten afwegen of hun feedback voldoende is. Dit kan leiden tot twijfel en extra tijdsinvestering en heeft dus gevolgen voor efficiënt werken.

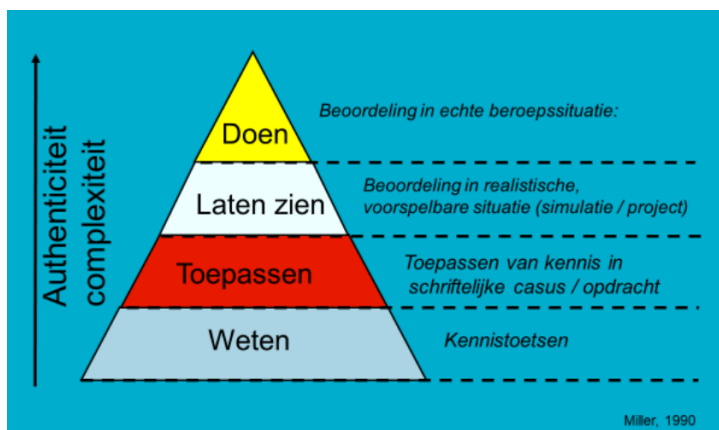
Aanbeveling	Een uitgebreidere instructie/leidraad voor het geven van feedback (of het maken van afspraken hierover) en indicatoren in de rubrics zouden de efficiëntie (=kwaliteitscriterium) en consistentie van de beoordeling kunnen verbeteren.
-------------	---

Fase 3: Constructie

In fase 3 van de toetscyclus construeer je de betreffende toets. Van belang bij het construeren van een toets is dat er niet alleen aandacht wordt besteed aan de toets zelf, maar ook aan de documenten die daarbij horen (Hogeschool Utrecht, z.d.-a). Omdat de betreffende toets al bestaat wordt de constructie ervan in dit hoofdstuk geëvalueerd.

De gekozen toetsvorm

Voor de betreffende toets is er gekozen voor een schrijfopdracht waarin studenten een zelfgekozen casus behandelen. Voor het bepalen van een geschikte toetsvorm geeft de piramide van Miller inzicht (Hogeschool Utrecht, z.d.-b):



Het gaat om een bestaande toets, waardoor niet meer te achterhalen is of de keuze voor de toetsvorm destijds bewust is gemaakt. Duidelijk is wel dat de huidige toetsvorm (schriftelijke opdracht) volgens de Taxonomie van Miller goed past bij het doel en het niveau dat getoetst wordt. Met andere woorden, de vorm sluit aan bij wat beoordeeld moet worden: het vermogen van de student om kennis daadwerkelijk toe te passen. En een passende toetsvorm is essentieel voor de validiteit van de toets (NHL Stenden, z.d.-b).

De opdrachtschrijving

Op Blackboard vinden de studenten de instructie (zie bijlage 1) voor de betreffende opdracht. Hierin is beschreven:

- hoe studenten de opdracht in moeten leveren
- op welke niveaus de criteria beoordeeld kunnen worden (below, at of above level)
- de opdrachtschrijving

In de praktijk blijkt dat studenten de opdracht over het algemeen goed begrijpen. Er is echter één specifiek onderdeel dat regelmatig vragen oproept, wat wijst op een gebrek aan transparantie over wat precies van de studenten wordt verwacht. Het betreft het dikgedrukte gedeelte in de volgende zin:

*“Explain how you as a tactical manager might act in **this situation** and why. What interventions would you undertake?”*

Studenten vragen vaak wat er bedoeld wordt met ‘deze situatie.’ Ik leg dan uit dat hiermee wordt verwezen naar de geïdentificeerde en beschreven teamdynamiek uit stap a) van de opdracht. Daarnaast moeten ze zich voorstellen dat ze in de rol van teamleider nadenken over welke interventie in deze ‘situatie’ of ‘staat van zijn’ nou behulpzaam zou zijn.

Het herhaaldelijk uitleggen van deze vraag kost tijd en schaadt de efficiëntie (=kwaliteitscriterium) van de toets. Soms merk ik pas bij de beoordeling dat studenten de vraag anders hebben geïnterpreteerd of zelfs genegeerd, wat ten koste gaat van de validiteit (=kwaliteitscriterium) van hun antwoorden. Bovendien raakt het ook de transparantie (=kwaliteitscriterium) als studenten niet goed begrijpen wat er van ze verwacht wordt.

Aanbeveling	Om efficiëntie (=kwaliteitscriterium), validiteit (=kwaliteitscriterium) en transparantie (=kwaliteitscriterium) richting de student te verbeteren moet in de opdrachtschrijving een aanpassing gedaan worden, waarmee voor de studenten duidelijker is wat er precies van ze verwacht wordt in stap c) van de opdracht.
-------------	--

Richtlijnen voor nakijken en feedback

Als we kijken naar de richtlijnen voor het nakijken van de toets, valt opnieuw op dat er weinig houvast is voor de HRM-experts die de opdracht beoordelen. Binnen Blackboard wordt voor elke ingeleverde opdracht binnen de HMSL gebruikgemaakt van een beoordelingsformulier. Dit formulier is als volgt opgebouwd:

- Algemeen vak: Een tekstvak voor ‘*feedback to learner*’, bedoeld om inhoudelijke feedback te geven die bijdraagt aan het leerproces van de student.
- Per opdrachtonderdeel: Een keuzevak waarmee de beoordelaar per onderdeel (criterium) kan aangeven of het onderdeel *above*, *at*, of *below level* is beoordeeld. Onder deze keuze is ruimte voor aanvullende feedback over dat specifieke onderdeel.
- Onder alle aparte onderdelen van de opdracht een weergave van de optelling van de verschillende onderdelen in punten (op een schaal van 1 tot 100) en nog een algemeen feedback vak.

Het beoordelingsformulier maakt duidelijk hoe het eindcijfer wordt berekend, maar bevat geen duidelijke instructies. Hoewel het prettig is dat er ruimte is voor toelichting (zoals waarom een bepaalde beoordeling wordt gegeven), blijft onduidelijk of de beoordelaar geacht wordt alle feedbackboxen in te vullen of hier een keuze in mag maken. Er zijn geen richtlijnen voor de hoeveelheid feedback die wordt verwacht.

Aanbeveling	Een uitgebreidere instructie/leidraad voor het geven van feedback (of het maken van afspraken hierover) zou beoordelaars kunnen helpen om efficiënter (=kwaliteitscriterium) te werken.
-------------	---

Rubric

Daarnaast valt ook in deze stap van de cyclus weer op dat een goed uitgewerkte rubrics ontbreekt. Een rubric is een beoordelingsmatrix die inzicht geeft in wat je van de student wilt lezen, zien of horen en hoe deze prestaties worden gemeten (Hogeschool Utrecht, z.d.-e). Cruciaal daarbij is dat de prestatieniveaus (below level, at level, above level) per criterium in beschrijvingen (indicatoren) worden uitgewerkt. Deze uitwerking ontbreekt bij deze opdracht, wat niet alleen de transparantie (=kwaliteitscriterium) richting studenten belemmert (wat wordt verwacht voor elk niveau?), maar ook de betrouwbaarheid (=kwaliteitscriterium). Zonder uitgewerkte niveaus bestaat immers het risico dat elke beoordelaar zelf bepaalt wat bij welk niveau hoort. Dit kan ertoe leiden dat niet elk product op dezelfde manier en met dezelfde criteria wordt beoordeeld.

Aanbeveling	Om de transparantie (=kwaliteitscriterium) richting de student en de betrouwbaarheid van de beoordeling (=kwaliteitscriterium) te verbeteren, zou het goed zijn de criteria met prestatieniveaus en beschrijvingen per niveau (indicatoren) uit te werken en vast te leggen.
-------------	--

Collegiale screening/ kalibratie

Daarnaast zijn er geen richtlijnen voor collegiale screening of kalibratie, hoewel het team de voordelen hiervan erkent. In de praktijk blijkt het ontzettend lastig om hier voldoende momenten voor te vinden. Dit benadrukt nog eens extra het belang van een goed uitgewerkte rubric en het vooraf maken van duidelijke afspraken en instructies over het geven van feedback.

Aanbeveling	Het gebruik van kalibratie of collegiale screening zou een waardevolle toevoeging kunnen zijn om consistentie en betrouwbaarheid (=kwaliteitscriterium) beter te waarborgen.
-------------	--

Fase 4: Toetsafname en beoordeling

De vierde fase van de toetscyclus betreft de daadwerkelijke uitvoering van de toets en de daadwerkelijke uitvoering van de beoordeling. Bij een toets kan de betrouwbaarheid en validiteit van de toets negatief beïnvloed worden door een aantal factoren, zoals bijvoorbeeld de locatie, omstandigheden en instructies (Hogeschool Utrecht, z.d.-f).

Locatie en omstandigheden

De opdracht wordt niet in een gecontroleerde omgeving, zoals een les of examenhal, gemaakt. Studenten bepalen zelf waar en wanneer ze aan de opdracht werken. Het enige vaste element is de deadline voor het inleveren via Blackboard. Hierdoor is het niet mogelijk om te controleren of de omstandigheden waarin studenten de opdracht maken optimaal waren; Het is immers aan hen zelf om deze te creëren. Wel worden studenten aan het begin van het semester geïnformeerd over de deadlines, waardoor zij voldoende tijd hebben om te plannen, aan de opdracht te werken en deze tijdig in te leveren.

Objectiviteit?

Na het inleveren van de opdracht kan de docent beginnen met beoordelen, waarbij gebruik wordt gemaakt van het beoordelingsformulier dat eerder in fase 3 is beschreven. Objectiviteit is zeer belangrijk om tot een zo betrouwbaar mogelijke beoordeling te komen (E-xamen, z.d.). Hulpmiddelen zoals rubrics kunnen bijdragen aan deze objectiviteit. Echter, zoals al in hoofdstuk 3 is vastgesteld, is de rubrics voor deze opdracht niet optimaal uitgewerkt. Dit heeft een negatieve invloed op de betrouwbaarheid (=kwaliteitscriterium) van de beoordeling, vooral op de interbeoordelaarsbetrouwbaarheid. Het verbeteren van de rubrics is dan ook essentieel.

Aanbeveling	Om de (interbeoordelaars)betrouwbaarheid te vergroten, is het belangrijk om de rubrics van de opdracht te verbeteren door indicatoren op te nemen die zijn afgestemd en goedgekeurd door de beoordelaars.
-------------	---

Beoordelaarseffecten

Maar, hoe goed de rubrics ook zijn, bij beoordelen van toetsen zal het altijd belangrijk blijven om rekening te houden met zogenaamde beoordelaarseffecten. Dit zijn invloeden op het oordeel die niets te maken hebben met de kwaliteit van het werk of de presentaties van de student (Hogeschool Utrecht, z.d.-d). Om deze effecten zo klein mogelijk te maken is het belangrijk bewustwording te creëren. Een van de manieren om dat te doen is te kalibreren. Daarbij wordt een opdracht door alle beoordelaars vooraf hardopdenkend beoordeeld (voorbesproken) of worden de beoordelingen achteraf besproken en vergeleken. Door dit te doen kan de interbeoordelaarsbetrouwbaarheid mogelijk verhoogd worden. In hoofdstuk 3 is al kort genoemd dat er in de praktijk geen kalibratie plaatsvindt. Dit benadrukt de noodzaak om kalibratie structureel in te voeren om de consistentie en betrouwbaarheid van beoordelingen te verbeteren.

Aanbeveling	Het gebruik van kalibratie of collegiale screening zou een waardevolle toevoeging kunnen zijn om consistentie en betrouwbaarheid (=kwaliteitscriterium) beter te waarborgen.
-------------	--

Plagiaatchecker

Bij het nakijken van de opdracht kan de docent middels een plagiaatchecker in Blackboard kijken of er plagiaat zou zijn gepleegd. Ook kan er (bijvoorbeeld) gebruik worden gemaakt van Zerogpt.com om te controleren of de tekst niet geschreven is door AI. Het gebruik van deze tools kan bijdragen aan de validiteit (=kwaliteitscriterium) van de toets. De opdracht moet immers 'meten' in hoeverre de student aan de uiteindelijke leeruitkomsten voldoet. Als er sprake zou zijn van plagiaat of gebruik van AI, meet je niet wat je beoogt te meten (de kennis en de vaardigheden van de student). Het is belangrijk dit boven water te halen en de student hierop te wijzen.

Fase 5: Toetsanalyse

De vijfde fase van de toetscyclus betreft het analyseren van de toets.

De woorden *betrouwbaar*, *valide*, *transparant* en *efficiënt* kwamen al eerder voorbij in de voorafgaande hoofdstukken. Deze vier termen samen vormen de kwaliteitscriteria van een toets (NHL Stenden, z.d.-b). Met een goede toets, waarbij (zo goed mogelijk) voldaan wordt aan deze kwaliteitscriteria, kan je deugdelijke beslissingen nemen (NHL Stenden, z.d.-c). In hoeverre een toets voldoet aan de kwaliteitscriteria kan op een kwalitatieve en/of een kwantitatieve manier worden beoordeeld. Bij voorkeur vindt er een combinatie van beide manieren plaats (NHL Stenden, z.d.-c). In dit hoofdstuk wordt eerst de kwalitatieve analyse, en dan de kwantitatieve analyse besproken.

Kwalitatieve analyse:

De vier kwaliteitscriteria

Bij “een kwalitatieve analyse kijk je naar het hele toetsproces. Op basis van de informatie uit alle stappen uit de toetscyclus doe je uitspraken over de rechtvaardigheid van de toetsresultaten. Je let daarbij op de vier kwaliteitscriteria: validiteit, betrouwbaarheid, transparantie en efficiëntie”(NHL Stenden, z.d.-b). De eerste hoofdstukken van dit document zijn al onderdeel van deze analyse. Die bevindingen worden in onderstaande paragraaf daarom per kwaliteitscriterium samengevat en waar nodig aangevuld.

Validiteit

“ Een valide toets “meet wat gemeten moet worden”. Dit betekent dat de toets de juiste informatie oplevert om te kunnen bepalen of een student de leerdoelen heeft behaald” (NHL Stenden, z.d.-b).

Tijdens het doen van de kwalitatieve analyse viel m.b.t. validiteit het volgende op:

- De toets sluit aan bij de fase van de opleiding qua gevraagde autonomie en complexiteit. De toetsvorm (schrijfofdracht) is passend bij het niveau en het doel van de toets. Dit draagt bij aan de validiteit omdat de toets meet of studenten hun kennis en vaardigheden kunnen toepassen (en dat is het geen we willen beoordelen).
- De indicatoren van de toets zijn gerelateerd aan de indicatoren van de leeruitkomst. Dit biedt een belangrijke basis voor de validiteit van de toets. De toets is immers bedoeld om de leeruitkomsten te toetsen en als bewijsstuk te dienen dat de student aan de leeruitkomst voldoet. Als de inhoud van de toets nauw aansluit bij de leeruitkomst, komt dit de validiteit ten goede.
- De toetsvorm (schriftelijke opdracht) past goed bij het niveau en het doel. Met andere woorden, de vorm sluit aan bij wat beoordeeld moet worden: het vermogen van de student om kennis daadwerkelijk toe te passen. Een passende toetsvorm is essentieel voor de validiteit van de toets.
- De mogelijkheid om te checken op plagiaat en AI-gebruik (met tools als de in blackboard ingebouwde plagiaatchecker en de website ZeroChatGPT.nl) draagt bij aan de validiteit. Het zorgt er namelijk voor dat je weet of de toets inderdaad meet wat het hoort te meten (de kennis en vaardigheden van de specifieke student, en niet die van een andere student of een programma).
- De validiteit wordt bedreigd wanneer studenten delen van de opdracht, zoals de term ‘*deze situatie*,’ verkeerd interpreteren. Dit probleem kan worden aangepakt door de opdrachtschrijving te verduidelijken.

Betrouwbaarheid

“Een betrouwbare toets betekent dat het toetsresultaat niet afhankelijk is van allerlei irrelevante omstandigheden. Denk aan de “strengheid” van een bepaalde docent, de interpretatie van beoordelingscriteria door verschillende docenten en omstandigheden tijdens de toetsafname.” (NHL Stenden, z.d.-b).

Tijdens het doen van de kwalitatieve analyse viel m.b.t. betrouwbaarheid het volgende op:

- De betrouwbaarheid van de beoordeling van de producten wordt beïnvloed door het gebrek aan standaardisatie tussen beoordelaars. Er zijn namelijk géén indicatoren in de rubrics opgenomen die duidelijk maken wanneer een onderdeel op welk prestatieniveau wordt beoordeeld. Het risico is dat elke beoordelaar zelf bepaalt welke prestatie hij op welk niveau verwacht.

- Er is geen kalibratieproces waarin beoordelaars kunnen afstemmen. Een kalibratieproces zou tot meer bewustwording kunnen leiden en zo het risico op het ontstaan van beoordelaarseffecten mogelijk kunnen verminderen.

Vergelijking feedback beoordelaar 1 en beoordelaar 2

In het kader van betrouwbaarheid vond ik het interessant om te onderzoeken in hoeverre de beoordeling en bijbehorende feedback van beoordelaars overeenkomen of verschillen. In bijlage 2 is voor drie opdrachten de verstrekte feedback weergegeven van zowel LC2 (dit is mijn eigen feedback) als LC1. Het is belangrijk om te vermelden dat de feedback niet één op één vergelijkbaar is, omdat de opdrachten door verschillende groepen zijn gemaakt. Met andere woorden, dezelfde opdracht is niet door beide beoordelaars beoordeeld. Door echter de verstrekte feedback naast elkaar te leggen voor drie typen opdrachten (een laag gescoorde opdracht, een nét voldoende opdracht en een hoog gescoorde opdracht), heb ik gekeken naar algemene verschillen in de manier waarop de twee beoordelaars feedback geven. Hierbij vielen de volgende punten op:

- Concreetheid en lengte van feedback: De feedback van de beoordelaar van LC1 is concreter en bondiger, terwijl de feedback van LC2 langer is en meer uitleg en aanwijzingen voor verbetering bevat.
- Focus en duidelijkheid van scores: Door de bondigheid van LC1 is het soms minder duidelijk op basis waarvan een bepaalde score (below, at of above level) wordt toegekend, in vergelijking met LC2. Daarnaast lijkt de beoordelaar van LC2 meer nadruk te leggen op theoretische onderbouwing (naast praktische toepassing), terwijl LC1 vooral focust op basisbegrip en praktische toepassing.
- Suggesties en verbeterpunten: De beoordelaar van LC2 geeft vaker suggesties voor verbetering en tips over waar studenten meer informatie kunnen vinden. Bijvoorbeeld, bij het toepassen van APA-richtlijnen legt LC2 uit waarom dit belangrijk is, terwijl LC1 alleen aangeeft of iets voldoende is toegepast, zonder verdere specificatie of toelichting.

Hoewel het vanwege de onmogelijkheid om een één-op-één vergelijking te maken lastig is om een officiële conclusie te trekken over de interbeoordelaarsbetrouwbaarheid tussen LC1 en LC2, wijzen bovenstaande bevindingen er wel op dat de beoordelaars mogelijk verschillende accenten leggen in hun beoordeling of andere aspecten belangrijker vinden.

- LC2 lijkt meer nadruk te leggen op theoretische onderbouwing, wat kan resulteren in een hogere score wanneer een opdracht op dat gebied sterk is.
- LC1 lijkt daarentegen meer gericht op praktische toepassing en basisbegrip, waardoor studenten die dat goed uitwerken mogelijk beter scoren bij deze beoordelaar.

Aanbeveling	Bovenstaande bevindingen onderstrepen nogmaals de noodzaak van duidelijkere rubrics en betere kalibratie tussen beoordelaars. Dit kan bijdragen aan een meer uniforme beoordeling en consistente feedback voor studenten, ongeacht wie de beoordeling uitvoert.
-------------	---

Transparantie:

*“Krijgen de studenten duidelijke informatie over de test, de criteria, normen, cesuur en resultaten?”
(NHL Stenden, z.d.-b)*

- In de opdrachtomschrijving op Blackboard kunnen studenten terugvinden op welke onderdelen (criteria) ze beoordeeld worden en op welke prestatieniveaus. Dit zorgt voor transparantie.

- Het afspreken van een maximaal aantal ‘attempts’ binnen een semester zorgt voor transparantie naar de student toe. De verwachtingen zijn duidelijk en ze weten waar ze aan toe zijn.
- In de TER is vastgelegd wanneer studenten hun resultaten terug mogen verwachten.
- Studenten kunnen op Blackboard een duidelijke opdrachtomschrijving vinden. In de praktijk blijkt echter wel dat er één onderdeel regelmatig vragen oproept: “*Explain how you as a tactical manager might act in **this situation** and why. What interventions would you undertake?*” Daar blijkt de transparantie dus niet optimaal.
- In de opdrachtomschrijving op Blackboard kunnen studenten geen informatie terugvinden over de cesuur of de weging van de verschillende onderdelen, en er zijn geen indicatoren per criterium opgenomen. Het zou de transparantie richting de student ten goede komen als dit ook in de opdrachtomschrijving wordt opgenomen.

Efficiëntie

“De kwaliteit van de toets is mede afhankelijk van een aantal randvoorwaarden en de afname van de toets in de praktijk. Is de toets efficiënt en effectief georganiseerd? Hebben beoordelaars voldoende tijd om tot een betrouwbaar oordeel te komen?” (NHL Stenden, z.d.-b).

- Het beperken van het aantal attempts draagt bij aan de efficiëntie van de toets.
- Het herhaaldelijk moeten uitleggen van één van de onderdelen van de opdrachtomschrijving is inefficiënt.
- Het ontbreken van duidelijke beoordelingsrichtlijnen (indicatoren in de rubrics) kan leiden tot inefficiënt werken. De docent moet steeds zelf bedenken waar hij of zij op wil beoordelen en welke weging van belang is.
- Het ontbreken van een richtlijn m.b.t. het geven van feedback kan leiden tot inefficiëntie. Mogelijkerwijs kan het ertoe leiden dat de docent tijdens het nakijken gaat twijfelen over hoeveel feedback nodig is en hier al zoekende veel tijd aan kwijt is.

Kwantitatieve analyse

Bij een kwantitatieve analyse kijk je naar de toetsresultaten. Door dit te doen kan je zien hoe studenten hebben gepresteerd en inschatten of er eventuele aanpassingen nodig zijn. Ik vond het lastig om te bedenken wat ik bij de kwantitatieve analyse nou precies kon doen; het betreft een formatieve toetsing, waarbij de scores niet direct worden gebruikt om prestaties summatief te beoordelen. Daardoor ligt de focus meer op de kwalitatieve feedback die studenten krijgen. Bovendien betreft het een schrijfopdracht, die toch subjectiever van aard is, en waarop dus ook hele verschillende antwoorden worden gegeven. Dat maakt het lastig te vergelijken. Het zou interessant zijn om te kijken naar de verschillende scores op niveau van criteria, maar deze data is niet uit blackboard te halen. Al met al was het lastig om kwantitatieve data voor een analyse te vinden. In gesprek met onze atelier begeleider van beoordelen kwam ik er ook achter dat het niet nodig is om een P-waarde, R-waarde of Cronbachs alfa te berekenen, en dat het goed is om met name op de kwalitatieve analyse te focussen. Maar ze gaf ook aan dat het wel inzichtelijk kan zijn om bij dit soort opdrachten toch even te kijken naar de toetsscores en naar beoordelaarsvergelijkingen.

Beoordelaarsvergelijking en zak-slaag grens bij deze formatieve toets?

Een goede (kwantitatieve) beoordelaarsvergelijking maken kan niet. Hoewel de scores enige trends in beoordelingsstijl van beoordelaars kunnen laten zien (zie hiervoor de vergelijking van beoordelaars onder betrouwbaarheid bij de kwantitatieve analyse), kan een echte vergelijking tussen beoordelaars alleen worden gemaakt als dezelfde producten door meerdere beoordelaars worden beoordeeld. En dat is nog niet gedaan. Het zou interessant zijn om dit in het komende semester een keer op te pakken in een

kallibratiesessie.

Hoe zit dat met aan analyse van de zak-slaaggrens en toetsscores? De betreffende opdracht wordt zoals eerder beschreven in groepjes gemaakt en ingeleverd via Blackboard. De docent leest de opdracht via Blackboard, voegt feedback toe, voorziet de opdracht van een cijfer (toetsscore) en vervolgens wordt de opdracht vrijgegeven. Al eerder genoemd, maar wellicht goed om te herhalen is, dat het, ondanks toekenning van een toetsscore om een formatieve toets gaat. De betreffende opdracht (inclusief feedback en toetsscore) kan desgewenst door studenten toegevoegd worden aan het portfolio. Óók als het een toetsscore lager dan 5.5 betreft. De toetsscores zeggen in die zin dus niets over het uiteindelijk wél of níét slagen van de student op de bijbehorende leeruitkomst (er is geen sprake van een officiële cesuur).

Het geeft echter wel een indicatie, of beter gezegd, het is een mogelijke voorspeller voor hoe goed de student de leeruitkomst in de vingers heeft. Tijdens het voorbereiden van het portfolio interview bekijken de portfolio assessoren de datapunten. Zo ook deze opdracht. Hoewel studenten datapunten mogen toevoegen die onder de 5.5 beoordeeld zijn, werkt het in de praktijk vaak wel zo dat een hogere score (en in ieder geval een score die als ‘voldoende’ gerekend wordt) door de assessor als waardevoller wordt beoordeeld. Het is in wezen een ‘informele drempel.’ Dit maakt de toetsscores, en de ‘informele drempel’ van 5.5 ondanks het formatieve karakter van de toets, dus toch relevant. Ze bieden inzicht in hoe goed een student de leeruitkomst beheerst.

Ik wil hier echter nog wel een nuancerende opmerking bij maken. Hoewel ik hierboven spreek over ‘de student’ is het belangrijk te beseffen dat de student individueel beoordeeld wordt op de leeruitkomsten (d.m.v. zijn/haar portfolio en portfoliogesprek), maar de weg ernaar toe is (veelal) groepswork is. Dat maakt het vaak onduidelijk wie welk deel van het werk heeft gedaan en lastig te beoordelen hoe elk groepslid de leeruitkomst individueel beheerst. De toetsscores geven dus eigenlijk weer hoe de groep als geheel de leeruitkomst begrijpt.

Omdat er géén resultaat van deze toets in Progress wordt gezet, kan er geen gebruik gemaakt worden van analysetools uit Progress. Wel is er de mogelijkheid om rapporten uit Blackboard te draaien. De mogelijkheden in deze zijn beperkt. Zo kan er bijvoorbeeld wel een weergave van de totaalscore worden gegeven, maar geen score per toetsitem (per criterium).

Om inzicht te krijgen in een (mogelijk) verschil tussen beoordelaars zijn (op het moment van draaien in juni 2024) voor de afgelopen 6 maanden zowel de resultaten van LC 2 (waar ik de beoordelaar ben) als LC1 (waar mijn collega Mariët de beoordelaar is) uitgedraaid. Het ‘totaal aantal’ betreft het totale aantal studenten. Goed om te beseffen is dat deze in groepen ingedeeld waren. Uit Blackboard kan je echter alleen individuele scores draaien.

<i>Eerste inlevermogelijkheid</i>	Totaal aantal	Laagste score	Hoogste score	Gemiddeld
Totaal aantal LC2	54	20.0 (5)	85.0 (5)	63.10
Totaal aantal LC1	29	45.0 (5)	85.0 (5)	64.48
<i>Tweede inlevermogelijkheid</i>	Totaal aantal	Laagste Score	Hoogste Score	Gemiddeld
Totaal aantal LC2	8	52.5 (4)	57.5 (4)	55.00
Totaal aantal LC1	10	57.5 (5)	70.0 (5)	63.75

Wat valt op? De gemiddelde score van LC1 en LC2 liggen vlak bij elkaar bij de eerste inlevermogelijkheid. LC2 heeft echter een veel breder bereik dan LC1. Hoewel het kan liggen aan het verschil in kwaliteit van de ingeleverde opdrachten, wijst het er mogelijkwijs op dat de beoordelaars de criteria verschillend interpreteren en een andere mate van striktheid aanhouden. LC2 lijkt over het algemeen strenger te beoordelen. Met het trekken van harde conclusies op basis van deze tabel moeten we voorzichtig zijn. De tabel lijkt echter wel te onderschrijven dat kalibratie en het uitwerken van indicatoren in de rubrics nodig zijn.

Fase 6: Feedback en doorgeven toetsresultaat

In Fase 6 van de cyclus kijk je naar het geven van feedback en het registreren van de resultaten. Wat is hier over afgesproken?

Bij de HMSL is in artikel 6.12 van de TER vastgelegd dat studenten binnen drie weken hun resultaten moeten terugkrijgen. Met een week wordt bedoeld een kalenderweek, met uitzondering van de schoolvakanties zoals die in het jaarplan van NHL Stenden zijn vastgelegd. Voor studenten biedt dit transparantie. Zij weten wanneer ze hun resultaten mogen terugverwachten. In het geval van deze opdracht wordt het resultaat via Blackboard vrijgegeven en van feedback voorzien. Het resultaat wordt niet in Progress opgenomen omdat het geen summatieve toets betreft.

Een formatieve toets, zoals de betreffende opdracht, is een belangrijk moment in het leerproces van de student. Je wil immers duidelijk maken waar de student staat ten opzichte van de leeruitkomst die straks tijdens het summatieve afgetoetst gaat worden (Hogeschool Utrecht, z.d.-g). Feedback is dan erg belangrijk en een goed beoordelingsformulier is daarbij een goede leidraad.

Zoals eerder aangegeven is er wat dat betreft verbetering nodig. Rubrics met duidelijk beschreven en gewaardeerde indicatoren zijn nodig om ervoor kunnen zorgen dat studenten beter inzicht krijgen in waarom ze welke feedback hebben gekregen en waarom er een bepaalde weging aan vast hangt. Tevens zou dit het werk van de beoordelaar efficiënter maken; die heeft dan wellicht minder woorden nodig om feedback te verzorgen.

Aanbeveling	Om de transparantie (=kwaliteitscriterium) richting de student te verbeteren, en het werk van de beoordelaar efficiënter te maken (=kwaliteitscriterium) zou het goed zijn de criteria met prestatieniveaus en beschrijvingen per niveau (indicatoren) uit te werken en vast te leggen.
-------------	---

Fase 7: Evalueren

In Fase 7 van de cyclus valueer je als docent(enteam) de gehele toetscyclus. Je kijkt naar wat er goed ging en wat er beter kan (Radboud Universiteit, z.d.)

Op dit moment vindt er géén officiële kallibratie en/of evaluatie plaats tussen de verschillende docenten die met de betreffende opdracht werken. Ook wordt er niet geëvalueerd onder studenten. De toetscyclus die in dit document doorlopen wordt zorgt voor zicht op verbetermogelijkheden en een aantal concrete verbeteracties (zie volgend hoofdstuk). Dat is een eerste stap. Voor de toekomst zou het echter goed zijn om een evaluatie tussen beoordelaars en een evaluatie onder studenten uit te vragen. Het kan mogelijk nog andere belangrijke inzichten geven over de kwaliteit van de opdracht (als geheel) en mogelijk nog extra input geven voor het (her)ontwerp van de opdracht. Het doorlopen van het toetsproces is immers cyclisch (UVA, z.d.) en zou eigenlijk continu opnieuw moeten plaatsvinden. Het is dan ook aan te bevelen om in te zetten op evaluaties door studenten en de betrokken docenten.

Aanbeveling	Om de kwaliteit van de toets en het toetsproces te blijven waarborgen is het aan te bevelen om na elk semester te evalueren, zowel met beoordelaars, als met studenten.
-------------	---

Conclusie en geformuleerde verbeteracties

In dit document is de toetscyclus doorlopen. Het doorlopen van de verschillende fasen heeft bijgedragen aan het onderzoeken van de kwaliteit van zowel de toetsing als het toetsproces. Daarnaast hebben de analyses geholpen bij het formuleren van aanbevelingen om de kwaliteit verder te verbeteren. Deze heb ik vertaald naar concrete acties:

1. De bestaande rubrics moet verder worden uitgewerkt door indicatoren toe te voegen en deze af te stemmen met en goed te laten keuren door mede beoordelaars. Na afstemming moeten deze worden opgenomen in de opdrachtinstructie, zodat ze helder zijn voor studenten, en in de rubrics, zodat ze beschikbaar zijn tijdens het beoordelen.
2. Er moet een overzicht van de weging van de verschillende onderdelen van de opdracht (criteria) gemaakt worden en dit moet toegevoegd worden aan de opdrachtoomschrijving.
3. Er moet met medebeoordelaars gesproken worden om richtlijnen voor het geven van feedback af te stemmen. Idealiter rolt hier een leidraad uit die we kunnen gebruiken bij het nakijken.
4. Stap c) van de opdrachtoomschrijving voor studenten moet worden aangepast, om specifieker te maken wat er van studenten wordt verwacht.
5. Er moet besproken worden hoe we kalibratiesessies of collegiale screenings binnen ons HRM expert team gaan organiseren.
6. Er moet besproken worden hoe we ons toetsproces elk semester kunnen gaan evalueren onder zowel studenten als docenten.

Hoewel er verschillende verbeteracties zijn geformuleerd richt ik mij eerst op het uitwerken van de indicatoren, in verband met haalbaarheid en de directe impact die deze actie kan hebben op meerdere kwaliteitscriteria. De overige acties hoop ik na het afronden van mijn BDB, in het tweede semester van AY24/25 (samen met mijn vakgroepgenoten) op te kunnen pakken. Het verbeteren van de rubrics door middel van het uitwerken van de indicatoren is in mijn portfolio in een apart document uitgewerkt.

Bijlages

Bijlage 1 Assignment EHL HRM AY 23-24

Professional assignment:

- a. identify and explain the team dynamics in an existing team of your choice. For example at your job or your design challenge group.
- b. use one of the tools/models (e.g. Mapstell or the levels of maturity in a team) we provided in class or choose your own tool or model from literature.
- c. Explain how you as a tactical manager might act in this situation and why. What interventions would you undertake? Explain for at least one intervention, what the intervention might look like, and why? Support your choices/reasoning with sources.
- d. Max. 800 words. Use at least academic 2 sources (APA-referencing) to support your reasoning.

The rubrics your assignment will be graded on:

<i>Identification and explanation of team dynamics</i>	<i>Above/at/below level</i>
<i>Analysis using tools and models</i>	<i>Above/at/below level</i>
<i>Comments on effective solutions/strategies</i>	<i>Above/at/below level</i>
<i>Link to course readings, effective use of sources and APA</i>	<i>Above/at/below level</i>

Bijlage 2: feedback Beoordelaar LC2 en Beoordelaar LC1 voor 3 opdrachten (below, at & above level)

LC 2 Groepje EHL2B1	LC1 EHL1A2
Score 20/100	45.00
Feedback to Learner (algemeen) Dear all, thank you for your HRM assignment. Unfortunately it is not sufficient. I have provided you with both in-text feedback as feedback in the rubrics. Unfortunately you have chosen not to hand in a first attempt earlier. This means you can not work on the feedback and try another time. I do hope that you will take the feedback into account and use it for future work. For the benefit of this: please especially take a second close look to your reference list and the feedback there. If you have any questions, please let me know. I wish you a good christmas holiday. Romée, I hope to see you again in class and to the others: I hope to meet you there! Best regards, Y	Feedback to Learner (Algemeen) Dear students, Thank you for handing in a first attempt. This assignment is not at level yet. I invite you to hand in a new attempt. Please let me know via email when you have handed it in, so I know there is something there for me to grade. Kind regards X
Identification and explanation of team dynamics Below Level You do talk about things that indeed belong to team dynamics (communication, interaction), but you didn't explain or introduce the topic of team dynamics in general? what is it? What are you talking about? Please elaborate and then make a connection with your own team.	Identification and explanation of team dynamics At level You share what your team looks like and who does what effectively. But is that team dynamics? please look up the definition of team dynamics in an academic source and reconsider your description.
Analysis using tools and models Below Level You chose Mapstell to explain the dynamics within your group. You made a start but didn't elaborate enough on Mapstell and how it helps you think about team dynamics (you only named colors). You could ask yourself questions like: What do these colors say about team dynamics? About the behavior I see? How does it work out? Are we missing a certain color/behavior in our group? Use references to back up your analysis.	Analysis using tools and models Below Level I see that you have included the Mapstell profiles to the assignment. However, you do not explain or apply it anywhere. E.g. in your tactical solution you mention the importance of communication. Would you for instance communicate differently with somebody from the blue or the yellow area? And how is Mapstell an explanation for what you see in the group dynamics? Or did yo perhaps use Mapstell to divide the tasks in the team? How then did you divide them and why?
Comments on Effective Solutions /Strategies	Comments on Effective Solutions /Strategies At Level

Below Level You have found an interesting example where indeed the behavior of one can impact the group dynamics as a whole. I don't think you have dived deep enough in theory about Belbin to understand how you can use it/to make the right connection with your problem. While I am not saying that choosing a role for one person (in this case Pelle) absolutely can not work out, it is not obvious to do this (and also so far that I know of not obvious according to Belbin). According to Belbin's theory, individuals have a natural preference for one or more roles. Roles can evolve over time and in different situations, but I would say that the role of compelter-finsiher and team worker probably isn't Pelle's preferred role. This would probably mean that he would have a lot of struggles trying to adapt to this role. Is that working to a solution? If you do believe this could work, please provide with explanation where you use theory about Belbin as reference so that I can check.	The description of a tactical action is sufficient, the survey (used to start up feedback in the team) can be effective. The writing is a little high over. I would have liked a more elaborate description, e.g. not only mentioning the survey but also sharing if in the end it worked for you or not, why and how did it (not) work?
Links to Course Readings or sources and Additional Research, links to learning outcomes Below Level You have tried to find sources to back up your analysis and solutions/strategy but I don't think the sources you provided have given you the right understanding of how Belbin works. Also, your reference list is a mess and therefore I am not able to find everything and check. Please provide correct APA referencing next time.	Links to Course Readings or sources and Additional Research, links to learning outcomes Below Level No in-text references. The sources are not academic.

Vergelijking feedback Beoordelaar LC2 en Beoordelaar LC1 m.b.t. de Teamdynamics opdracht voor HRM

LC 2 Groepje EHL2A3	LC1 EHL1B4
Score 57.5/100	57.5/100
Feedback to Learner (algemeen) Dear Students, thank you for your assignment for HRM. It is just at level. Your part about the interventions was quite strong and I want to compliment you with that. But there is one big problem that I want to address, and I strongly advise you to work on this, as you will need it when you are entering the next level of your education. It is important for academic integrity and in avoiding plagiarism: Your use of APA referencing is not correct. You quote other articles, but don't use (" ") double quotes. Instead you make it seem like you are	Feedback to Learner (Algemeen) Dear students, Thank you for handing in this assignment. Overall, the assignment is just at level. Even though you identify the model of Tuckman, the actual application is lacking, missing depth. Please find more feedback in the rubrics above. Kind regards X

paraphrasing. Being able to distinguish between quoting and paraphrasing in APA style is really important for several things. For example integrity, but also for clarity and showing comprehension. I would advise you to practice more with proper paraphrasing, as it shows that you can interpret source material and put it in your own words. If you have any questions, please reach out to me. I wish you all the best in finalizing your semester! Best regards, Y	
Identification and explanation of team dynamics At level You have given definitions of team dynamics and have explained how communication and conflict resolving works in your group. You have shared what your team looks like and who does what.	Identification and explanation of team dynamics At Level <u>Geen feedback toegevoegd</u>
Analysis using tools and models At Level You have used Tuckmans model to explain the phases of team dynamics you went through as a team. It shows that you have basic understanding of these phases.	Analysis using tools and models Below Level You have chosen a model, but you have not actively applied it. So, what behaviors might you expect to see in a stage (model Tuckman) and what did you see in the group that you analyzed. To what conclusion does that lead you on where they are developmentally?
Comments on Effective Solutions /Strategies At level You have sufficiently used Belbins team roles theory to explain how you want to build appropriate solutions for the problem that sometimes occurs within your team. I like the fact that you think you can link it to higher motivation and involvement, but you have written it as a statement. Can you back this up with a relevant source?	Comments on Effective Solutions /Strategies At level Well chosen intervention.
Links to Course Readings or sources and Additional Research, links to learning outcomes Below Level You have used sources but you unfortunately make a lot of mistakes in APA. Be aware of the	Links to Course Readings or sources and Additional Research, links to learning outcomes At Level Sufficiently linked to sources.

fact that you are mostly quoting, which is using exact the same text of someone else. If you do this, you have to show it and give the credits to the original author.	
--	--

Vergelijking feedback Beoordelaar LC2 en Beoordelaar LC1 m.b.t. de Teamdynamics opdracht voor HRM

LC 2 Groepje EHL2A4	LC1 EHL1A3
Score 85/100	Score 85/100
Feedback to Learner (algemeen) Dear students, thank you for your HRM assignment. It was really interesting to read and I really appreciate the fact that you did great research on the topic of team dynamics. For your learning, please read the feedback that was provided in the rubrics. Well done! Y	Feedback to Learner (Algemeen) Dear students, Thank you for handing in this assignment. It is a job well done, above level. You show clear insight in what team dynamics are, you choose and apply a relYnt model to make sense of what is happening in a real life situation and you are able to come with a effective tactical plan to address the issues in the team dynamics Kind regards X
Identification and explanation of team dynamics Above level You have shared insights into your team and have used Mapstell to help you with thinking about it. You have given a definition of team dynamics, although I think you have missed something there: you have mentioned why it is so important, but forgot to mention what are the elements of team dynamics? I really appreciate the fact that you have also given background about Mapstell by including history of DISC model. You have put a lot of effort into exploring what team dynamics is about.	Identification and explanation of team dynamics At Level Sufficiently explained.
Analysis using tools and models At Level You have efficiently used Mapstell to help you analyse your team. It could have been a bit more elaborative. For example, you write about what behaviour you can expect (Mapstell). Do you see this in real life as well? Do problems occur?	Analysis using tools and models Above Level Well done on finding, selecting and actively applying a model that is relevant and that was not discussed in class. Good to see you take initiative.
Comments on Effective Solutions /Strategies	Comments on Effective Solutions /Strategies
	Above Level

At level You have not mentioned a specific problem occurring, but you have explained what communication style the different team members need from their team leader in order to perform, which is indeed very important for a positive and productive team environment and in shaping team dynamics. It was interesting to read.	Well documented and thought out example of what you would do if you were the tactical manager. Your approach makes sense and is feasible. Reasoning well supported with research.
Links to Course Readings or sources and Additional Research, links to learning outcomes Above Level You have done great in finding relevant academic sources and applying APA style efficiently.	Links to Course Readings or sources and Additional Research, links to learning outcomes At Level Well linked. Please also include an APA reference in the bibliography for the quotes used from class slides.